



This is a repository copy of *Does early child language predict internalizing symptoms in adolescence? An investigation in two birth cohorts born 30 years apart.*

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/172544/>

Version: Published Version

Article:

Thornton, E., Patalay, P., Matthews, D. orcid.org/0000-0003-3562-9549 et al. (1 more author) (2021) Does early child language predict internalizing symptoms in adolescence? An investigation in two birth cohorts born 30 years apart. *Child Development*, 92 (5). pp. 2106-2127. ISSN 0009-3920

<https://doi.org/10.1111/cdev.13615>

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 including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Does Early Child Language Predict Internalizing Symptoms in Adolescence? An Investigation in Two Birth Cohorts Born 30 Years Apart

Emma Thornton 
University of Liverpool

Praveetha Patalay 
University College London

Danielle Matthews 
University of Sheffield

Colin Bannard 
University of Manchester

Language is vital for social interaction, leading some to suggest early linguistic ability paves the way for good adolescent mental health. The relation between age-5 vocabulary and adolescent internalizing symptoms was examined in two U.K. birth cohorts that are nationally representative in terms of sex, ethnicity, and socioeconomic status: the 1970 British Cohort Study (BCS; $N = 11,640$) and the Millennium Cohort Study (MCS born ~2001; $N = 14,754$). In the BCS, no relation between receptive vocabulary and age-16 self-reported symptoms was observed ($\beta = 0.00 [-0.03; 0.03]$). In the MCS, better expressive vocabulary was associated with *more* age-14 self-reported symptoms ($\beta = 0.05 [0.02; 0.07]$). The direction of this effect was reversed for parent-reported symptoms. All effect sizes were small. The relation between childhood vocabulary and internalizing symptoms varies by generation and reporter.

Early language skills are frequently claimed to be an important contributing factor to later mental health, with calls being made for early language interventions to prevent later mental health problems (Bercow, 2018; Miller, Machlin, McLaughlin, & Sheridan, 2020; Oxford University Press, 2018). Theoretically, language ability could be important for mental health because it is the primary medium for social interaction and because it supports self-regulation (Redmond & Rice, 1998; Salmon, O’Kearney, Reese, & Fortune, 2016). However, while there is some work to suggest poorer mental health outcomes for children with Developmental Language Disorder (DLD; Yew & O’Kearney, 2013), relatively

little empirical work has tested this association in the general population. If an association holds across the continuum of language ability, this would support calls for widespread early language intervention to improve adolescent mental health.

This article focuses on adolescent internalizing symptoms (emotional difficulties), which include symptoms of the most common mental health problems such as anxiety and depression. These stand in contrast to externalizing symptoms (behavioral problems) such as poor impulse control and aggression (Willner, Gatzke-Kopp, & Bray, 2017), which we do not examine in this article. Adolescence is a critical period for susceptibility to internalizing mental ill health (McLaughlin & King, 2015). If early language ability were to have any direct effect on mental health, there is good reason to expect there to be evidence of its impact on internalizing symptoms at this pivotal developmental stage. Since adolescent mental health difficulties (a) predict social exclusion, stigma, poor educational attainment, risky behaviors, and poor physical

We are grateful to the cohort members of the 1970 British Cohort Study, the Millennium Cohort Study, and their families. We are also grateful to the Centre for Longitudinal Studies (CLS), for the data collection and management, and to the UK Data Service for making them available. It must be noted that neither CLS nor the UK Data Service have any responsibility, or took any part in the analysis or interpretation of these data. We also thank Professor Julian Pine (The University of Liverpool) and Dr Vanessa Moulton (CLS, UCL) for their extremely helpful comments on a draft of this manuscript. We also thank Hannah Sawyer and Alice Newton-Fenner for proof reading the final version of this manuscript. Danielle Matthews was supported by the British Academy grant MD\170025. Emma Thornton was funded by the ESRC, NWSSDTP grant number ES/P000665/1.

Correspondence concerning this article should be addressed to Emma Thornton, Department of Psychology, Eleanor Rathbone Building, Bedford Street South, Liverpool L69 7ZA, United Kingdom. Electronic mail may be sent to hlethor2@student.liverpool.ac.uk.

© 2021 The Authors

Child Development published by Wiley Periodicals LLC on behalf of Society for Research in Child Development

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

0009-3920/2021/xxxx-xxxx

DOI: 10.1111/cdev.13615

health (Clayborne, Varin, & Colman, 2019), (b) often persist into adulthood (Fergusson, Horwood, Ridder, & Beautrais, 2005), and (c) are becoming more prevalent (Patalay & Gage, 2019), the prevention of mental ill health in adolescence is a priority (Thapar, Collishaw, Pine, & Thapar, 2012).

Early language skills could underpin adolescent mental health in at least two ways. First, good early language skills, specifically vocabulary and narrative skills, are critical for self-regulation and emotional understanding (see Salmon et al., 2016), which are in turn important for internalizing mental health (Robson, Allen, & Howard, 2020; Trentacosta & Fine, 2010). Second, language facilitates social interaction and is potentially a major determinant of our ability to relate to others and maintain relations with them, which likely supports mental health. This has long been proposed to be important for children with DLD, a developmental disorder where language ability does not fall within the typical range despite otherwise normal development (Bishop et al., 2017). Children with DLD often adapt to the communicative demands of real-world social environments by relying on adults to mediate interactions and by engaging in reduced levels of initiation and assertive negotiation with peers (see the social adaptation model: Redmond & Rice, 1998). Since positive peer interaction is known to be important for adolescent mental health (Thapar et al., 2012), we might expect that a greater degree of language difficulty would put children at risk of later internalizing difficulties.

A number of studies have found that children with DLD when aged 4–7 years are at increased risk of later mental health difficulties when aged 8–19 years, compared with their typically developing peers (Beitchman et al., 2001; Conti-Ramsden & Botting, 2008; Conti-Ramsden, Mok, Pickles, & Durkin, 2013; Wadman, Botting, Durkin, & Conti-Ramsden, 2011; although see Redmond & Rice, 2002; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006 for counterevidence). These findings are hard to interpret since, despite group differences, continuous measures of language ability do not always predict internalizing symptoms in the children studied (Conti-Ramsden & Botting, 2008; Wadman et al., 2011). A meta-analysis conducted by Yew and O’Kearney (2013) noted that children with early language difficulties experience emotional problems of increased severity and frequency compared to their typically developing counterparts. However, very few studies to date have controlled for baseline emotional or behavioral difficulties and so we cannot be sure that language

problems explain unique variance in later mental health difficulties. Nonetheless, recent research with the Millennium Cohort Study (MCS; which did account for such factors), found that those ‘at risk’ of DLD at age 5 (operationalized as having low vocabulary scores and/or parent-reported language difficulties) were more likely to have parent-reported internalizing symptoms at age 11 (St Clair, Forrest, Yew, & Gibson, 2019; see also Forrest, Gibson, Halligan, & St Clair, 2020). Furthermore, evidence from the 1970 British Cohort Study (BCS) suggests 5-year-olds with language difficulties are more likely to self-report internalizing symptoms at age 34 (Schoon, Parsons, Rush, & Law, 2010).

The key question for this study was whether the relation between early vocabulary and adolescent mental health that is observed in many studies in children with, or at risk of, DLD extends to the general population when looking across the full continuum of vocabulary ability. Current evidence with regard to this question presents a mixed picture. Westrupp et al. (2019) found that lower vocabulary at 4–5 years predicted greater internalizing symptoms at the age of 8–9 years but found no association between childhood vocabulary and internalizing symptoms in adolescence (14–15 years). In contrast, other studies have found that poorer language skills in the general population throughout childhood (ages 4–10) are associated with more internalizing symptoms in adolescence (ages 14–15; Bornstein, Hahn, & Suwalsky, 2013; Miller et al., 2020). A recent meta-analysis (that included clinical and nonclinical samples) suggested that there is a small, negative association between language ability and internalizing symptoms (Hentges, Devereux, Graham, & Madigan, 2021). However, there are a number of reasons why further research is warranted with large cohort studies run across generations in nationally representative samples.

First, few studies to date have adequately controlled for factors such as early child and parent mental health difficulties and socioeconomic status (SES). Vocabulary size and processing speed are positively associated with social advantage from 18 months in the United States and the United Kingdom (Fernald, Marchman, & Weisleder, 2013; McGillion, Pine, Herbert, & Matthews, 2017; Pace, Luo, Hirsh-Pasek, & Golinkoff, 2017), a relation that persists throughout the life span (Sullivan, Moulton, & Fitzsimons, 2021). We therefore included robust SES confounders across our analyses.

Children born into more deprived backgrounds also have a higher risk of mental health problems

(Reiss, 2013). Indeed, SES reflects a host of important life experiences and cultural differences that can affect mental health (Power et al., 2007). In this study, we tested whether any specific association between language and mental health remained once SES and other relevant childhood confounders, such as maternal and childhood mental health, were taken into account. We also report unadjusted models to give the full picture regarding the influence of these covariates.

Second, in studies to date, information about mental health has been obtained either by asking the individual concerned (e.g., Conti-Ramsden & Botting, 2008; Conti-Ramsden et al., 2013; Wadman et al., 2011), by asking others such as parents (e.g., Forrest et al., 2020; St Clair et al., 2019), or by some conflation of these measures (e.g., Bornstein et al., 2013; Miller et al., 2020). When considering internalizing symptoms (such as feelings of low mood or worrying), individuals themselves are uniquely well positioned to know how they are feeling. Many self-report measures have been validated for use with clinical and community samples (e.g., Sharp, Goodyer, & Croudace, 2006; Thabrew, Stasiak, Bavin, Frampton, & Merry, 2018) and self-reports are the recommended measure to use according to The Good Childhood Report (2019). In contrast, parents and other adults may not know the full extent of internalizing symptoms, unless the adolescent discloses such feelings to them. Indeed, parent- and self-report measures are not highly correlated (typical correlations $\sim .2$; Rescorla et al., 2013). In this study, self-report was preregistered as our primary outcome measure. In additional analyses, we then tested whether the choice of self-report over parent-report affected our findings.

Finally, the relation between language and mental health could plausibly be changing over historical time (Yew & O’Kearney, 2013). It has been argued that the transition to a knowledge-based economy has increased the economic importance of cognitive resources, including language (Beddington et al., 2008). At the same time, adolescent internalizing problems have become more prevalent (Bor, Dean, Najman, & Hayatbakhsh, 2014; Patalay & Gage, 2019). The current research, therefore, explored the relation between vocabulary and adolescent internalizing mental health in two large, nationally representative cohort studies with cohort members born 30 years apart: the BCS (children born in 1970) and the MCS (children born in 2000–2002). This cross-cohort comparison allowed us to investigate the relation across a time period that has seen an increase in both reliance on cognitive

ability, and in the prevalence of internalizing mental health difficulties.

We preregistered two main analyses to assess whether early vocabulary is associated with self-reported adolescent internalizing symptoms in the general population. The first analysis assessed this with the BCS and the second with the MCS. To better understand the findings and connect them with existing literature, we also report two exploratory analyses. The first repeated the main analyses, but with the vocabulary predictor dichotomized (whether or not the child had a language difficulty, operationalized as scoring 1 *SD* below the mean for vocabulary). This permitted comparison with prior work that has sought to identify children with language difficulties in this way (e.g., Schoon et al., 2010). The second exploratory analysis repeated the main analyses, but with self-reported adolescent internalizing symptoms considered as a binary outcome, according to clinical threshold cut-offs. This allowed us to check whether the relation between vocabulary and internalizing symptoms differed for those with clinical levels of internalizing symptoms. Finally, the Supporting Information reports three analyses that assessed the role of vocabulary when parent-reported adolescent internalizing symptoms were considered as the outcome in each cohort, and when an adult outcome point was considered for the BCS. The latter analysis allows comparison with Schoon et al.’s (2010) findings of an association between early vocabulary (dichotomous variable: difficulty or not) and adult mental health in the BCS (see Supporting Information, section 10).

Across all analyses, we adjusted for demographic, SES, and childhood psychosocial variables in order to better capture the unique role that early childhood vocabulary plays in internalizing symptoms. We hypothesized that after accounting for sociodemographic and childhood psychosocial factors, lower vocabulary scores would be associated with higher internalizing symptom scores (i.e., poorer mental health).

Method

Data

Data from two national birth cohort studies were used: the BCS and the MCS. The BCS follows 16,571 children born in England, Scotland, and Wales during one week in 1970 (Elliott & Shepherd, 2006) and has four childhood sweeps (ages 0, 5, 10, and 16 years). More information about this cohort study can be found here: <https://cls.ucl.ac.uk/cls->

studies/1970-british-cohort-study/. The MCS follows 19,244 young people born across England, Scotland, Wales, and Northern Ireland in 2000–2002 (Connelly & Platt, 2014) and there are currently six sweeps (ages 9 months, 3, 5, 7, 11, and 14 years). The age 14 sweep of the MCS took place in 2015, and therefore this cohort represents contemporary adolescence. More information about this cohort study can be found here: <https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/>.

Participants

For the BCS, information about all babies born between April 5 and April 11, 1970 was requested (this was not restricted to babies born in the NHS; Institute of Child Health, 1970). The sample was supplemented with children who were born in the eligible week and had subsequently moved to the United Kingdom; there were an additional 79 new cohort members at age 5, 294 at age 10 and 65 at age 16 (CLS website: <https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study/>).

For the MCS, a stratified clustered sample design was used, which specifically over-recruited subgroups of the population (ethnic minorities, disadvantaged areas, and the smaller U.K. countries). Eligible children (living in the United Kingdom at age 9 months, born within the eligible time period—September 1, 2000–August 31, 2001 for England and Wales, and November 23, 2000–January 11, 2002 for Scotland and Northern Ireland—and receiving child benefit at age 9 months) were identified by government child benefit records and sampled from electoral wards (Connelly & Platt, 2014). Seventy-two percent of eligible families responded to the 9 months sweep of data collection. The original sample was supplemented in the age 3 sweep with families who were eligible to be included, but were not recruited due to recently moving to the eligible address; this resulted in an additional 692 families being interviewed (Connelly & Platt, 2014).

Families with multiple births in the cohorts were excluded due to possible differences in the language learning environments experienced by these children (BCS: 189 pairs of twins and 1 set of triplets—2.30%; MCS: 251 pairs of twins, 11 sets of triplets, and 6 families with two singleton cohort members—2.84%; Thorpe, Rutter, & Greenwood, 2003).

For the BCS, we selected singleton cohort members with complete responses for the English Picture Vocabulary Test (EPVT; Brimer & Dunn, 1962; age 5). This resulted in a sample of 11,640

individuals. The majority of cohort members in our analytic sample were of White ethnicity (96%) and spoke only English (98%).

For the MCS, we considered singleton cohort members with complete responses on the British Ability Scale, 2nd ed. (BAS–II) naming vocabulary scale (age 5; Elliott, Smith, & McCulloch, 1996), resulting in a final sample of 14,754 cohort members. Eighty-nine percent of our analytic sample were of a White ethnicity, and 90% spoke only English in the home.

Measures

Predictor Variable: Age 5 Vocabulary

For the BCS, receptive vocabulary was measured at age 5 using the EPVT (Brimer & Dunn, 1962; see Supporting Information, section 1 for test details). The EPVT has been reported to have a reliability coefficient of .96 (Osborn, Butler, & Morris, 1984). For the MCS, expressive vocabulary was measured using the naming vocabulary subtest of the BAS–II (Elliott et al., 1996), which was administered to cohort members during the third sweep (aged around 5 years; see Supporting Information, section 1). The Naming Vocabulary subscale of the BAS has been reported to have a reliability coefficient of .65 for 5-year-olds (Elliott, Smith, & McCulloch, 1997). Note that receptive and expressive vocabulary measures tend to be moderately to highly correlated (e.g., Conway et al., 2017).

Due to the nature of the naming vocabulary test, MCS cohort members did not complete the same items, as progression through the test depends on their performance and poor performance may result in the administration of an easier set of items. Therefore, in our analyses, we used ability scores, adjusted for item difficulty. The same set of items were administered to all BCS cohort members and raw scores were therefore used. Because MCS cohort members were born over a 2-year period (2000–2002), they were different ages when they completed the naming vocabulary test (mean age of 62.51 months, range 52.87–75.52 months). Additionally, fieldwork in the BCS age 5 follow-up was conducted over 6 months in 1975, and cohort members were thus different ages when they completed the EPVT (mean of 60.92 months, ranging from 58.78 to 75.52 months). We, therefore, adjusted for age in months at the time of the test, in both cohorts. In both tests, higher scores indicate a higher ability. All scores and ages were converted to z-scores for analyses.

Outcome Variable: Adolescent Internalizing Symptoms

For the BCS, total scores on the nine-item Malaise Inventory (Rutter, Tizard, & Whitmore, 1970) were used as a measure of internalizing symptoms at age 16. Scores ranged from 0 to 9. For the MCS, cohort members were given the 13-item Short Mood and Feelings Questionnaire (SMFQ; Angold et al., 1995) at age 14. Scores ranged from 0 to 26. For both scales, higher scores indicate greater severity of internalizing symptoms. These two scales have similar items relating to the same domains, such as tiredness, restlessness, and mood, and both are reliable and valid indicators of internalizing symptoms (Daviss et al., 2006; Rutter et al., 1970). In the current samples, there was an alpha of .7 (BCS Malaise Inventory) and .93 (MCS SMFQ). All scores were converted to z-scores for analyses.

Potential Confounding Variables

Biological and SES variables. Biological risk variables included in all models were the child's birth weight (in grams, converted to z-scores for analyses), gestational age in days (converted to z-scores for analyses), and sex (male = 0, female = 1). We also included ethnicity and the main language spoken in the home. SES variables included in all models were the highest level of parental education achieved and highest occupational status in the household at birth (a three-category measure with a fourth category for unemployment. BCS: (a) professional and managerial, (b) skilled, (c) semi-skilled and unskilled; MCS: (a) higher managerial, (b) intermediate, (c) routine and manual). The MCS has a richer set of indicators of SES, allowing us to include two further SES variables for MCS analyses: UK OECD weighted income quintiles (taken from the first sweep, an indication of household income; 1 = lowest, 5 = highest) and net total wealth (converted to z-scores for analyses). We derived the latter measure by summing information collected at age 11 (MCS5) about net housing wealth (house value net of outstanding mortgage) and net financial wealth (total savings net of any owed debts).

Mother and child psychosocial variables. Maternal psychosocial variables included whether the cohort member's mother was a teenage mother at their birth (0 = yes, 1 = no), the marital status of the mother at birth (0 = partnered, 1 = not partnered) and maternal depression when children were aged 5. In the BCS, this was assessed using the Malaise Inventory (full version; Rutter et al., 1970). In the MCS, this was assessed using the Kessler K6 scale

(Kessler et al., 2003). Items on the two scales are similar; for example, both ask questions regarding feelings of low mood and restlessness. These variables were converted to z-scores for analyses.

Child psychosocial variables were internalizing and externalizing difficulties at the age of 5. In the BCS, cohort member's parents completed the Rutter "A" scale (Rutter et al., 1970). For this study, a neurotic score and an antisocial score were calculated, as detailed by Rutter et al. (1970) and used as indicators of internalizing and externalizing behavioral difficulties, respectively. In the MCS, parents completed the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), which was developed as the successor to the Rutter scales. Items on the SDQ emotional symptoms and conduct problems subscales are similar to those of the neurotic and antisocial subscales of the Rutter "A" scale. Total scores from the emotional symptoms and conduct problems subscales were calculated as indicators of internalizing and externalizing behavior problems, respectively. For both scales, a higher score indicates increased difficulties. These variables were converted to z-scores for analyses.

Data Analysis

The main analyses consisted of two multiple linear regressions: (a) BCS data, with age 16 self-reported internalizing symptoms as the outcome and (b) MCS data, with age 14 self-reported internalizing symptoms as the outcome. These confirmatory analyses were preregistered at the Open Science Framework website (OSF number: osf.io/a94bh).

Missing Data Strategy

Sampling weights were applied to the analyses of MCS data, to account for the stratified clustered design of the data and the oversampling of subgroups. The BCS does not have a complex survey design and therefore sample weights were not required for this cohort. Missing data in all analyses were accounted for with multiple imputation using chained equations with the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011). Each data set was imputed 25 times, as this was greater than the percentage of missing cells in both cohorts (7.67% BCS, 10.26% MCS). Across our chosen samples for each analysis, no data were missing for the main predictor variable (vocabulary score) or sex. All analyses were conducted in RStudio (RStudio Team, 2020).

Analysis Plan

Initially, the raw association between vocabulary and internalizing symptoms was estimated to assess whether or not there was an association before the addition of potential confounding variables. Subsequently, to determine whether there was a relation between age 5 vocabulary and adolescent internalizing symptoms after adjusting for SES and childhood psychosocial variables, the following nested models were estimated for both BCS and MCS data. Biological and sociodemographic factors were added in the first model. Mother and child psychosocial variables were then added in a second model. The vocabulary predictor was added to a third model, and quadratic and cubic terms were added to the vocabulary predictor in a fourth model to test for any nonlinearities. Regression estimates were pooled based on Rubin's rules (Rubin, 1984). Mean centering was carried out for all continuous variables for all analyses.

The model containing biological and SES variables was initially compared to a model with no predictors. Each model was then compared to the previous model. The new predictors in each model were added to the existing predictors in the previous model and therefore our models were nested within each other. Improvements in fit were assessed using model comparisons for imputed data, using the method of Meng and Rubin (1992). If an improvement in model fit was seen when adding the main variable of interest (age 5 vocabulary), this would indicate that language predicted unique variance in adolescent internalizing symptoms. If an improvement in model fit was seen after adding the quadratic and cubic terms to vocabulary, this would suggest nonlinearities in the relation between age 5 vocabulary and adolescent internalizing symptoms. Pooled partial R^2 values are reported for all variables, computed using the method outlined by Harel (2009).

Sensitivity and Supplementary Analyses

In order to assess whether the ethnic make-up of our selected samples could be driving any observed effects, we ran two sensitivity analyses: (a). BCS: age 5 vocabulary predicting age 16 internalizing symptoms: White, English-speaking sample; and (b). MCS: age 5 vocabulary predicting age 14 internalizing symptoms: White, English-speaking sample, with matching potential confounding variables to the BCS analysis. In order to assess whether the different items tapping internalizing symptoms for

the two cohorts could be driving any differences, we ran a third sensitivity analysis using a harmonized matched subset of items from the self-reported internalizing symptoms subscales from the BCS and MCS. These can be found in Supporting Information.

In order to allow comparison with existing literature, we also carried out three supplementary analyses: (1) BCS: age 5 vocabulary predicting age 16 *parent-reported* internalizing symptoms; (2) MCS: age 5 vocabulary predicting age 14 *parent-reported* internalizing symptoms; (3) BCS: age 5 vocabulary predicting *age 34* internalizing symptoms. There are multiple potential reporters for adolescent mental health. Rates of agreement between parent and self-reported symptoms of adolescent internalizing symptoms are known to be low (Rescorla et al., 2013) and studies to date have varied in the measure used. Analyses 1 and 2 were therefore carried out in order to assess whether the size and direction of any associations differed as a function of the reporter. Analysis 3 was completed to complement our adolescent analyses—to see whether or not any relation persisted into adulthood in the BCS sample, allowing for comparison with Schoon et al. (2010). Main findings for each can be found below, full details and results can be found in Supporting Information.

Results

Descriptive Statistics

Descriptive statistics can be found in Table 1. These were estimated across 25 imputed data sets.

For the BCS, differences in the proportions between the full cohort sample and the analytical sample in this article (i.e., everyone with a vocabulary score at the age of 5) are negligible (see Table S1). For the MCS, differences between the full cohort sample and the selected analytical sample are also negligible for most variables (see Table S2). However, there are more unemployed parents in the full sample compared to the analytical sample.

As expected, based on demographic trends in the United Kingdom, there were more White ethnicity participants in the BCS compared to the MCS and more parents in the MCS had university level qualifications (higher degree, first degree, diploma in education; 38.96%) compared to BCS (first degree, postgraduate degree, national diploma or certificate, membership of a professional institution, city and guilds full technical certificate, certificate of education, state registered nurse; 16.84%).

Table 1
 Mean (SD), Proportions (%) and 95% CIs for Analysis Samples in the BCS (Born 1970) and MCS (Born ~2001) Cohorts

BCS: variable names	BCS M (SD) or % [95% CI] N = 11,640	MCS: variable names	MCS M (SD) or % [95% CI] N = 14,754
Mental health		Mental health	
Age 16 self-reported internalizing symptoms	3.77 (2.20) [3.73; 3.81]	Age 14 self-reported internalizing symptoms	5.62 (5.90) [5.53; 5.72]
Age 16 parent-reported internalizing symptoms	1.95 (1.89) [1.92; 1.99]	Age 14 parent-reported internalizing symptoms	1.98 (2.12) [1.95; 2.02]
Language		Language	
Age 5 language score (EPVT)	35.32 (10.81) [35.12; 35.51]	Age 5 language (naming vocabulary)	109.21 (15.61) [108.95; 109.46]
Age of CM at the time of vocabulary test (months)	60.92 (1.28) [60.89; 60.94]	Age of CM at time of vocabulary test (months)	62.51 (2.92) [62.46; 62.56]
Biological risk variables		Biological risk variables	
Birth weight (g)	3,331.09 (517.01) [3,321.70; 3,340.48]	Birth weight (g)	3,388.62 (1,959.92) [3,357.00; 3,420.25]
Gestational age (days)	281.86 (16.44) [281.56; 282.16]	Gestational age (days)	276.13 (13.44) [275.91; 276.35]
Sex (female)	48.21%	Sex (female)	48.93%
Ethnicity (White UK)	96.22%	Ethnicity (White)	88.61%
Ethnicity (minority)	3.78%	Ethnicity (mixed)	2.91%
		Ethnicity (Indian)	1.78%
		Ethnicity (Pakistani/Bangladeshi)	3.47%
		Ethnicity (Black/Black British)	2.22%
		Ethnicity (Other, including Chinese)	1%
Sociodemographic variables		Sociodemographic variables	
Language used in home (English)	97.63%	Main language in home (English)	90.20%
Language used in home (other than English)	2.37%	Main language in home (English and another language)	7.86%
		Main language in home (only another language)	1.94%
Occupation (professional and managerial)	20.95%	Occupation (NS-SEC higher managerial)	46.21%
Occupation (skilled manual and nonmanual)	61.75%	Occupation (NS-SEC intermediate)	18.68%
Occupation (semi-skilled/unskilled)	16.33%	Occupation (NS-SEC routine and manual)	26.17%
Occupation (unemployed)	0.97%	Occupation (unemployed)	8.95%
Parent education (degree +)	13.7%	Parent education (higher degree)	7.64%
Parent education (certificate of education)	1.66%	Parent education (first degree)	19.15%
Parent education (SRN)	1.67%	Parent education (diploma in higher education)	12.69%
Parent education (A levels)	7.88%	Parent education (A levels)	10.05%
Parent education (O levels)	21.18%	Parent education (O levels/GCSE grades A–C)	33.05%
Parent education (vocational qualification)	13.2%	Parent education (GCSE grades D–G)	7.36%
Parent education: no qualifications	40.71%	Parent education (none of these/other including overseas)	10.07%
		Income (lowest quintile)	17.34%
		Income (second quintile)	18.87%
		Income (third quintile)	20.05%
		Income (fourth quintile)	21.60%
		Income (highest quintile)	22.14%

Table 1 Continued

BCS: variable names	BCS	MCS: variable names	MCS
	M (SD) or % [95% CI] N = 11,640		M (SD) or % [95% CI] N = 14,754
		Total net wealth (£)	193,416.77 (518,334.2) [185,052.83; 201,780.72]
Childhood psychosocial variables		Childhood psychosocial variables	
Teen mum (yes)	8.85%	Teen mum (yes)	5.89%
Marital status (not partnered)	5.37%	Marital status (not partnered)	32.96%
Maternal depression (CM age 5)	4.32 (3.63) [4.25; 4.39]	Maternal depression (CM age 5)	3.14 (3.77) [3.08; 3.20]
Age 5 CM externalizing difficulties	1.8 (1.65) [1.77; 1.83]	Age 5 CM externalizing difficulties	1.23 (1.56) [1.21; 1.26]
Age 5 CM internalizing difficulties	1.5(1.5) [1.48; 1.53]	Age 5 CM internalizing difficulties	1.34 (1.56) [1.31; 1.36]

Note. Means, SDs, proportions and 95% CIs are pooled across 25 imputed data sets (both cohorts) and are sample weighted (MCS only). BCS = 1970 British Cohort Study; MCS = Millennium Cohort Study; EPVT = English Picture Vocabulary Test; CM = cohort member; NS-SEC = National Statistics Socio-Economic Classification; SRN = state registered nurse; GCSE = General Certificate of Secondary Education.

Does Age 5 Vocabulary Predict Age 16 Internalizing Symptoms in the BCS (Born 1970)?

In an unadjusted model (i.e., not including any potential confounding variables), there was a significant negative relation between vocabulary and self-reported internalizing symptoms, such that higher vocabulary scores were associated with fewer internalizing symptoms ($\beta = -0.03$ [-0.06; -0.01]; see Table 3). To test whether this relation held when potential confounding factors were included, we first tested whether two sets of potential confounding variables predicted internalizing symptoms and then whether vocabulary explained variance over and above these variables. Compared to a model with no predictors, a model with biological and SES control variables significantly improved the model fit, ($Dm(14, 1,336.91) = 10.57, p < .001$; see Table S3). Compared to a model with only these variables, a model that also included mother and child psychosocial variables had a significantly improved fit, ($Dm(5, 374.65) = 5.24, p = .001$). Compared to a fully adjusted model, adding receptive vocabulary scores (Model 3) did not significantly improve the model fit, ($Dm(2, 138.26) = 0.04, p = .965$; see Table 2). We examined a model with quadratic and cubic terms which did not improve the model fit, ($Dm(2, 226.84) = 0.34, p = .710$), suggesting the absence of nonlinear relations between age 5 vocabulary and age 16 internalizing

symptoms. These results suggest that age 5 vocabulary does not predict any unique variance in age 16 self-reported internalizing symptoms in this cohort after accounting for potential confounding variables. Given the unadjusted relation between age 5 vocabulary and age 16 internalizing symptoms, we ran a post hoc analysis whereby we added the vocabulary predictor to a model containing biological and SES variables. This was to check which potential confounder(s) removed the relation. Compared to a model with only biological and SES variables, adding receptive vocabulary scores did not significantly improve the model fit, ($Dm(2, 137.73) = 0.28, p = .756$). Sex was the only significant predictor in this model ($\beta = 0.31$ [0.25; 0.37], see Table S4), and we, therefore, conclude that the appearance of a relation between vocabulary and internalizing symptoms in the unadjusted analysis is spurious and due to the collinearity of sex and vocabulary size in this cohort (see Patalay & Fitzsimons, 2018; Stolarova et al., 2016 for further evidence that both mental health and vocabulary show sex differences).

Sensitivity analysis 1 (restricting the analysis to a White, English-speaking subsample) revealed a similar pattern of results (see Supporting Information). However, a different pattern of results was observed in supplementary analysis 1, which considered *parent-reported* adolescent internalizing symptoms as the outcome variable. In a fully

Table 2

β [95% CIs] for Fully Adjusted Regression Models (Final Model) for BCS (Born 1970) Sample (N = 11,640) and MCS (Born ~2001) Sample (N = 14,754)

BCS: variable names	BCS β [95% CI] <i>Partial R</i> ² N = 11,640	MCS: variable names	MCS β [95% CI] <i>Partial R</i> ² N = 14,754
Biological risk variables		Biological risk variables	
Birth weight (g)	-0.02 [-0.05; 0.01] <i>p</i> = .26 .0002	Birth weight (g)	0.00 [-0.03; 0.04] <i>p</i> = .81 0
Gestational age (days)	0.00 [-0.03; 0.03] <i>p</i> = .99 0	Gestational age (days)	0.01 [-0.01; 0.03] <i>p</i> = .49 0
Sex (male)	Reference	Sex (male)	Reference
Sex (female)	0.32 [0.26; 0.38]*** <i>p</i> ≤ .001 .0247	Sex (female)	0.54 [0.50; 0.58]*** <i>p</i> < .001 .0728
Ethnicity (White)	Reference	Ethnicity (White)	Reference
Ethnicity (minority)	0.05 [-0.11; 0.20] <i>p</i> = .56 0	Ethnicity (mixed)	0.02 [-0.09; 0.12] <i>p</i> = .76 .0008
		Ethnicity (Indian)	-0.11 [-0.27; 0.04] <i>p</i> = .15
		Ethnicity (Pakistani/Bangladeshi)	-0.21 [-0.34; -0.08]*** <i>p</i> < .001
		Ethnicity (Black/Black British)	-0.12 [-0.25; 0.01] <i>p</i> = .06
		Ethnicity (Other, including Chinese)	-0.11 [-0.31; 0.09] <i>p</i> = .28
Sociodemographic variables		Sociodemographic variables	
Language used in home (English)	Reference	Main language in home (English)	Reference
Language used in home (other than English)	0.01 [-0.17; 0.19] <i>p</i> = .95 0	Main language in home (English and another language)	0.01 [-0.09; 0.11] <i>p</i> = .83 0
		Main language in home (only another language)	0.01 [-0.16; 0.17] <i>p</i> = .93
Occupation (professional and managerial)	Reference	Occupation (NS-SEC higher managerial)	Reference
Occupation (skilled manual and nonmanual)	0.00 [-0.06; 0.07] <i>p</i> = .96 .0001	Occupation (NS-SEC intermediate)	0.01 [-0.05; 0.07] <i>p</i> = .76 0
Occupation (semi-skilled/unskilled)	-0.01 [-0.10; 0.08] <i>p</i> = .84	Occupation (NS-SEC routine and manual)	0.02 [-0.04; 0.08] <i>p</i> = .49
Occupation (unemployed)	0.10 [-0.21; 0.40] <i>p</i> = .53	Occupation (NS-SEC unemployed)	0.03 [-0.06; 0.13] <i>p</i> = .46
Parent education (degree +)	Reference	Parent education (higher degree)	Reference
Parent education (certificate of education)	0.07 [-0.12; 0.25] <i>p</i> = .48 .0002	Parent education (first degree)	-0.05 [-0.13; 0.02] <i>p</i> = .18 .0001

10 Thornton, Patalay, Matthews, and Bannard

Table 2 Continued

BCS: variable names	BCS β [95% CI] <i>Partial R</i> ² N = 11,640	MCS: variable names	MCS β [95% CI] <i>Partial R</i> ² N = 14,754
Parent education (SRN)	−0.01 [−0.20; 0.18] <i>p</i> = .91	Parent education (diploma in higher education)	−0.05 [−0.13; 0.04] <i>p</i> = .30
Parent education (A levels)	0.01 [−0.09; 0.11] <i>p</i> = .83	Parent education (A levels)	−0.04 [−0.13; 0.05] <i>p</i> = .37
Parent education (O levels)	0.02 [−0.07; 0.11] <i>p</i> = .62	Parent education (O levels/GCSE grades A–C)	−0.06 [−0.14; 0.03] <i>p</i> = .17
Parent education (vocational qualification)	−0.03 [−0.14; 0.08] <i>p</i> = .59	Parent education (GCSE grades D–G)	−0.02 [−0.13; 0.10] <i>p</i> = .79
Parent education: no qualifications	0.02 [−0.06; 0.11] <i>p</i> = .61	Parent education (none of these/other including overseas)	−0.05 [−0.16; 0.05] <i>p</i> = .33
		Income (lowest quintile)	Reference
		Income (second quintile)	0.03 [−0.04; 0.10] <i>p</i> = .41 .0006
		Income (third quintile)	−0.03 [−0.10; 0.04] <i>p</i> = .43
		Income (fourth quintile)	−0.04 [−0.11; 0.04] <i>p</i> = .34
		Income (highest quintile)	−0.08 [−0.15; 0.00] <i>p</i> = .05
		Total net wealth (£)	−0.02 [−0.04; −0.00]* <i>p</i> = .03 .0004
Childhood psychosocial variables		Childhood psychosocial variables	
Teen mum (no)	Reference	Teen mum (no)	Reference
Teen mum (yes)	0.10 [0.01; 0.19]* <i>p</i> = .03 .0007	Teen mum (yes)	0.01 [−0.12; 0.14] <i>p</i> = .88 0
Marital status (partnered)	Reference	Marital status (partnered)	Reference
Marital status (not partnered)	0.01 [−0.12; 0.14] <i>p</i> = .92 0	Marital status (not partnered)	0.07 [0.02; 0.13]*** <i>p</i> < .001 .001
Maternal depression (CM age 5)	0.02 [−0.01; 0.05] <i>p</i> = .29 .0001	Maternal depression (CM age 5)	0.07 [0.04; 0.09]*** <i>p</i> < .001 .0038
Age 5 CM externalizing difficulties	0.04 [0.01; 0.07]** <i>p</i> = .01 .0013	Age 5 CM externalizing difficulties	0.05 [0.02; 0.07]*** <i>p</i> < .001 .0016
Age 5 CM internalizing difficulties	0.05 [0.02; 0.08]*** <i>p</i> ≤ .001 .0025	Age 5 CM internalizing difficulties	0.02 [−0.00; 0.04] <i>p</i> = .07 .0003

Table 2 Continued

BCS: variable names	BCS	MCS: variable names	MCS
	β [95% CI] <i>Partial R</i> ² N = 11,640		β [95% CI] <i>Partial R</i> ² N = 14,754
Age 5 language		Age 5 language	
Age of CM at time of taking vocabulary test (months)	0.00 [−0.03; 0.03] <i>p</i> = .82 0	Age of CM at time of taking vocabulary test (months)	0.01 [−0.01; 0.03] <i>p</i> = .33 0
Age 5 vocabulary	0.00 [−0.03; 0.03] <i>p</i> = .94 0	Age 5 vocabulary	0.05 [0.02; 0.07]*** <i>p</i> < .001 .0017
<i>R</i> ²	.0332	<i>R</i> ²	.0893

Note. These are the results for the final model (Model 3) in the hierarchical regression. See Supporting Information, section 3 for the full regression model tables. BCS = 1970 British Cohort Study; MCS = Millennium Cohort Study; CM = cohort member; NS-SEC = National Statistics Socio-Economic Classification; SRN = state registered nurse; GCSE = General Certificate of Secondary Education. **p* < .05. ***p* < .01. ****p* < .001.

adjusted model, this analysis revealed a significant *negative* relation between age 5 vocabulary and parent-reported adolescent internalizing symptoms ($\beta = -0.06$ [−0.08; −0.03]; see Table 3). This suggests that lower vocabulary scores in childhood were predictive of more parent-reported internalizing symptoms in adolescence (Table S10). A sensitivity analysis with parent-reported symptoms that considered only White, English-speaking cohort members yielded a similar pattern of results.

Measures of self-reported internalizing symptoms are also available in the adulthood sweeps of the BCS. In order to investigate the longer-term effects of age 5 vocabulary on internalizing symptoms, we ran an analysis with age 34 internalizing symptoms as the outcome variable, extending the findings of Schoon et al. (2010), by considering vocabulary as a continuous predictor of age 34 internalizing symptoms. In a fully adjusted model, this analysis revealed a significant *negative* relation between age 5 vocabulary and age 34 internalizing symptoms ($\beta = -0.07$ [−0.09; −0.04]; see Supporting Information, section 10), such that those with lower vocabulary scores in childhood self-reported more internalizing symptoms in adulthood. This differs from the findings of the preregistered analysis with age 16 self-reported symptoms as the outcome (see Figure 1).

Does Age 5 Vocabulary Predict Age 14 Internalizing Symptoms in the MCS (Born ~2001)?

In an unadjusted model, a significant *positive* relation between vocabulary and self-reported internalizing symptoms was observed such that higher

vocabulary scores were associated with *more* internalizing symptoms ($\beta = 0.03$ [0.01; 0.06]; see Table 3). To test whether this relation held when control factors were included, we first tested whether two sets of control variables predicted internalizing symptoms in the MCS and then whether vocabulary explained variance over and above these variables. Compared to a model with no predictors, a model with biological and sociodemographic control variables significantly improved the model fit, ($Dm(24, 4,312.52) = 36.69, p < .001$). Compared to a model with only these variables, a model that also adjusted for mother and child psychosocial variables gave a significantly improved fit, ($Dm(5, 429.41) = 14.43, p < .001$; see Table S5). Compared to a fully adjusted model, adding expressive vocabulary scores in Model 3 accounted for significantly more variance in the outcome, ($Dm(2, 331.61) = 9.93, p < .001$; see Table 2), such that higher vocabulary scores were associated with more internalizing difficulties in adolescence. We examined a model with quadratic and cubic terms which did not improve the model fit, ($Dm(2, 366.39) = 0.03, p = .975$), suggesting the absence of nonlinear relations between age 5 vocabulary and age 14 internalizing symptoms. In sum, for children born in 2000–2002, age 5 vocabulary ability explains some unique variance in age 14 internalizing symptoms, such that better childhood vocabulary ability predicts *poorer* adolescent internalizing symptoms. The effect size of the vocabulary predictor ($\beta = 0.05$ [0.02; 0.07]) indicates that a 1 *SD* increase in vocabulary was associated with an increase of 5% of a standard deviation in internalizing symptoms. Despite being small in size, this effect is of

12 Thornton, Patalay, Matthews, and Bannard

Table 3

Main Exposure (Age 5 Language)—Outcome (Internalizing Symptoms) Relationship in Each Analysis

Analysis details		BCS Age 5 language (β coef [95% CI] <i>p</i> value)	MCS Age 5 language (β coef [95% CI] <i>p</i> value)
Self-report: unadjusted	Age 5 language predicting self-reported adolescent mental health, unadjusted model	−0.03 [−0.06; −0.01]*** <i>p</i> < .001	0.03 [0.01; 0.06]*** <i>p</i> < .001
Self-report: fully adjusted	Main, preregistered analysis. Age 5 language ability predicting self-reported adolescent mental health, after adjusting for demographic, socioeconomic, and childhood indicators of poor mental health	0.00 [−0.03; 0.03] <i>p</i> = .94	0.05 [0.02; 0.07]*** <i>p</i> < .001
White subset	Analysis conducted as a sensitivity check. Model built in the same way as the main self-report analysis, this time considering only English-speaking cohort members of a White ethnicity	0.01 [−0.03; 0.04] <i>p</i> = .72	0.03 [0.00; 0.05]* <i>p</i> = .02
Parent-report	Supplementary analysis. Models built in the same way as the main analysis. Here, the outcome variable was parent-reported mental health, to assess whether the main analysis results changed when considering a different perspective	−0.06 [−0.08; −0.03]*** <i>p</i> < .001	−0.03 [−0.05; −0.01]*** <i>p</i> < .001
White subset (parent-report)	Analysis conducted as a sensitivity check. Model built in the same way as the parent-report analysis, this time considering only English-speaking cohort members of a White ethnicity	−0.06 [−0.08; −0.03]*** <i>p</i> < .001	−0.04 [−0.07; −0.02]*** <i>p</i> < .001
Exploratory 1: binary language ^a : those scoring 1 <i>SD</i> below the mean defined as language difficulties (0 = normal language, 1 = language difficulties)	Exploratory analysis 1. Models built in the same way as the main self-report analysis. This time language considered as a binary predictor, to consider those specifically with language difficulties	−0.02 [−0.10; 0.05] <i>p</i> = .54	−0.04 [−0.09; 0.02] <i>p</i> = .22
Sensitivity: binary language ^a : those scoring 1.5 <i>SD</i> below the mean defined as language difficulties (0 = normal language, 1 = language difficulties)	Sensitivity check with more stringent cut off point. Models built in the same way as the main self-report analysis	−0.03 [−0.15; 0.08] <i>p</i> = .59	−0.11 [−0.21; −0.02]** <i>p</i> = .01

Table 3 Continued

Analysis details		BCS Age 5 language (β coef [95% CI] <i>p</i> value)	MCS Age 5 language (β coef [95% CI] <i>p</i> value)
Exploratory 2: binary internalizing symptoms (self-report) OR [95% CIs]	Exploratory analysis 2. Internalizing symptoms outcome considered as binary. Models built in the same way as the main self-report analysis	1.00 [0.93; 1.08] <i>p</i> = .95	1.16 [1.07; 1.25]*** <i>p</i> < .001
Sensitivity: complete case for language and internalizing symptoms (self-report)	The main analysis considered all of those cohort members with just complete scores for the language predictor. However, due to debate around whether or not the outcome variable should be imputed, we also conducted a sensitivity check whereby we considered complete cases for language and the self-reported mental health outcome	0.00 [−0.03; 0.03] <i>p</i> = .91	0.05 [0.03; 0.08]*** <i>p</i> < .001
Age 34 mental health as the outcome variable	Supplementary analysis. Models built in the same way as the main preregistered analysis	−0.07 [−0.09; −0.04]*** <i>p</i> < .001	

Note. Binary language ^a: coefficient for poor language, normal language = reference group. β coefficients for the White subset, binary language, and parent-reported outcome models are taken from the fully adjusted models (see Supporting Information). OR = Odds Ratio.

p* < .05. *p* < .01. ****p* < .001.

comparable magnitude to other predictors in the model (maternal mental health (β = 0.07 [0.04; 0.09]), childhood externalizing symptoms (β = 0.05 [0.02; 0.07]) and is larger than that of childhood parent-rated internalizing symptoms (β = 0.02 [−0.00; 0.04]); see Table S5).

A sensitivity analysis restricted to a White, English-speaking subsample, with matched BCS potential confounding variables revealed that a model with the vocabulary predictor was a significantly better fit than a model including only the biological, SES, mother, and childhood psychosocial variables: better vocabulary scores in childhood were associated with more self-reported internalizing symptoms in adolescence (see Supporting Information).

Supplementary analysis 2 considered *parent-reported* adolescent internalizing symptoms as the outcome variable and here there was a significant *negative* relation between age 5 vocabulary and

adolescent internalizing symptoms, in a fully adjusted model such that better vocabulary scores in childhood were predictive of *fewer* parent-reported internalizing symptoms in adolescence (β = −0.03 [−0.05; −0.01]; see Table 3; Table S12).

Overall, switching from self-report to parent-report of adolescent internalizing symptoms changes the direction of effect, such that good early vocabulary predicts *fewer* internalizing symptoms in both cohorts (see Figure 2). As can be seen from Figure 3, the self- and parent-reported internalizing symptoms measures are significantly different from one another (the confidence intervals for each do not overlap).

Exploratory Analyses

The following post hoc, exploratory analyses were conducted to better understand the above results in the context of the broader literature.

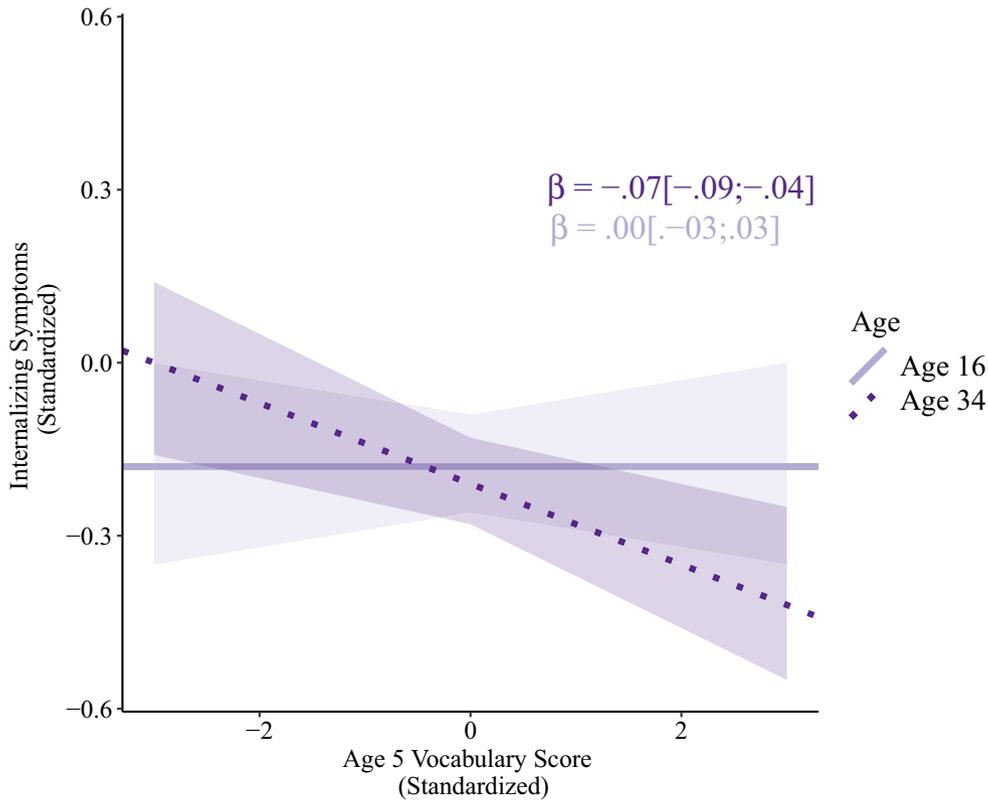


Figure 1. Age 34 and age 16 internalizing symptoms, predicted from age 5 vocabulary score.

Note. The scale of the standardized vocabulary measure ranges from -1.97 (5th percentile) to 1.45 (95th percentile). The age 16 internalizing symptoms measure standardized scale ranges from -1.71 (5th percentile) to 1.92 (95th percentile). The age 34 internalizing symptoms measure standardized scale ranges from -1.16 (5th percentile) to 1.75 (95th percentile).

Model comparisons and tables for exploratory analyses can be found in Supporting Information, section 13.

Age 5 Vocabulary as a Binary Predictor

Schoon et al. (2010) analyzed the BCS and found that vocabulary difficulties at age 5 (a dichotomous variable, where a difficulty was defined as vocabulary 1 standard deviation below the mean) were associated with poor mental health at age 34. Along with studies of DLD, Schoon et al.'s (2010) findings led us to predict that vocabulary ability across the full continuum would be negatively associated with internalizing symptoms in the general population in adolescence. However, the main preregistered results suggest that this is not the case when the full continuum of vocabulary ability is considered. To test whether the predicted association holds in adolescence when a dichotomized vocabulary predictor (vocabulary difficulty or not) is used, we ran two further models with data from the BCS and the MCS.

The absence of vocabulary difficulties was used as the reference category. Models were built in the same way as the main analyses. The vocabulary predictor was dichotomized at 1 *SD* below the mean, in line with the methodology of Schoon et al. (2010). However, some research has classified language difficulty as 1.5 *SD* below the mean (Norbury et al., 2016), and we therefore also dichotomized the vocabulary predictor using this cut off as a sensitivity analysis (see Supporting Information, section 14).

In the BCS sample, 1,872 cohort members (16% of the cohort) had vocabulary scores 1 *SD* below the mean. Results for this analysis can be found in Table S18. There was no relation between vocabulary and internalizing symptoms when vocabulary was considered as a binary predictor in a fully adjusted model ($\beta = -0.02$ [$-0.10; 0.05$]; see Table 3). This remained the case when the more stringent cut off of 1.5 *SD* below the mean was used (1,114 cohort members (10%) had scores 1.5 *SD* below the mean. See Supporting Information, section 14).

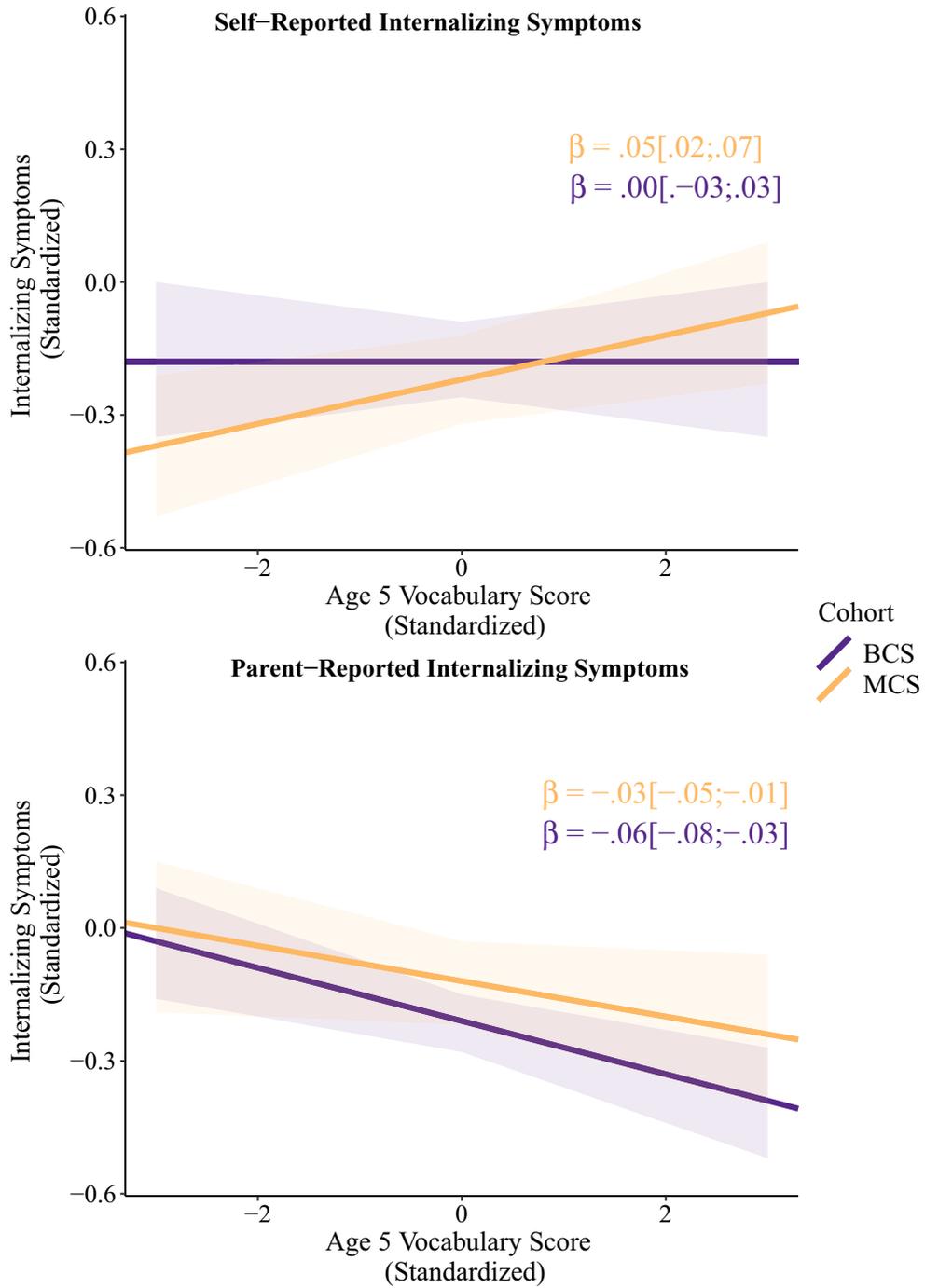


Figure 2. Adolescent self- and parent-reported internalizing symptoms, predicted from age 5 vocabulary score.
 Note. For British Cohort Study (BCS) data, the scale of the standardized vocabulary measure ranges from -1.97 (5th percentile) to 1.45 (95th percentile). The self-reported internalizing measure standardized scale ranges from -1.71 (5th percentile) to 1.92 (95th percentile). The parent-reported internalizing symptoms measure standardized scale ranges from -1.03 (5th percentile) to 2.13 (95th percentile). For Millennium Cohort Study (MCS) data, the scale of the standardized vocabulary measure ranges from -1.79 (5th percentile) to 1.44 (95th percentile). The self-reported internalizing measure standardized scale ranges from -0.95 (5th percentile) to 2.12 (95th percentile). The parent-reported internalizing symptoms measure standardized scale ranges from -0.96 (5th percentile) to 1.83 (95th percentile).

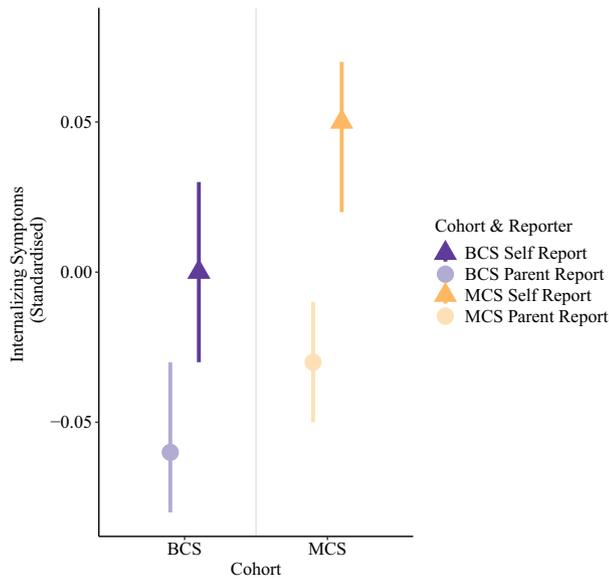


Figure 3. β [95% CIs] for internalizing symptoms (self- and parent-reported) in BCS and MCS. BCS = British Cohort Study; MCS = Millennium Cohort Study.

In the MCS sample, 2,919 cohort members (20%) had vocabulary scores 1 *SD* below the mean. Results for this analysis can be found in Table S19. There was no significant relation between age 5 vocabulary and adolescent internalizing symptoms in a fully adjusted model when this cut off was considered ($\beta = -0.04$ [-0.09; 0.02]; see Table 3). However, when the more stringent cut off of 1.5 *SD* below the mean was used, there was a significant *negative* relation between age 5 binary vocabulary and age 14 internalizing symptoms ($\beta = -0.11$ [-0.21; -0.02]. See Supporting Information, section 14). This suggests that poor vocabulary was predictive of *fewer* internalizing symptoms at age 14. While the effect size was again small, this unexpected direction of effect is consistent with the outcome of the main preregistered analyses reported earlier. When using this cut off, 8% of MCS cohort members were classed as having vocabulary difficulties (1,204 cohort members). This maps on to national prevalence levels for DLD, which are estimated to be around 7.5% (Norbury et al., 2016).

Binary Internalizing Symptoms as the Outcome Variable

In our main preregistered analyses, we found that for BCS cohort members, there was no relation between age 5 vocabulary and internalizing

symptoms. For MCS cohort members, there was a positive relation between age 5 vocabulary and self-reported internalizing symptoms. Therefore, in the second set of exploratory analyses, we investigated whether these trends remained when we considered those with clinical levels of internalizing symptoms, (scores ≥ 4 on the Malaise inventory in the BCS and scores ≥ 12 on the SMFQ in the MCS), with binary logistic regressions, whereby 0 = nonclinical levels and 1 = clinical levels of internalizing symptoms.

In the BCS sample, 4,188 cohort members had scores ≥ 4 on the Malaise inventory. This analysis revealed that the odds of having clinical levels of internalizing symptoms in adolescence did not differ as a function of vocabulary (see Table 3). This finding is in line with the main preregistered analysis for the BCS, which also suggests no relation between early vocabulary and the continuous internalizing symptoms measure.

In the MCS sample, 2,013 cohort members had scores ≥ 12 on the SMFQ. This analysis revealed that for every *SD* unit increase in vocabulary, there was a 16% increase in the odds of having clinical levels of internalizing symptoms (OR = 1.16 [1.07; 1.25], see Table 3; Table S21). This is in line with the finding of the main preregistered analysis, whereby MCS cohort members with better vocabulary in childhood were found to have more internalizing symptoms in adolescence. However, it is worth noting that compared to a model with all potential confounding variables, adding expressive vocabulary scores did not significantly improve the model fit (see Supporting Information, section 13).

Discussion

In preregistered analyses, we assessed whether early vocabulary in the general population is associated with self-reported internalizing symptoms in adolescence. The overall finding was that in a cohort of children born in 1970, there was no significant relation between early vocabulary and self-reported adolescent internalizing symptoms once a comprehensive set of potential confounding variables was included. However, a supplementary analysis revealed that a relation emerged in adulthood, such that better early vocabulary predicted fewer self-reported adult internalizing symptoms. This finding was in line with Schoon et al. (2010). Conversely, in the more recently born MCS children (born ~ 2001), better early vocabulary predicted *worse* self-reported adolescent internalizing symptoms, an effect that remained in a fully adjusted

model. In general, findings for both cohorts did not differ when vocabulary or internalizing symptoms measures were treated as dichotomous measures. Overall, our results suggest that the relation between early vocabulary and self-reported adolescent internalizing symptoms varies by generation in the United Kingdom.

Given the low rates of agreement between self-report and parent report (correlations of typically $\sim .2$; Patalay & Fitzsimons, 2018; Rescorla et al., 2013), we investigated whether or not the outcomes of the preregistered analyses (which focused on self-reported symptoms) differed when parent-reported symptoms were considered as the outcome variable. Across cohorts, parents tended to report fewer internalizing symptoms if their child had better language early on. This finding is in line with St Clair et al. (2019) and a recent meta-analysis (which did not differentiate studies on the basis of reporter; see Hentges et al., 2021). Thus, for the MCS, the direction of effect reversed when parent reports were considered instead of self-reports. Similar trends have been noted in the literature. For instance, a socioeconomic gradient is observed in parent-reported child mental health, but not in child-reported mental health (Johnston, Propper, Pudney, & Shields, 2014). In contrast, no significant differences by ethnic group are observed at age 14 in MCS cohort members based on parent report, but substantial ethnic differences are observed based on self-report (Patalay & Fitzsimons, 2018). Some previous research looking at this relation with general population samples has used an outcome measure where self- and parent-reported internalizing symptoms have been combined into one overall measure (Bornstein et al., 2013; Miller et al., 2020). However, it remains unclear what a combined measure of self- and parent-reported symptoms represents given their low correlation. The current findings, where the direction of effect differs as a function of reporter, suggest that studies with a combined outcome measure should be interpreted with caution.

Given the differences in direction of effect as a function of reporter, an important question is whether one reporter is more reliable for identifying internalizing symptoms in adolescence. There are strong arguments to be made for self-report measures. First, the sociodemographic patterns of self-report better match the latest national prevalence estimates based on clinical diagnoses, which are arguably the gold standard. This suggests that they better reflect population patterns in diagnosed

mental health difficulties, compared to symptoms reported by others (Sadler et al., 2018). Second, generally speaking, young people are competent reporters of their own mental health (Sharp et al., 2006) and it is argued that they should be considered the primary reporter when assessing internalizing mental health (The Good Childhood Report, 2019). Third, from a longitudinal perspective, self-report measures allow direct comparison with adult outcomes in the BCS (e.g., Schoon et al., 2010), where only self-reported measures are available, which is the norm in research on adult mental health. For these reasons, we preregistered the self-report measure as the primary outcome and, while the current direction of findings for the recent cohort is surprising, we consider it important to take seriously the possibility that good vocabulary is not straightforwardly predictive of good mental health.

The finding that better childhood abilities predict more internalizing symptoms in adolescence in the MCS is counterintuitive, and there are a number of possible explanations for this direction of effect. For example, academic pressure may have increased in recent years: schoolwork, examinations and feeling pressured are commonly reported stressors among adolescents (Gray, Galton, McLaughlin, Clarke, & Symonds, 2011). It is possible that language ability is positively associated with pressure to succeed academically, resulting in adolescents with more advanced language abilities having a higher risk of feeling stressed and experiencing poor mental health. Adding to possible increases in academic pressure is the widening of social and generational inequalities in Britain over this period (Corlett, Clarke, Mccurdy, Rahman, & Whittaker, 2019), which increases the importance of academic qualifications in achieving economic stability in adulthood (Green, Anders, Henderson, & Henseke, 2020).

Limitations and Strengths

This research used vocabulary as the sole measure of language ability. As a result, we might not have captured the kinds of language problems that lead to mental health difficulties. Recent research suggests that different measures of formal language tend to load on to the same factor (Fricke et al., 2017), so vocabulary is likely a good proxy for broader language ability. However, there is some evidence that pragmatic language skills cluster separately (Wilson & Bishop, 2019) and that they might be more directly related to mental health (Brenne & Rimehaug, 2019; Ketelaars, Cuperus, Jansonius, & Verhoeven, 2010). Likewise, we focused only on

one domain of mental ill-health—internalizing symptoms—and it might be that a different picture would emerge if externalizing symptoms were analyzed (e.g., Chow & Wehby, 2018).

While receptive and expressive vocabulary tend to be moderately to highly correlated (e.g., Conway et al., 2017), the difference in self-reported findings between cohorts could be attributed to the use of a receptive vocabulary measure in the BCS and an expressive vocabulary measure in the MCS. However, one would have thought that if the difference could be attributed to the use of different vocabulary measures, a similar cohort difference would emerge for the parent-reported outcome, which was not the case.

We were careful to include a robust set of confounders based on previous research, including childhood SES, biological risk factors, and both childhood and maternal mental health. However, we acknowledge that, given the weakness of the observed associations between early vocabulary and adolescent internalizing symptoms, it is possible that taking into account a strong unmeasured confounder could result in the associations disappearing.

As with any longitudinal data analysis, missing data had to be accounted for. Those with mental health difficulties in one sweep of cohort studies are less likely to take part in the next sweep. Furthermore, males, particularly of a lower SES, tend to be underrepresented in subsequent sweeps (Elliott & Shepherd, 2006; Mostafa & Wiggins, 2014). Therefore, missing data could introduce bias into the results. To combat this, missing data were accounted for using multiple imputations, which are considered a “best effort” approach (Little & Rubin, 2002). Although we have aimed to capture the full continuum of vocabulary abilities, higher rates of attrition may occur among those with language difficulties, and it is, therefore, possible that our results underestimate effects. However, we have imputed missing data to minimize bias due to attrition.

Finally, as with any study, it is likely that some measurement error is present (van Smeden, Lash, & Groenwold, 2020). However, we have no reason to expect any differential or multivariate error for our variables, or to expect large amounts of nondifferential error for the standardized, reliable measures we make use of.

Despite these limitations, the strengths of this research lie in the large and nationally representative samples with researcher-collected vocabulary measures, that make it possible to test the association between early vocabulary and later

internalizing symptoms, while taking into account important control variables. As such, findings are generalizable to the United Kingdom. However, our cross-cohort comparison revealed that this relation has changed between generations. It could therefore also differ as a function of cultural and socio-economic conditions across the globe, meaning our findings for contemporary adolescents may not be generalizable beyond the United Kingdom.

Finally, supplementary analyses allowed us to look at this relation when taking parents’ perspectives on their adolescent’s internalizing symptoms, and enabled us to look at the relation across the life course in the 1970 born cohort, by considering adulthood internalizing symptoms. The use of two nationally representative birth cohorts allowed us to compare this relation in two different generations born 30 years apart, during a period when mental health difficulties were on the rise.

Implications

There are several implications of this work. First, it has been claimed that early language ability is important for later mental health (Bercow, 2018). Empirical findings in support of this position would suggest a need for public health interventions to promote early language in the wider population rather than exclusively in clinical populations with language difficulties. However, our findings suggest that good early vocabulary does not necessarily protect adolescents from internalizing difficulties. Furthermore, where a relation does exist, effect sizes are small, and, for contemporary adolescents, in the opposite direction to that predicted. This research suggests that while public health interventions to promote early language are well founded for educational reasons (e.g., Fricke et al., 2017), caution is needed when looking for means to improve adolescent internalizing mental health.

Second, in line with Schoon et al. (2010), higher vocabulary scores in early childhood do appear to be related to better internalizing mental health in *adulthood*. However, it remains to be seen whether this is still the case in the more recent cohort. Given the absence of an analogous relation for adolescent internalizing mental health, it would appear that the link between early vocabulary and adult internalizing mental health might not be direct (through adolescent mental health); but might instead operate via education and labor market outcomes in early adulthood. Possible pathways need to be tested with future research.

Third, although the effect size for vocabulary in the contemporary cohort was small, the finding that better vocabulary ability in childhood was associated with *poorer* adolescent internalizing mental health should not be dismissed. Rather, potential adverse associations of cognitive ability and mental health should be entertained as a possibility in current generations, and reasons for such associations should be studied.

Finally, given the change in direction of effect as a function of reporter, it is vital to understand the measurement and reporting of adolescent internalizing mental health in greater detail. In the meantime, studies should ideally not be based solely on one reporter and reporter effects should be more actively considered.

Conclusion

The use of two cohort studies enabled us to test whether there is an association between early vocabulary and adolescent internalizing mental health, and if so, how any relation may have changed over 30 years. In the BCS (born in 1970), no relation was observed for self-reported adolescent internalizing mental health once controls were accounted for. In the contemporary generation (born ~2001), MCS data indicate that, if anything, better childhood vocabulary predicts *poorer* self-reported adolescent internalizing mental health, regardless of whether vocabulary was considered as a continuous or binary predictor. Thus, the relation between age 5 vocabulary ability and adolescent internalizing symptoms varies with generation. When parent-reported adolescent symptoms were considered, lower childhood vocabulary scores predicted *poorer* adolescent internalizing mental health in both cohorts. Therefore, the relation also varies as a function of reporter. In all analyses effect sizes were small. In sum, the relation between childhood vocabulary and adolescent internalizing symptoms varies by generation and reporter—good early language skills may not be protective for contemporary adolescents' internalizing mental health.

References

- Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. (1995). The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*, 237–249. Retrieved from [http://www.scalesandmeasures.net/file/s/files/Short%20Mood%20and%20Feelings%20Questionnaire%20\(1995\).pdf](http://www.scalesandmeasures.net/file/s/files/Short%20Mood%20and%20Feelings%20Questionnaire%20(1995).pdf)
- Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., . . . Thomas, S. M. (2008). The mental wealth of nations. *Nature*, *455*, 1057–1060. <https://doi.org/10.1038/4551057a>
- Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., . . . Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 75–82. <https://doi.org/10.1097/00004583-200101000-00019>
- Bercow. (2018). *Bercow: Ten years on*. Royal College of Speech and Language Therapists. Retrieved from <https://www.bercow10yearson.com/wp-content/uploads/2018/04/Bercow-Ten-Years-On-Summary-Report.pdf>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., . . . Whitehouse, A. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*, 1068–1080. <https://doi.org/10.1111/jcpp.12721>
- Bor, W., Dean, A. J., Najman, J., & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian and New Zealand Journal of Psychiatry*, *48*, 606–616. <https://doi.org/10.1177/0004867414533834>
- Bornstein, M. H., Hahn, C.-S., & Suwalsky, J. T. D. (2013). Language and internalizing and externalizing behavioral adjustment: Developmental pathways from childhood to adolescence. *Developmental Psychology*, *25*, 857–878. <https://doi.org/10.1017/S0954579413000217>
- Brenne, E., & Rimehaug, T. (2019). Pragmatic language impairment general and specific associations to mental health symptom dimensions in a child psychiatric sample. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *7*, 3–12. <https://doi.org/10.21307/sjcpp-2019-001>
- Brimer, M. A., & Dunn, L. M. (1962). *Manual for the English Picture Vocabulary Tests: Test 1 (age range 5: 0–8: 11), test 2 (age range 7: 0–11: 11)*. Educational Evaluation Enterprises.
- Chow, J. C., & Wehby, J. H. (2018). Associations between language and problem behavior: A systematic review and correlational meta-analysis. *Educational Psychology Review*, *30*, 61–82. <https://doi.org/10.1007/s10648-016-9385-z>
- Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic review and meta-analysis: Adolescent depression and long-term psychosocial outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*, 72–79. <https://doi.org/10.1016/j.jaac.2018.07.896>
- Connelly, R., & Platt, L. (2014). Cohort profile: UK Millennium Cohort Study (MCS). *International Journal of Epidemiology*, *43*, 1719–1725. <https://doi.org/10.1093/ije/dyu001>

- Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*, 516–525. <https://doi.org/10.1111/j.1469-7610.2007.01858.x>
- Conti-Ramsden, G., Mok, P. L. H., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities*, *34*, 4161–4169. <https://doi.org/10.1016/j.ridd.2013.08.043>
- Conway, L. J., Levickis, P. A., Mensah, F., McKean, C., Smith, K., & Reilly, S. (2017). Associations between expressive and receptive language and internalizing and externalizing behaviours in a community-based prospective study of slow-to-talk toddlers. *International Journal of Language and Communication Disorders*, *52*, 839–853. <https://doi.org/10.1111/1460-6984.12320>
- Corlett, A., Clarke, S., Mccurdy, C., Rahman, F., & Whitaker, M. (2019). *The living standards audit 2019* (Vol. 44). Retrieved from <https://www.resolutionfoundation.org/publications>
- Daviss, W. B., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*, 927–934. <https://doi.org/10.1111/j.1469-7610.2006.01646.x>
- Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales Second Edition (BAS II) early years*. London, UK: NFER-Nelson.
- Elliott, C. D., Smith, P., & McCulloch, K. (1997). *British Ability Scales-Second Edition (BAS II): Technical manual*. London, UK: Nelson.
- Elliott, J., & Shepherd, P. (2006). Cohort profile: 1970 British Birth Cohort (BCS70). *International Journal of Epidemiology*, *35*, 836–843. <https://doi.org/10.1093/ije/dyl174>
- Fergusson, D. M., Horwood, L. J., Ridder, E. M., & Beautrais, A. L. (2005). Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry*, *62*, 66–72. <https://doi.org/10.1001/archpsyc.62.1.66>
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234–248. <https://doi.org/10.1111/desc.12019>
- Forrest, C. L., Gibson, J. L., Halligan, S. L., & St Clair, M. C. (2020). A cross-lagged analysis of emotion regulation, peer problems, and emotional problems in children with and without early language difficulties: Evidence from the Millennium Cohort Study. *Journal of Speech, Language, and Hearing Research*, *63*, 1227–1239. https://doi.org/10.1044/2020_JSLHR-19-00188
- Fricke, S., Burgoyne, K., Bowyer-Crane, C., Kyriacou, M., Zosimidou, A., Maxwell, L., . . . Hulme, C. (2017). The efficacy of early language intervention in mainstream school settings: A randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*, 1141–1151. <https://doi.org/10.1111/jcpp.12737>
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Gray, J., Galton, M., McLaughlin, C., Clarke, B., & Symonds, J. (2011). *The supportive school*. Newcastle upon Tyne, UK: Cambridge Scholars.
- Green, F., Anders, J., Henderson, M., & Henseke, G. (2020). Private benefits? External benefits? Outcomes of private schooling in 21st century Britain. *Journal of Social Policy*, *49*, 724–743. <https://doi.org/10.1017/S0047279419000710>
- Harel, O. (2009). The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation. *Journal of Applied Statistics*. <https://doi.org/10.1080/02664760802553000>
- Hentges, R. F., Devereux, C., Graham, S. A., & Madigan, S. (2021). Child language difficulties and internalizing and externalizing symptoms: A meta-analysis. *Child Development*, 1–25. <https://doi.org/10.1111/cdev.13540>
- Institute of Child Health. (1970). *BCS70 birth sweep survey report*. Bristol, UK: University of Bristol. Retrieved from <https://cls.ucl.ac.uk/wp-content/uploads/2017/07/birth.pdf>
- Johnston, D. W., Propper, C., Pudney, S. E., & Shields, M. A. (2014). The income gradient in childhood mental health: All in the eye of the beholder? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *177*, 807–827. <https://doi.org/10.1111/rssa.12038>
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Groerger, J. C., Hiripi, E., . . . Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*, 184–189. <https://doi.org/10.1001/archpsyc.60.2.184>
- Ketelaars, M. P., Cuperus, J., Jansonius, K., & Verhoeven, L. (2010). Pragmatic language impairment and associated behavioural problems. *International Journal of Language and Communication Disorders*, *45*, 204–214. <https://doi.org/10.3109/13682820902863090>
- Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.
- McGillion, M., Pine, J. M., Herbert, J. S., & Matthews, D. (2017). A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*, 1122–1131. <https://doi.org/10.1111/jcpp.12725>
- McLaughlin, K. A., & King, K. (2015). Developmental trajectories of anxiety and depression in early adolescence. *Journal of Abnormal Child Psychology*, *43*, 311–323. <https://doi.org/10.1007/s10802-014-9898-1>
- Meng, X. L., & Rubin, D. B. (1992). Performing likelihood ratio tests with multiply-imputed data sets. *Biometrika*, *79*, 103–111. <https://doi.org/10.1093/biomet/79.1.103>

- Miller, A. B., Machlin, L., McLaughlin, K. A., & Sheridan, M. A. (2020). Deprivation and psychopathology in the Fragile Families Study: A 15-year longitudinal investigation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 62, 382–391. <https://doi.org/10.1111/jcpp.13>
- Mostafa, T., & Wiggins, R. D. (2014). *Handling attrition and non-response in the 1970 British Cohort Study*. CLS Working Paper 2014/2, Institute of Education, University of London, London, UK.
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 57, 1247–1257. <https://doi.org/10.1111/jcpp.12573>
- Osborn, A. F., Butler, N. R., & Morris, A. C. (1984). *The social life of Britain's five-year-olds: A report of the Child Health and Education Study*. London, UK: Routledge & Kegan Paul.
- Oxford University Press. (2018). *Why closing the word gap matters: Oxford language report*. Oxford, UK: Author.
- 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>
- Patalay, P., & Fitzsimons, E. (2018). Development and predictors of mental ill-health and wellbeing from childhood to adolescence. *Social Psychiatry and Psychiatric Epidemiology*, 53, 1311–1323. <https://doi.org/10.1007/s00127-018-1604-0>
- Patalay, P., & Gage, S. H. (2019). Changes in millennial adolescent mental health and health-related behaviours over 10 years: A population cohort comparison study. *International Journal of Epidemiology*, 48, 1650–1664. <https://doi.org/10.1093/ije/dyz006>
- Power, C., Atherton, K., Strachan, D. P., Shepherd, P., Fuller, E., Davis, A., . . . Stansfeld, S. (2007). Life-course influences on health in British adults: Effects of socioeconomic position in childhood and adulthood. *International Journal of Epidemiology*, 36, 532–539. <https://doi.org/10.1093/ije/dyl310>
- Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing Research*, 41, 688–700. <https://doi.org/10.1044/jslhr.4103.688>
- Redmond, S. M., & Rice, M. L. (2002). Stability of behavioral ratings of children with SLI. *Journal of Speech, Language, and Hearing Research*, 45, 190–201. [https://doi.org/10.1044/1092-4388\(2002/014\)](https://doi.org/10.1044/1092-4388(2002/014))
- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science and Medicine*, 90, 24–31. <https://doi.org/10.1016/j.socscimed.2013.04.026>
- Rescorla, L. A., Ginzburg, S., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., . . . Verhulst, F. C. (2013). Cross-informant agreement between parent-reported and adolescent self-reported problems in 25 societies. *Journal of Clinical Child and Adolescent Psychology*, 42, 262–273. <https://doi.org/10.1080/15374416.2012.717870>
- Robson, D. A., Allen, M. S., & Howard, S. J. (2020). Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review. *Psychological Bulletin*, 146, 324–354. <https://doi.org/10.1037/bul0000227>
- RStudio Team. (2020). *RStudio: Integrated development for R*. Boston, MA: RStudio, PBC. Retrieved from <http://www.rstudio.com/>
- Rubin, D. B. (1984). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. London, UK: Longman Publishing Group.
- Sadler, K., Vizard, T., Ford, T., Marchesell, F., Pearce, N., Mandalia, D., . . . Goodman, R. (2018). *Mental health of children and young people in England, 2017*. NHS Digital. Retrieved from <https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017>
- Salmon, K., O’Kearney, R., Reese, E., & Fortune, C. A. (2016). The role of language skill in child psychopathology: Implications for intervention in the early years. *Clinical Child and Family Psychology Review*, 19, 352–367. <https://doi.org/10.1007/s10567-016-0214-1>
- Schoon, I., Parsons, S., Rush, R., & Law, J. (2010). Children’s language ability and psychosocial development: A 29-year follow-up study. *Pediatrics*, 126(1), e73–e80. <http://doi.org/10.1542/peds.2009-3282>
- Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. *Journal of Abnormal Child Psychology*, 34, 379–391. <https://doi.org/10.1007/s10802-006-9027-x>
- Snowling, M. J., Bishop, D. V. M., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47, 759–765. <https://doi.org/10.1111/j.1469-7610.2006.01631.x>
- St Clair, M. C., Forrest, C. L., Yew, S. G. K., & Gibson, J. L. (2019). Early risk factors and emotional difficulties in children at risk of developmental language disorder: A population cohort study. *Journal of Speech, Language, and Hearing Research*, 62, 2750–2771. https://doi.org/10.1044/2018_JSLHR-L-18-0061
- Stolarova, M., Brielmann, A. A., Wolf, C., Rinker, T., Burke, T., & Baayen, H. (2016). Early vocabulary in relation to gender, bilingualism, type, and duration of childcare. *Advances in Cognitive Psychology*, 12, 130–144. <https://doi.org/10.5709/acp-0192-6>

- Sullivan A., Moulton V., Fitzsimons E. (2021). The inter-generational transmission of language skill. *The British Journal of Sociology*, 72(2), 207–232. <https://doi.org/10.1111/1468-4446.12780>
- Thabrew, H., Stasiak, K., Bavin, L. M., Frampton, C., & Merry, S. (2018). Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents. *International Journal of Methods in Psychiatric Research*, 27, 1–9. <https://doi.org/10.1002/mpr.1610>
- Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *Lancet*, 379, 117–130. [https://doi.org/10.1016/S0140-6736\(11\)60871-4](https://doi.org/10.1016/S0140-6736(11)60871-4)
- The Children's Society. (2019). *The good childhood report 2019*. London, UK: Author.
- Thorpe, K., Rutter, M., & Greenwood, R. (2003). Twins as a natural experiment to study the causes of mild language delay: I: Design; Twin-singleton differences in language, and obstetric risks. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44, 326–341. <https://doi.org/10.1111/1469-7610.00126>
- Trentacosta, C. J., & Fine, S. E. (2010). Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*, 19, 1–29. <https://doi.org/10.1111/j.1467-9507.2009.00543.x>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Van Smeden, M., Lash, T. L., & Groenwold, R. H. H. (2020). Reflection on modern methods: Five myths about measurement error in epidemiological research. *International Journal of Epidemiology*, 49, 338–347. <https://doi.org/10.1093/ije/dyz251>
- Wadman, R., Botting, N., Durkin, K., & Conti-Ramsden, G. (2011). Changes in emotional health symptoms in adolescents with specific language impairment. *International Journal of Language and Communication Disorders*, 46, 641–656. <https://doi.org/10.1111/j.1460-6984.2011.00033.x>
- Westrupp, E. M., Reilly, S., McKean, C., Law, J., Mensah, F., & Nicholson, J. M. (2019). Vocabulary development and trajectories of behavioral and emotional difficulties via academic ability and peer problems. *Child Development*, 91. <https://doi.org/10.1111/cdev.13219>
- Willner, C. J., Gatzke-Kopp, L., & Bray, B. C. (2017). The dynamics of internalizing and externalizing comorbidity across the early school years. *Developmental Psychopathology*, 28, 1033–1052. <https://doi.org/10.1017/S0954579416000687>
- Wilson, A. C., & Bishop, D. V. M. (2019). “if you catch my drift. . .”: Ability to infer implied meaning is distinct from vocabulary and grammar skills. *Wellcome Open Research*, 4, 68. <https://doi.org/10.12688/wellcomeopenres.15210.1>
- Yew, S. G. K., & O’Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 54, 516–524. <https://doi.org/10.1111/jcpp.12009>

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s website:

Appendix S1. Supplementary Materials