Investigating how vocabulary relates to different dimensions of family socioeconomic circumstance across developmental and historical time.

Emma Thornton[1]
Danielle Matthews[2]
Praveetha Patalay[3]
Colin Bannard[4]

Manchester Institute of Education, University of Manchester, United Kingdom.[1]
Department of Psychology, University of Sheffield, United Kingdom.[2]
Social Research Institute and Department of Population Science and Experimental Medicine, University College London, United Kingdom.[3]
Department of Linguistics and English Language, University of Manchester, United Kingdom.[4]

Abstract: Social inequalities in child vocabulary persist, despite decades of efforts to understand and reduce them. Different dimensions of socioeconomic circumstances (SEC), such as parent education, income, occupational status, wealth, and relative neighbourhood deprivation, are likely to represent different mechanisms of effects on child vocabulary. We investigate which aspects of SEC relate to vocabulary, and whether relations are stable over developmental and historical time. Data from two large, national datasets were analysed: the 1970 British Cohort Study (born 1970; N= 14,851) and the Millennium Cohort Study (born 2000-01; N=17,070). Substantial individual differences in vocabulary (ages 3–14) were explained by multiple indicators each making a unique contribution, most notably parent education (partial R^2:6.4%-8.5%), income (partial R^2: 4.3%-6.4%), and occupation (partial R^2: 5.3-8.1). Inequalities were generally stable over developmental and historical time. However, findings suggest a need to focus on widening inequalities at the start and end of compulsory schooling.

Keywords: Language, cognitive ability, child, adolescent, cross-cohort, generation, ontogeny, social inequality, social class.

Corresponding author(s): Emma Thornton, Manchester Institute of Education, University of Manchester, Oxford Road, Manchester, M13 9PL, UK. Email: Emma.Thornton@manchester.ac.uk

ORCID ID(s): https://orcid.org/0000-0003-4623-9538 (ET); https://orcid.org/0000-0003-3562-9549 (DM); https://orcid.org/0000-0002-5341-3461 (PP); https://orcid.org/0000-0001-5579-5830 (CB).

Citation: Thornton, E., Matthews, D., Patalay, P., & Bannard, C. (2024). Investigating how vocabulary relates to different dimensions of family socioeconomic circumstance across developmental and historical time. *Language Development Research, 4*(1), 80–174. https://doi.org/10.34842/mhqh-9g10

Introduction

Children need good language skills in order to be able to access education and, in turn, the labour market (Law, Charlton, & Asmussen, 2017; Oxford University Press, 2018). For decades, studies have observed social inequalities in vocabulary size (Hart & Risley, 1995; Pace, Luo, Hirsh-Pasek, & Golinkoff, 2017) and policy makers have sought educational interventions to reduce these disparities (Bercow, 2018). Yet randomised controlled trials suggest that such interventions have mixed success (Law, Charlton, Dockrell, et al., 2017). To assist in better directing future research and better targeting interventions, we address three fundamental questions using large, nationally representative, longitudinal UK datasets. First, are all indicators of socioeconomic circumstance (SEC) equal in predicting vocabulary outcomes? Second, does the relation between SEC and language development stay constant over developmental time? And third, is the relation between SEC and language development changing over historical time as our economy becomes increasingly knowledge-based and hourglass-shaped?

While caregiver education, occupational status, income, wealth, and neighbourhood disadvantage statistics are all often used as interchangeable indicators of SEC, each dimension reflects access to different resources that may affect language development (Duncan & Magnuson, 2012). Some have argued that caregiver education is the most relevant SEC indicator for language development as it is most directly related to the *quality* of the language learning environment and/or language related genetic factors (Hirsh-Pasek et al., 2015; Hoff, 2013; Hoff, Laursen, & Bridges, 2012). However, no empirical work has explicitly tested this claim in nationally representative samples and there are plausible pathways by which other indicators of SEC may also exert effects on vocabulary. First, income may affect language development through the availability of learning resources in the household (Duncan, Magnuson, & Votruba-Drzal, 2017; Washbrook & Waldfogel, 2011). Second, the family stress model posits that economic difficulty can influence parenting through its harmful effect on emotions, behaviours and relationships (Conger & Donnellan, 2007). This in turn can affect language development via the interactions parents have with their children (Perkins, Finegood, & Swain, 2013). Therefore, family wealth could be a protective mechanism, acting as a safeguard against any negative effects of sudden income losses, such as unexpected unemployment (Grinstein-Weiss, Williams Shanks, & Beverly, 2014; Killewald, Pfeffer, & Schachner, 2017). Third, occupational status reflects one's social position in the labour market, as well as power and status (Sullivan, Ketende, & Joshi, 2013). It is thought that people's social networks generally consist of people who are similar to them in terms of occupational status, known as occupational homophily. (Griffiths, Lambert, & Tranmer, 2011; McPherson, Smith-Lovin, & Cook, 2001). This may be indirectly related to language development, as children will adopt language used by their parents when talking to them and when talking to individuals in their social network (Sullivan, 2007). Finally, developmental theory emphasises how the immediate caregiving environment is nested within broader societal and cultural spheres (Bronfenbrenner, 1979; Rowe & Weisleder, 2020). As a proxy for this wider environment, neighbourhood-level statistics (such as the UK Indices of Multiple Deprivation) may additionally predict language development (Bennetts et al., 2022; Neuman, Kaefer, & Pinkham, 2018). Directly comparing the predictive value of different SEC indicators can help us understand why vocabulary inequalities exist and which mechanisms to further explore and target if aiming to support development. Our first goal was thus to test whether five key indicators of SEC (caregiver education, income, wealth, occupational status and neighbourhood deprivation) each predict unique variance in child vocabulary and how much relative variance they predict.

Compelling arguments have been made in favour of early intervention to prevent social disadvantage affecting language before children reach formal education (e.g., Doyle, Harmon, Heckman, & Tremblay, 2009), yet there is also evidence that the SEC gap in vocabulary is pronounced among adolescents (Spencer, Clegg, & Stackhouse, 2012; Sullivan & Brown, 2015). In fact, we do not know if or when the word gap shrinks or widens as children grow up. Nor do we know whether the predictive value of different SEC indicators remains stable over developmental time. For example, while caregiver education may be important during the early years, it has been proposed that family wealth may be a more important predictor of outcomes in adolescence and early adulthood. This might be because wealth facilitates access to high quality secondary education or other forms of academic support (Pfeffer, 2018). It is thus possible that the relative effect of different dimensions of SEC changes throughout development. Our second goal was therefore to test whether social disparities in language development have narrowed or widened over developmental time, from early childhood to mid-adolescence, for a contemporary generation born at the start of the 21st Century.

Large societal changes in the UK have seen an increase in the proportion of parents who have attended university, and a reconfiguration of the economy such that fewer people are in middle-ranked jobs, with more in lower grade employment on the one hand and in the higher managerial and professional occupations on the other (often characterised as a move to an hourglass economy; Bolton, 2012; Holmes & Mayhew, 2012). Many more jobs are now also knowledge-based, making language and cognitive skills of great importance for the UK economy (Beddington et al., 2008; Deloitte, 2016), and putting pressure on parents to support their children's cognitive development to open doors to the labour market. Income inequality increased in the UK in the 1980s and 1990s, and at the start of the millennium, income polarisation appeared to increase (those with the highest average incomes appeared to experience the largest increases, whilst those with lower average incomes experienced declines in their income; Dorling et al., 2007). These broad shifts in society have the potential to change the association between different measures of SEC and language development. Our third goal was thus to test whether the relations between different SEC indicators and language development have become more or less pronounced over historical time, comparing children born at the turn of this century with those born in 1970.

In a series of pre-registered analyses, we met the first two goals by analysing data from the Millennium Cohort Study (17,070 children born between 2000-02; MCS2001). We then compared these contemporary trends with those in a cohort born 30 years prior using data from the 1970 British Cohort Study (15,817children born in 1970; BCS1970, and 16,020 children in the MCS2001). Both studies contain measures of vocabulary at multiple ages and we use these as indicators of general language ability. Since different measures of formal language tend to load on to the same factor (Fricke et al., 2017), vocabulary is likely to be a good proxy for broader language ability. Nonetheless, an exclusive focus on vocabulary has implications for the conclusions we can draw, and we return to this in the discussion section.

Method

Data

We used data from two large nationally representative UK birth cohort studies: the Millennium Cohort Study (MCS2001 cohort) and the 1970 British Cohort Study (BCS1970 cohort). Addressing research questions 1-3 involved analyses of the MCS2001 cohort data only, due to the availability of multiple SEC indicators in this cohort, allowing us to examine the unique contribution of different SEC indicators to inequalities in language ability in a contemporary cohort. In addressing research question 4 we used data from the MCS2001 and BCS1970 cohorts in a cross-cohort comparison. The use of these two datasets for a cross-cohort comparison allowed us to examine inequalities in language ability in two generations born 30 years apart, during a period which has seen changes to occupational and educational structures in the UK.

MCS2001. The Millennium Cohort Study is a longitudinal birth cohort study of 19,518 young people, from 19,244 families, born across England, Scotland, Wales and Northern Ireland between 2000-02 (Connelly & Platt, 2014). To date there have been seven sweeps of data collection conducted when cohort members were aged 9 months and ages 3, 5, 7, 11, 14, and 17. More information on the MCS2001 cohort can be found here: https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/.

BCS1970. The 1970 British Cohort Study is a longitudinal birth cohort study of 16,571 children who were born during one week in 1970 in England, Scotland and Wales (Elliott & Shepherd, 2006). It has 4 childhood sweeps (data collected at birth and 5, 10 and 16 years). More information on the BCS1970 cohort can be found here: https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study/

Sample Selection. We selected all cohort members with a response on at least one of the language tasks at the time points considered – ages 3, 5, 11 or 14 (RQ 1-3, MCS2001 cohort only) and age 5, 10 or 16 (BCS1970) and ages 5, 11 or 14 MCS2001) for the cross-cohort comparison. Where cohort members were twins, triplets, or there were multiple cohort members from the same family, one of these members was selected at random.

Measures

Vocabulary Measures (MCS2001 Cohort Only)

The MCS2001 cohort members completed a battery of cognitive tests throughout childhood and into early adolescence. Full details about the completed vocabulary tests can be found in Appendix A.

At ages 3, 5 and 11, subscales of the British Ability Scale II (BAS II) were completed (Elliott, Smith, & McCulloch, 1996). The British Ability Scales consist of a series of tests measuring cognitive ability and educational attainment, between ages 2 years 6 months to 7 years 11 months. Progression through these tests depends on performance, and poor performance may result in a different, easier set of items being administered. Cohort members were born over a 1.5 year period (September 2000-January 2002) and assessed over a range of months, so age at the time of testing may differ between cohort members. Therefore, we used t-scores (as published in the data), which are adjusted for item difficulty and age. These were converted to z scores for analyses.

Ages 3 & 5. Cohort members completed the Naming Vocabulary BAS II subscale, as a measure of expressive vocabulary. Cohort members were shown a series of images and were asked to name each item in the image (Moulton et al., 2020).

Age 11. Cohort members completed the Verbal Similarities BAS II subscale. This is a measure of verbal reasoning and verbal knowledge. Sets of three words were read out to the cohort member, usually by the interviewer, and cohort members had to say how the words were related to each other (Moulton, 2020).

Age 14. Word Activity task. This test was a subset of items from the Applied Psychology Unit (APU) Vocabulary Test (Closs, 1986). Cohort members were given a list of 20 target words, each presented alongside 5 other words. Cohort members had to choose the word which meant the same, or nearly the same as the target word, from the 5 options (Moulton, 2020). Total scores out of 20 were converted into z scores for analyses.

Vocabulary Measures (Cross-Cohort Comparison)

For the cross-cohort comparison, we considered vocabulary at three time points in each cohort: age 5 (both cohorts; defined as early language ability), ages 10/11 (BCS1970 and MCS2001 cohorts respectively, referred to as late childhood language ability) and ages 16/14 (BCS1970 and MCS2001 cohorts respectively, referred to as adolescent language ability). There is no age 3 data for the BCS1970 cohort, hence the earliest language measure considered in the cohort comparisons is age 5.

Early Language Ability. For the BCS1970 cohort, receptive vocabulary was measured at age 5 using the English Picture Vocabulary Test (EPVT), a UK version of the Peabody Picture Vocabulary Test (Brimer & Dunn, 1962; Dunn, Dunn, Bulheller, & Häcker, 1965). Cohort members were shown 56 sets of four diverse images and heard a specific word associated with each set of four images. They were asked to select one picture that matched the presented word and were awarded one point for every correct response. For the MCS2001 cohort, expressive vocabulary was measured using the naming vocabulary sub-test of the BAS II (Elliott et al., 1996). We adjusted for age in months at the time of the test in both cohorts. All scores and ages were converted to z scores for analyses.

Late Childhood Language Ability. When the BCS1970 cohort members were aged 10, they completed the BAS word similarities subscale (Elliott, Murray, & Pearson, 1979). The test was made up of 21 items, each of which consisted of three words. The teacher read these sets of items out loud and cohort members had to a) name another word that was consistent with the three words in the item and b) state how the words were related. In order to receive a point, cohort members had to correctly answer both parts of the question (Moulton, 2020). Details on the scoring of this vocabulary measure and the SPSS syntax used can be found in appendix 3 of "Childhood Cognition in the 1970 British Cohort Study" (Parsons, 2014). When MCS2001 cohort members were aged 11, they completed the BAS II verbal similarities subscale (detailed above). As already mentioned, test scores for the MCS2001 cohort were adjusted for item difficulty. In both cohorts, we controlled for age at the time of the test and converted all scores to z scores.

Adolescent Language Ability. When aged 16, BCS1970 cohort members completed the APU Vocabulary Test (Closs, 1986). This consisted of 75 items: an item consisted of a target word, presented with a multiple-choice list, from which cohort members had to select a word that meant the same as the target word (Moulton, 2020). These items got progressively harder throughout the test. Details on the scoring of this vocabulary test can be found in appendix 3 (Parsons, 2014). When MCS2001 cohort members were aged 14, they completed the Word Activity Task (detailed above). Words used in the Word Activity Task were a subset of the words used in the BCS1970cohort Vocabulary Test, which cohort members completed aged 16 (Moulton, 2020). Scores were adjusted for age and converted to z scores for analyses.

Measures Of Socioeconomic Position (MCS2001 Cohort Only)

Five indicators of family SEC were used: parent education, family income, wealth, occupational status, and relative neighbourhood deprivation. Operationalisation of these variables is as follows:

Parent Education. As a measure of parent's education when cohort members were aged 3, highest parent NVQ (National Vocational Qualification) level was used (both academic and vocational qualifications derived into NVQ levels 1-5, with level 5 equating to higher qualifications). It is worth noting that the NVQ levels derived in MCS2001 data differ from those defined by the UK Government (https://www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels). In the MCS2001 data, these are:

NVQ level 0: none of these/other qualifications
NVQ level 1: GCSE grades D-G, NVQ/ SVQ/ GSVQ level 1
NVQ level 2: GCSE grades A-C, trade apprenticeships, NVQ/ SVQ/ GSVQ level 2
NVQ level 3: A/ AS/ S levels, NVQ/ SVQ/ GSVQ level 3
NVQ level 4: first degree, diplomas in higher education, professional qualifications at degree level
NVQ level 5: higher degree

To contextualise for readers not familiar with the UK system, GCSEs (or the Scottish equivalent) are subject-specific qualifications. The majority of children will take 9 GCSEs in the academic year they turn 16. A-levels are also subject specific and most people continuing in school on an academic route will specialise to take three subjects at the age of 18. A range of non-vocational qualifications are available at both stages, yielding the mapping noted above. We compared how well *maternal education* and *highest household education* (i.e., the educational qualification of the most qualified parent in the household) predicted vocabulary at each age (see Appendix B) and, based on findings that highest household education consistently accounted for the most variance in vocabulary at each age, we use a measure of highest parent education in our analyses.

Family Income. Here we used UK OECD weighted income quintiles at child age 3 (an indication of household income 1=lowest, 5=highest, accounting for family size). If data was missing, OECD weighted income quintiles at child age 9 months were used instead.

Wealth. Here we used a measure of total net wealth, taken from the age 11 sweep of the MCS2001 cohort — when cohort members were aged 11, parents reported on their savings and assets, total debts owed, the value of their house and the amount of outstanding mortgage owed on their home for the first time. This measure was derived from 4 variables: amount outstanding on all mortgages, house value, amount of investments and assets, and amount of debts owed. Outstanding mortgages were subtracted from the house value, to give a measure of housing wealth. In cases where families were not homeowners, they were given a housing wealth value of 0. Debts owed were taken from the amount of investments and assets, to give a measure of financial wealth. In cases where families reported having no savings or debts, they were given a financial wealth value of 0. Housing wealth and financial wealth were then summed to give an overall measure of total net wealth. Our measure of wealth was heavily positively skewed, in line with the distribution of wealth in the general population, which is heavily influenced by extreme values of the top 1% (Killewald, 2017). Total net wealth was therefore split into quintiles for our analyses.

Occupational Status. Here we used the highest household occupational status (National Statistics Socioeconomic Classification (NS-SEC) 3 categories: higher managerial; intermediate; and routine, with a fourth category for those who were unemployed) at child age 3 years. If data were missing, occupational status at child age 9 months was used instead.

Relative Neighbourhood Deprivation. Indices of multiple deprivation (IMD) are the government official measure of relative deprivation (Mclennan et al., 2019). Based on an individual's postcode (at the level of the street), these are used to rank small areas or neighbourhoods in England, Scotland, Wales, and Northern Ireland from the least deprived to the most deprived area. The IMD is a broad conceptualisation of deprivation, including a wide variety of living circumstances, rather than just a lack of income for adequate financial resources, which often defines people living in poverty. However, people can be considered deprived if they do not have access to any type of resource, not just income (Mclennan, 2019). Therefore, we used IMD deciles at child age 3 (with 1= most deprived and 10=least deprived) as a measure of relative neighbourhood deprivation.

Measures Of Socioeconomic Position (Cross-Cohort Comparison)

The SEC indicators used in RQ1-RQ3 include the full set of five SEC indicators (parent education, income, wealth, occupational status, and neighbourhood deprivation), enabling us to consider the multi-faceted nature of SEC. However, they are not all directly comparable to the data available in the BCS1970 cohort. Therefore, for RQ4, we used a subset of SEC indicators to ensure comparability, to the best of our ability, across the two cohorts. Harmonisation of these measures can be found in Table 1; data harmonisation is the process of making data from different sources (such as different cohorts) more similar to improve comparability between cohorts (O'Neill, Kaye, & Hardy, 2020).

Parent Education. The highest academic qualification achieved by a parent in the household when the cohort member was aged 5. Where this information is missing, information from previous sweeps was used.

Occupational Status. Highest household occupational status at child age 5. For the BCS1970 cohort, this was ascertained with the Registrar General's classification. For the MCS2001 cohort, the NS-SEC classification system was used. Where this information is missing, information from previous sweeps was used.
Family Income. UK OECD weighted income quintiles at child age 10 (BCS1970) and 11 (MCS2001) were used as an indication of household income 1=lowest, 5=highest, accounting for family size). The BCS1970 first measured family income when cohort members were aged 10, hence we take this information from the age 10 (BCS1970) and age 11 (MCS2001) sweeps for the cross-cohort comparison.

Potential Confounders

We adjusted for gender (male= 0, female=1), ethnicity and whether English was spoken as an additional language (EAL) in the home (1= only English, 2=English and another language, 3=Only another language). Harmonisation of these measures for RQ4 can be found in Table 1.

Data Analysis

All analyses were pre-registered on the Open Science Framework website (https://osf.io/482zw/).

Missing Data Strategy.

Missing data in all analyses was accounted for with multiple imputation using chained equations with the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011).

Analysis of MCS2001 Cohort Only. Each dataset was imputed 25 times, as this was greater than the percentage of missing data (10.6%)(White, Royston, & Wood, 2011). There was no missing data for gender or neighbourhood deprivation, and the percentage of missing data was less than 1% for ethnicity and EAL status. 14.71% of vocabulary scores at age 3 were missing, 12.41% of age 5 vocabulary scores were missing, 23.92% of age 11 vocabulary scores were missing, and 36.88% of age 14 vocabulary scores were missing. Full proportions of missing data can be found in Appendix C We conducted a series of sensitivity checks whereby we repeated the analyses on a dataset which had complete cases for vocabulary at ages 3, 5, 11 and 14 Missing data among the components of our wealth variable were also high (30.73% (outstanding mortgage); 27.57% (house valuation); 39.85% (total savings); and 28.99 (total debts owed). We therefore conducted sensitivity analyses where we considered all cohort members with a response to at least one wealth component variable and at least two wealth variables. Overall, these sensitivity checks revealed a similar pattern of results to the main analyses; results are available upon request. Combined sampling and attrition weights were applied to the data to account for the stratified clustered design of MCS2001 cohort data and the oversampling of subgroups, as well as for missing data due to attrition.

Cross-Cohort Comparison. Each dataset was again imputed 25 times, as this was greater than the percentage of missing data in each cohort (6.7% MCS2001 cohort, 21.3% BCS1970 cohort (White, Royston, & Wood, 2011). For the MCS2001 cohort, 6.67% of age 5 vocabulary scores were missing, 18.93% of age 11 vocabulary scores were missing, and 32.74% of age 14 vocabulary scores were missing. For the BCS1970 cohort, 20.12% of age 5 vocabulary scores were missing, 6.89% of age 10 vocabulary scores were missing, and 63.92% of age 16 vocabulary scores were missing (as a result of the teachers strike in 1986). Full proportions of missing data in both cohorts can be found in Appendix C. Again, combined sampling and attrition weights available in MCS2001 data were applied to data from this cohort. The BCS1970 cohort does not have the same sample design as the MCS2001cohort and thus sample weights are not necessary. However, attrition weights to account for non-response between birth and age 5 were created and included in analyses for BCS1970 cohort data (Appendix D for details).

Analyses

Analytic Sample. To address the first two research questions in a contemporary cohort, we analysed the data of 17,070 children in the MCS2001 (all cohort members with a response on at least one of the language tasks at ages 3, 5, 11 or 14). 49.05% of cohort members were female, 85.97% were of White ethnicity and 88.49% did not speak English as an additional language. Demographic differences between the children included in the analytic samples for Research Questions 1-3 and the full MCS cohort are negligible (see Table S2, Appendix E).

For the cross-generation comparison, we analysed the data of 14,851children in the BCS1970, and 16,020 children in the MCS2001 with harmonised measures (cohort members with a response on at least one vocabulary task administered in early childhood, late childhood and/or adolescence; see Table 1 for details of harmonisation). 49.45% of BCS1970 cohort members were female, 93.52% were of White ethnicity and 94.97% did not speak English as an additional language. In the cross-cohort comparison, 48.67% of MCS cohort members were female, 86.03% were of White ethnicity and 88.64% did not speak English as an additional language. Demographic differences between the children included in the analytic samples for Research Question 4 and the full MCS2001 and BCS1970 cohorts were also negligible (see Table S3, Appendix E).

Descriptive Statistics. Descriptive statistics were calculated across the 25 imputed datasets. Analytical samples were compared to the full cohort samples to see if there were any differences in characteristics of those included in the analyses. Mean language scores for each SEC group are reported (see Table 2).

Inequalities in vocabulary at ages 3, 5, 11 and 14: what is the variation captured by each indicator of SEC individually? Language scores at ages 3, 5, 11 and 14 were considered as separate outcome variables. For each age, separate models with each SEC predictor in turn (parent education, income, wealth, occupational status, and neighbourhood deprivation, each in a separate model) were built to assess the unadjusted relationship between each predictor and language at each time point. Potential confounding variables were then added to each of the models.

A drop-one analysis was used to assess the unique contribution of each predictor; a model with all 5 SEC predictors was compared to models with each predictor removed in turn. This was done for each age (3, 5, 11 and 14). Improvements in fit were assessed using model comparisons for imputed data, using the method of Meng and Rubin (Meng & Rubin, 1992). If the five-predictor model was a better fit to the data than the four-predictor model following the removal of an SEC indicator, then the SEC variable that was dropped can be said to account for significant unique variance in language ability at that age. Partial R^2 values for each SEC indicator are reported, indicating the proportion of variance explained by each SEC predictor, above that of the potential confounding variables.

How does a composite measure of overall socioeconomic position perform relative to individual measures and combinations of measures? A latent composite factor of

SEC was created using confirmatory factor analysis (see Appendix F for details). This composite factor was then included as the predictor variable in four separate regression models (each one considering vocabulary at each age), adjusting for the potential confounding variables. Relative AIC values were used to compare the marginal predictive value of each SEC predictor. These were calculated for each imputed dataset for each single-predictor model, the composite model and a model with all indicators included simultaneously (Schomaker & Heumann, 2014), and means and confidence intervals of these values across the imputed datasets are reported. This allowed us to consider whether the composite measure provides an equivalent or better fit to the data, compared to all predictors included simultaneously, and in relation to each individual predictor.

How does the relationship between SEC measures and vocabulary change over developmental time? (Vocabulary at ages 3, 5, 11 and 14). Here we addressed whether or not one's position in the language distribution changes at each age, and how much of this is a function of SEC. The models from RQ1 were used to answer this question. Due to the different measures of language ability available at each age, we were unable to model longitudinal changes in language development. However, because the outcome variable of language ability at each age is standardised to the same scale, the coefficients are directly comparable. We also compared the standardised coefficients from the models in RQ2, which consider our composite factor of SEC, allowing us to establish the best predictor across developmental time.

How has the relationship between SEC measures and vocabulary changed with historical time? (Comparison of two nationally representative cohorts, born 30 years apart). We had 3 separate outcome variables in each cohort (early childhood language ability, late childhood language ability, and adolescent language ability). We built three regression models per outcome, one with occupational status as the predictor variable, one with parent education as the predictor variable, and finally, one with family income as the predictor variable. Because our measures of language ability were standardised within each cohort, we were able to directly compare coefficients between cohorts and establish the rate of inequality in language ability at each age in the two cohorts.

Table 1. ***Cross-cohort harmonisation of variables***

Measure	BCS1970	MCS2001	Harmonised
Age 5 language ability	EPVT. Continuous measure.	Naming vocabulary. Continuous measure.	Total vocabulary score: continuous cohort specific standardised *z* score
Late childhood language ability	Age 10. BAS word similarities	Age 11. BAS II verbal similarities	Total vocabulary score: continuous cohort specific standardised *z* score
Adolescent language ability	Age 16. Vocabulary Test	Age 14. Word activity task,	Total vocabulary score: continuous cohort specific standardised *z* score. Note that a harmonised version of the BCS1970 Vocabulary Test with the same words included in the MCS2001 Word activity task was also created, however this correlated 0.93 with the full BCS1970 measure, so we did not conduct this sensitivity analysis.
Occupational status at birth	Age 5. Registrar General's classification. 5 classes: 1. professional 2. managerial, other professionals 3. non-manual skilled, skilled manual 4. semi-skilled workers 5.unskilled workers 6. Full/part time students or volunteers with no paid employment	Age 5. NS-SEC 5 classes: 1. Higher managerial/admin/professional 2. intermediate 3. small employers/self-employed 4. lower supervisory and technical occupations 5. semi-routine and routine	Composite variable, with a 4th category for unemployment: BCS1970: Professional & Managerial Skilled Semi-skilled and unskilled Unemployed

	Note: students/volunteers were categorised as unemployed as they have no paid employment.	This 5-class version was collapsed into a 3-class version, as shown here: https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010#classes-and-collapses	MCS2001: Higher managerial Intermediate Routine Unemployed *Note: The convention used in the MCS2001 was used for the occupational status variables from both cohorts, for ease.*
Parental education: highest educational qualification (highest household level)	No qualifications Vocational qualifications O levels A-levels State registered nurse Certificate of education Degree +	None of these qualifications GCSE grades D-G O level/GCSE grades A-C A/AS/ S Levels Diplomas in higher education First degree Higher degree Other academic qualifications (incl.overseas)	No qualifications/low level qualifications O levels/GCSE grades A*-C A levels/earning a degree – post 16 education university level qualifications
Family Income	Weekly Income Bands (midpoint for each band) (Age 10) Under £35 pw (£17) £35 - £ 49 pw (£42) £50-£99 pw (£74.50) £100 - £149 pw (124.50)	Annual Income Bands (midpoint for each band) (Age 11) < £ 3,000 (£1500) £3,000- £6,999 (£5000)	OECD equivalisation was applied to the midpoint of each income band in each cohort separately, and these equivalized values were converted into quintiles to give OECD equivalised quintiles:

	£150 - £199 pw (174.50)	£ 7,000 - £ 10,499 (£8750)	
	£200 - £ 249 pw (224.50)	£ 10,500 - £ 12,499 (£11500)	Quintile 1 (Most Deprived)
	> £250 pw (£275)	£ 12,500 - £ 13,999 (£13250)	Quintile 2
		£ 14,000 - £ 14,999 (£14500)	Quintile 3
		£ 15,000 - £ 19,499 (£17250)	Quintile 4
		£ 19,500 - £ 23,499 (£21500)	Quintile 5 (Least Deprived)
		£ 23,500 - £ 27,499 (£25500)	
		£ 27,500 - £ 30,499 (£29000)	
		£ 30,500 - £ 34,499 (£32500)	
		£ 34,500 - £ 39,999 (£37250)	
		£ 40,500 - £ 47,999 (£44250)	
		£ 48,000 - £ 53,999 (£51000)	
		£ 54,000 - £ 62,999 (£58500)	
		£ 63,000 - £ 82,999 (£73000)	
		£ 83,000 - £ 114,999 (£99000)	
		£ 115,000 - £ 149,999 (£132500)	
		more than 150,000 (£150000)	
Ethnicity	European UK	White	Categorical measures collapsed into 0=White, 1=Minority
	European Other	Mixed	
	West Indian	Indian	
	Indian-Pakistani	Pakistani and Bangladeshi	
	Other Asian	Black or Black British	
	African	Other Ethnic group (incl. Chinese, Other)	
	Other		

Language spoken at home	English Welsh-Gaelic Hindi-Urdu Greek-Turkish Chinese-Oriental African Language European Language	Yes - English only Yes - English and other language(s) No - other language(s) only	Categorical measures collapsed into 0= Monolingual English 1= Other language

Results

Which SEC Measures Predict Child Vocabulary?

As can be seen in Table 2, for every SEC measure, the mean vocabulary score is greater with each increase in SEC group, with the highest mean vocabulary scores in the highest SEC group.

To assess the unique contribution of each predictor at each age, a model with all five SEC predictors was compared to models with each predictor removed in turn. Improvements in fit were assessed using model comparisons for imputed data, using the method of Meng and Rubin (Meng & Rubin, 1992). This drop-one analysis revealed that caregiver education, income, wealth, and occupational status accounted for significant unique variance in vocabulary at all ages (see Appendix G). Neighbourhood statistics accounted for significant variance in vocabulary at ages 3, 5 and 11.

Figure 1 presents partial R^2 values indicating the proportion of variance explained by each SEC predictor, above that of potential confounding variables (sex, ethnicity, and whether English is spoken as an Additional Language (EAL) in the home). Caregiver education explains the largest proportion of variance in vocabulary at each age (between 6.4% and 8.5% of variance), closely followed by income and occupational status, and at ages 11 and 14, wealth. Relative neighbourhood deprivation consistently contributes the least variance in vocabulary scores, regardless of age.

Reducing individual indicators to a single composite factor may afford us an efficient way of communicating and understanding inequalities in vocabulary but we do not yet know whether such composites explain more variance than certain SEC indicators considered alone, and/or are equivalent to models with each predictor considered separately. Confirmatory factor analysis was therefore used to create a composite variable of SEC (see Appendix F), which was then included as the predictor in an adjusted model predicting language at ages 3, 5, 11 and 14. Regardless of age, compared to each individual measure, the composite factor was a better fit to the data (see Table S4 in Appendix H), and explained 7.4-10.2% of variance in language across ages. However, a model with each SEC measure included simultaneously explained more variance than a model with just the composite measure and control variables (see Table S5 in Appendix H). This indicates that if one needs to identify a single variable for use in analyses, then a composite variable would be a better choice than any of the original individual predictors. In the absence of such a constraint, including a set of multiple predictors would be preferable.

Table 2: ***Means (±SD) and 95% CIs for language scores in each SEC group at each age (MCS2001 cohort)***

	Proportion (%) or Mean(±SD) [95% CIs]			
SEC Indicator	*Age 3 Vocabulary*	*Age 5 Vocabulary*	*Age 11 Vocabulary*	*Age 14 Vocabulary*[1]
Parent Education				
Parent education (NVQ1)	45.24(10.28) [44.61;45.87]	49.78(10.51) [49.14;50.43]	54.97(10.14) [54.35;55.6]	6.12(2.38) [5.97;6.27]
Parent education (NVQ2)	47.91(10.63) [47.59;48.23]	52.79(10.29) [52.48;53.1]	56.83(9.9) [56.53;57.12]	6.53(2.35) [6.46;6.6]
Parent education (NVQ3)	49.62(10.64) [49.23;50.01]	54.24(10.14) [53.86;54.61]	58.36(9.35) [58.01;58.7]	6.81(2.43) [6.72;6.9]
Parent education (NVQ4)	52.35(10.74) [52.07;52.63]	57.54(10.18) [57.28;57.81]	60.76(8.97) [60.53;60.99]	7.57(2.65) [7.5;7.64]
Parent education (NVQ5)	53.47(11.47) [52.82;54.11]	59.56(10.48) [58.97;60.14]	63.26(8.66) [62.77;63.74]	8.53(2.9) [8.37;8.69]
Parent education (none of these/overseas)	41.3(11.55) [40.79;41.8]	46.4(11.66) [45.9;46.91]	54.11(10.9) [53.64;54.58]	5.96(2.27) [5.86;6.06]
Income				
Income (Quintile 1)	44.26(11.49) [43.9;44.62]	49.45(11.3) [49.1;49.8]	55.7(10.62) [55.37;56.03]	6.28(2.35) [6.2;6.35]
Income (Quintile 2)	47.31(11.09) [46.99;47.64]	52.19(10.71) [51.88;52.5]	57.05(9.83) [56.76;57.33]	6.67(2.46) [6.6;6.75]
Income (Quintile 3)	51.18(10.65) [50.83;51.54]	55.97(10.18) [55.63;56.31]	59.05(9.35) [58.74;59.36]	7.08(2.54) [7;7.17]
Income (Quintile 4)	52.58(10.38) [52.22;52.94]	57.44(10.06) [57.1;57.79]	60.37(9.21) [60.05;60.69]	7.51(2.69) [7.42;7.61]
Income (Quintile 5)	53.65(10.32) [53.19;54.12]	59.48(9.78) [59.04;59.92]	62.64(8.46) [62.26;63.02]	7.99(2.79) [7.86;8.12]
Wealth				
Wealth (Quintile 1)	46.5(11.05) [46.19;46.82]	51.55(10.68) [51.25;51.86]	56.09(10.18) [55.8;56.38]	6.52(2.44) [6.45;6.59]

Wealth (Quintile 2)	46.71(11.29) [46.23;47.19]	51.49(11.11) [51.02;51.96]	56.56(10.15) [56.13;56.99]	6.48(2.4) [6.38;6.58]
Wealth (Quintile 3)	49.63(11.2) [49.26;50.01]	54.31(10.76) [53.95;54.67]	58.64(9.51) [58.32;58.96]	6.93(2.5) [6.85;7.02]
Wealth (Quintile 4)	50.75(11.18) [50.37;51.12]	55.68(10.75) [55.32;56.04]	59.59(9.58) [59.27;59.91]	7.16(2.57) [7.08;7.25]
Wealth (Quintile 5)	52.54(10.99) [52.17;52.91]	58.09(10.59) [57.74;58.45]	61.49(8.96) [61.19;61.79]	7.78(2.8) [7.69;7.88]
Occupational Status				
Occupational Status (Unemployed)	44.18(11.07) [43.82;44.54]	48.91(10.9) [48.56;49.27]	55.03(10.61) [54.69;55.38]	6.21(2.4) [6.13;6.29]
Occupational Status (Routine)	47.33(11.09) [46.99;47.67]	52.21(10.7) [51.88;52.54]	56.82(9.92) [56.52;57.13]	6.57(2.38) [6.5;6.65]
Occupational Status (Intermediate)	50.12(10.97) [49.74;50.5]	54.67(10.63) [54.3;55.04]	58.7(9.42) [58.38;59.03]	6.88(2.46) [6.8;6.97]
Occupational Status (higher managerial)	52.75(10.64) [52.48;53.01]	58.28(9.96) [58.03;58.53]	61.28(8.87) [61.06;61.5]	7.74(2.71) [7.67;7.8]
Relative Neighbourhood Deprivation				
Relative neighbourhood deprivation (most deprived)	43.7(11.64) [43.28;44.13]	48.69(11.2) [48.27;49.1]	54.91(10.6) [54.52;55.3]	6.27(2.39) [6.18;6.36]
Relative neighbourhood deprivation (10 - < 20%)	45.77(11.82) [45.3;46.25]	50.54(10.97) [50.09;50.98]	57.07(10.08) [56.67;57.48]	6.59(2.43) [6.49;6.69]
Relative neighbourhood deprivation (20 - < 30%)	48.01(11.1) [47.53;48.5]	53.13(10.6) [52.66;53.59]	57.64(9.94) [57.2;58.07]	6.74(2.54) [6.63;6.85]
Relative neighbourhood deprivation (30 - < 40%)	49.07(11.21) [48.54;49.61]	53.77(10.53) [53.27;54.27]	58.38(10.08) [57.9;58.86]	6.88(2.58) [6.76;7]
Relative neighbourhood deprivation (40 - < 50%)	49.56(10.97) [49;50.12]	54.49(10.89) [53.94;55.04]	58.38(9.12) [57.92;58.84]	6.95(2.53) [6.82;7.08]
Relative neighbourhood deprivation (50 - < 60%)	50.5(10.92) [49.93;51.06]	55.55(10.47) [55.01;56.1]	58.89(9.92) [58.37;59.4]	7.04(2.54) [6.91;7.17]

Relative neighbourhood deprivation (60 - < 70%)	51.48(10.58) [50.88;52.08]	56.35(10.37) [55.76;56.94]	60.16(9.96) [59.59;60.72]	7.25(2.7) [7.09;7.4]
Relative neighbourhood deprivation (70 - < 80%)	52.14(10.49) [51.56;52.72]	57.49(10.57) [56.91;58.08]	60.15(9.03) [59.65;60.65]	7.5(2.67) [7.35;7.65]
Relative neighbourhood deprivation (80 - < 90%)	52.19(10.33) [51.64;52.73]	57.55(10.2) [57.01;58.09]	60.16(9.08) [59.68;60.64]	7.48(2.57) [7.34;7.61]
Relative neighbourhood deprivation (least deprived)	53.61(9.94) [53.09;54.13]	58.93(9.55) [58.43;59.43]	61.45(8.68) [61;61.9]	7.75(2.79) [7.6;7.89]

[1]Note: different standardised vocabulary tests were used at different ages, hence the lower mean score at 14 years.

Does the relationship between SEC and child vocabulary change over developmental time from age 3 to 14 years?

Figure 2 shows the relationships between each SEC indicator and vocabulary at each age (coefficients and 95% CIs plotted; see also Table S6, Appendix I). Because vocabulary scores were converted into z scores, the coefficients indicate the change in vocabulary in units of standard deviation (SD) associated with different levels of each predictor. A steeper slope indicates greater inequalities. Inequalities in vocabulary size are consistently narrowest at age 3, and widen by age 5. They then persist throughout childhood and into adolescence, regardless of the SEC indicator used. The relation between SEC and age 14 vocabulary displays a discontinuity not seen for the other ages, with the line appearing shallow for the lower SEC groups and steeper between the higher SEC groups. It is nonetheless clear that across childhood, inequalities in vocabulary have not substantially changed in this cohort; gaps in vocabulary size have not narrowed over time.

Given that the SEC measures used in the above analyses were collected when cohort members were aged 3, it is plausible that this pattern of results is due to the proximity of the SEC measures to the developmental stage at which vocabulary was measured. Therefore, we conducted a sensitivity analysis with age 14 SEC indicators predicting age 14 vocabulary. Overall, despite some inequalities appearing to be wider based on age 14 SEC measures, the proximity of the SEC measure to age 14 vocabulary does not affect the main pattern of results (see Appendix J).

Does the relationship between SEC and child vocabulary change with historical time?

The caregivers of children in the MCS2001 cohort are noticeably different to those of the BCS1970 cohort when compared on the basis of the SEC measures available for both cohorts. More parents of the BCS1970 cohort held no or low-level qualifications compared to parents of the MCS2001 cohort (which is to be expected given changes in the age of compulsory schooling; see Table 3). Furthermore, proportionally more parents from the BCS1970 cohort were in intermediate occupations, whereas more parents from the MCS2001 cohort were in either routine or higher managerial occupations (which is expected given that the UK is becoming more of an hourglass economy; see Table 3; Holmes & Mayhew, 2012). For all SEC measures, the mean vocabulary score was greater with each increase in SEC group in both cohorts, with a higher mean score in the highest SEC groups (see Table S9, Appendix K).

As can be seen in Figure 3, vocabulary scores generally increased with SEC regardless of indicator and cohort (also see Table S10, Appendix K). The overall picture is thus one of continuity of social inequality across the generations. Nonetheless, compared to their BCS1970 counterparts, MCS2001 cohort members whose parents had university level qualifications were at a clearer advantage in terms of their language ability in early childhood and adolescence. In contrast, inequalities in vocabulary based on occupational status and income are wider for the BCS1970 cohort at all ages, as indicated by the steeper slopes for this cohort. As can be seen from partial R^2 values (Figure 4), inequalities are substantial in both cohorts. There is no evidence of a decrease in SEC inequalities over the 30-year period and there is even some evidence that inequalities may have widened in early childhood, with SEC indicators explaining more variance in the MCS2001 cohort for this age point. Whereas for the BCS1970 cohort SEC indicators explained most variance in late childhood, for the contemporary MCS2001 cohort, SEC indicators explained most variance in early childhood.

To examine whether our findings were robust to changes in the distribution of education and occupation measures or to the ethnic composition of the UK during the period separating the BCS1970 and MCS2001 cohorts, we conducted two sensitivity checks. First, highest household occupational status and highest household educational attainment were converted to Ridit scores to aid comparability across cohorts (see Appendix L; Donaldson, 1998). Second, we restricted our analyses to those of a White ethnicity only (see Appendix M). Neither analysis resulted in a change in the pattern of results observed.

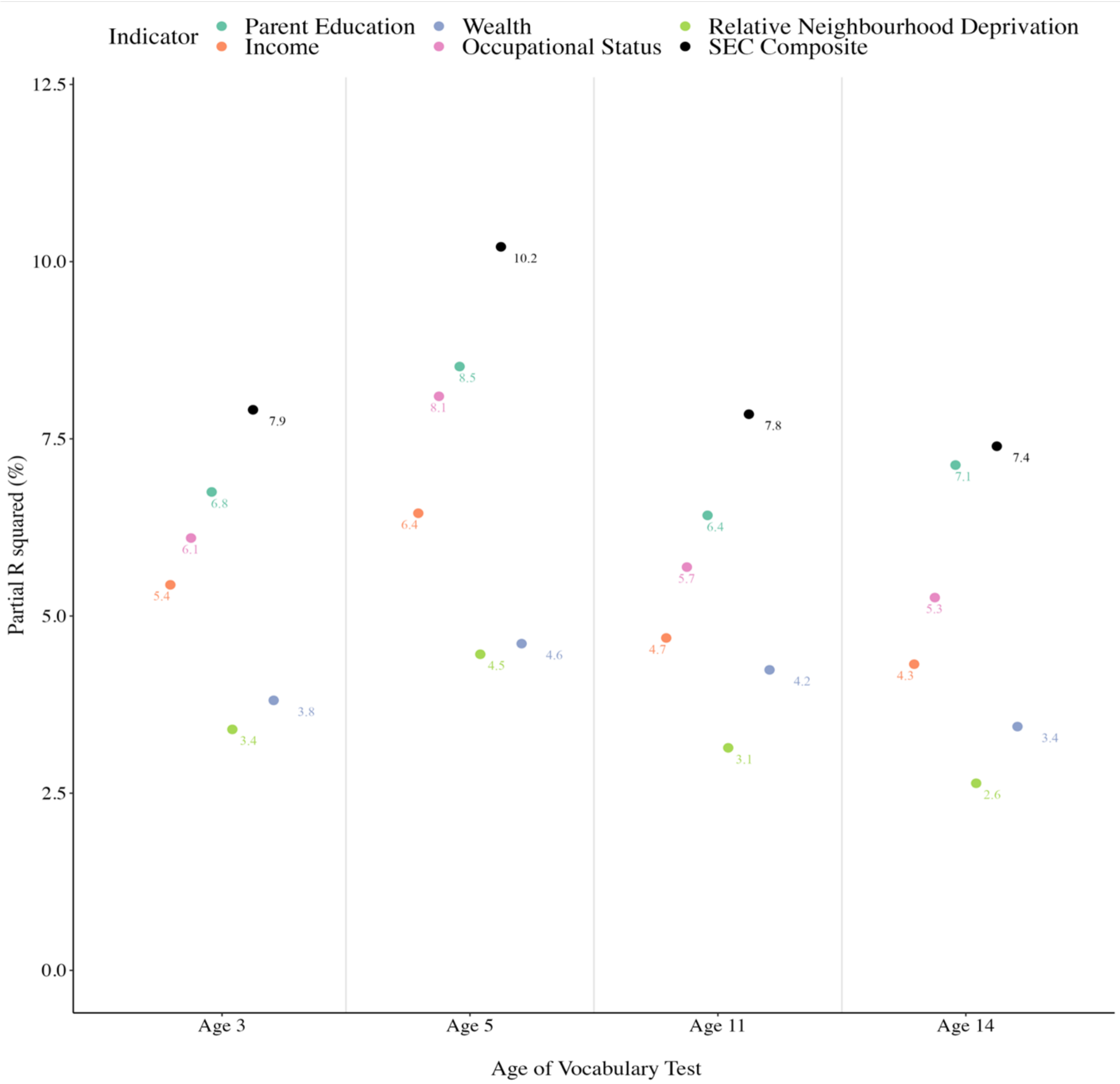


Figure 1. ***Variance explained by SEC indicators in predicting vocabulary in MCS2001 cohort.*** Partial R^2 values for separate models predicting vocabulary at ages 3, 5, 11 and 14, for 5 separate SEC indicators and a composite SEC indicator. Models adjusted for potential confounding variables of sex, ethnicity and English as an additional language (EAL).

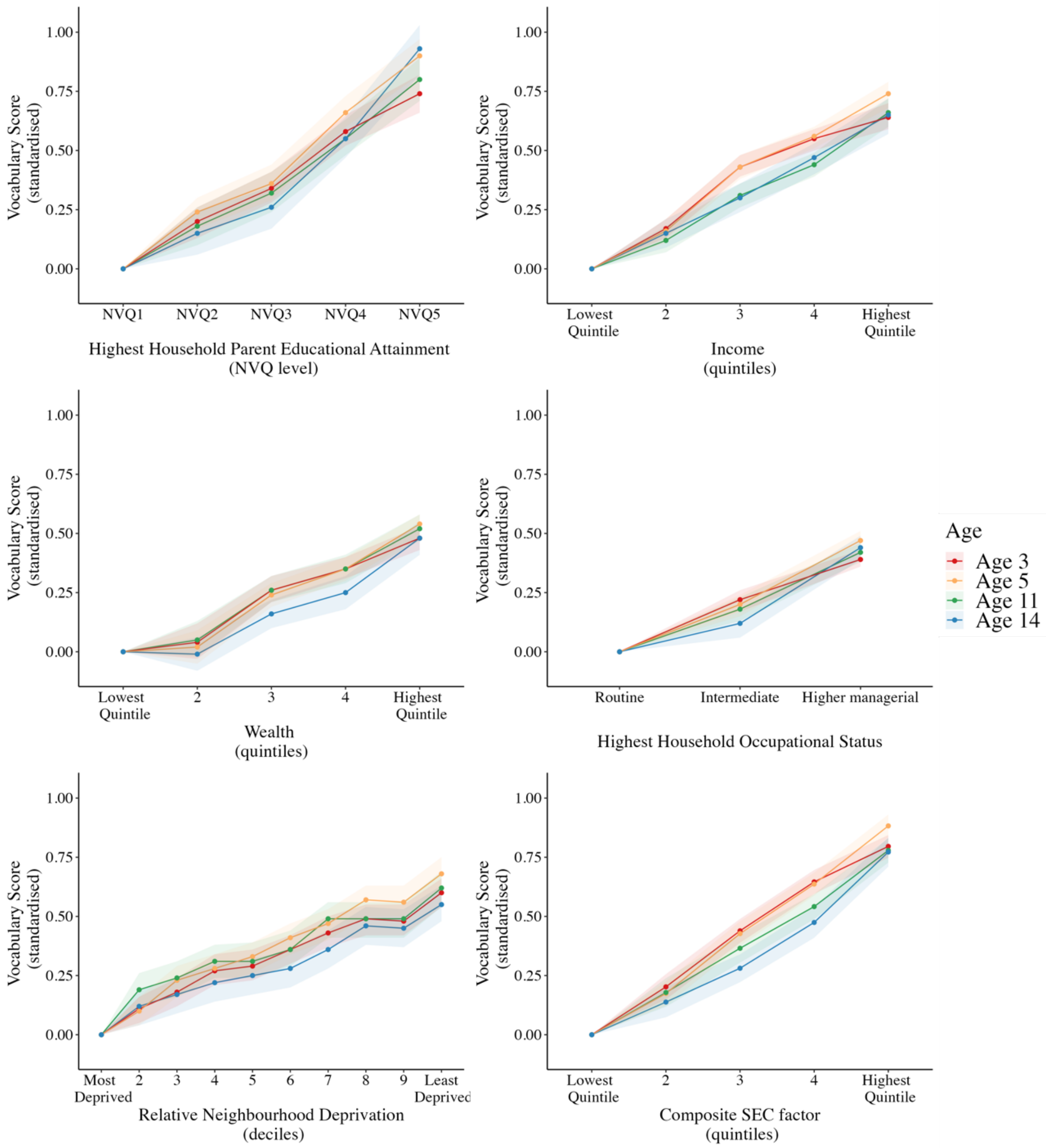
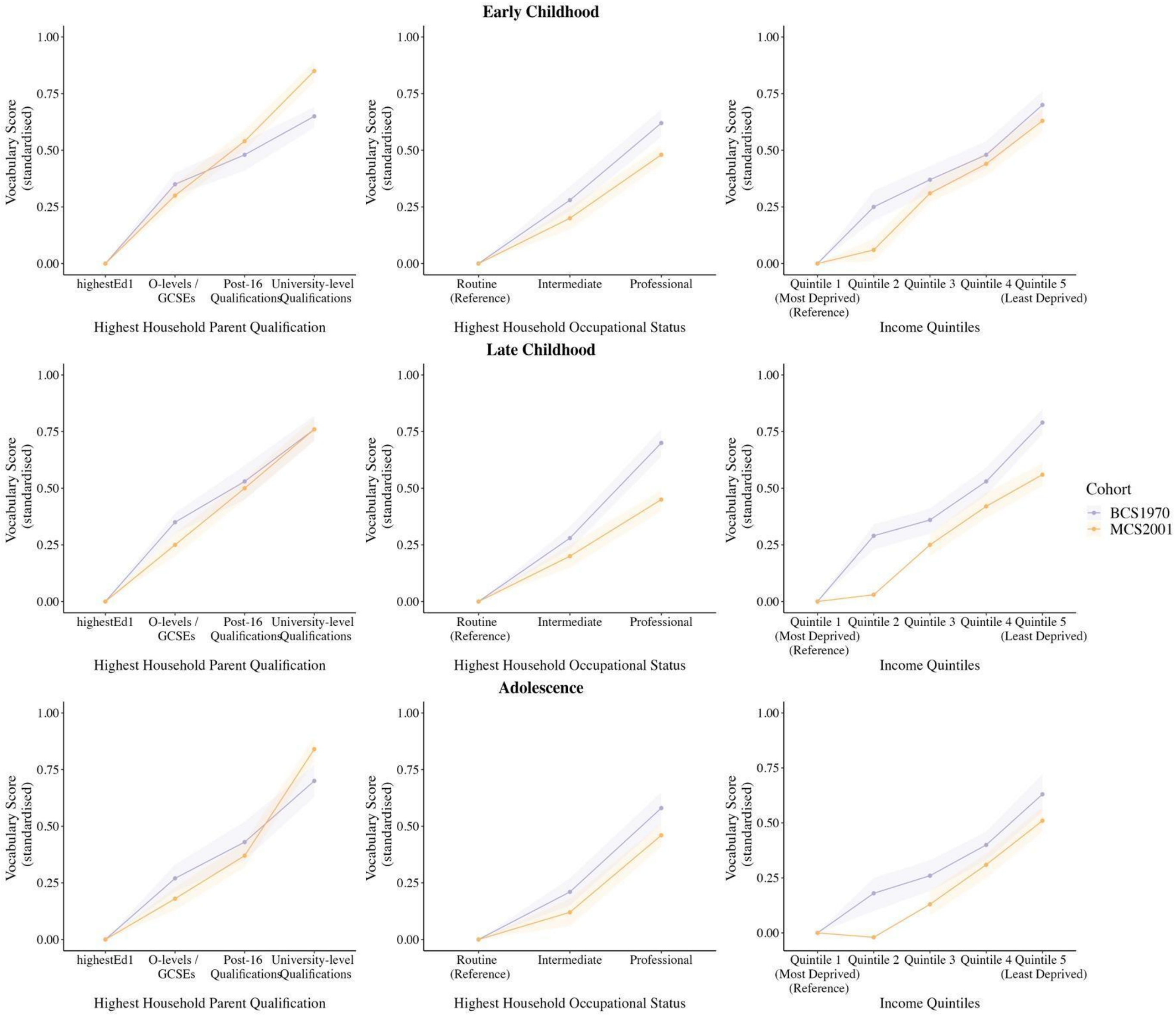


Figure 2: ***Associations between SEC indicators and vocabulary at ages 3, 5, 11 and 14 in the MCS2001 cohort.*** β coefficients and 95% confidence intervals for vocabulary at ages 3, 5, 11 and 14,

plotted as a function of each SEC indicator. Coefficients adjusted for potential confounding variables of sex, ethnicity, and English as an additional language (EAL).

Table 3: ***Descriptive Statistics in MCS2001 and BCS1970 for the cross-cohort comparison***

	Proportion (%) or Mean(±SD) [95% CIs]	
Variable	*BCS1970 (N = 14,851)*	*MCS2001 (N = 16,020)*
Demographics		
Sex (Male)	50.55	51.33
Sex (Female)	49.45	48.67
Ethnicity (White)	93.52	86.03
Ethnicity (Minority)	6.48	13.97
Language Status (English only)	94.97	88.64
Language Status (English as Additional Language)	5.03	11.36
Socioeconomic Circumstances		
Parent Education (no/low level)	54.49	21.14
Parent Education (O-levels/GCSEs grades A*-C)	20.23	32.1
Parent Education (ost-16 quals)	7.66	21.85
Parent Education (university level quals)	17.62	24.92
Income Quintile 1	21.31	19.67
Income Quintile 2	19.81	19.58
Income Quintile 3	20.84	20.44
Income Quintile 4	20.68	20.07
Income Quintile 5	17.36	20.24
Occupational Status (routine)	14.32	22.47
Occupational Status (intermediate)	50.88	18.98
Occupational Status (higher managerial)	33.63	38.76



Occupational Status (unemployed)	1.16	19.78

Descriptive statistics combined across 25 imputed datasets. Descriptive statistics are sample and attrition weighted (MCS2001 cohort) and attrition weighted (BCS1970 cohort)

Figure 3: *Associations between SEC and language ability in the MCS2001 and BCS1970 cohorts in early childhood, late childhood, and adolescence.* Vocabulary in early childhood (top), late childhood (middle) and adolescence (bottom), plotted as a function of highest household parent education (left), highest household occupational status (middle), and income (right) in two cohorts. Data are β coefficients and 95% confidence intervals. Coefficients adjusted for potential confounding variables (sex, ethnicity, English as an additional language and age at time of vocabulary test).

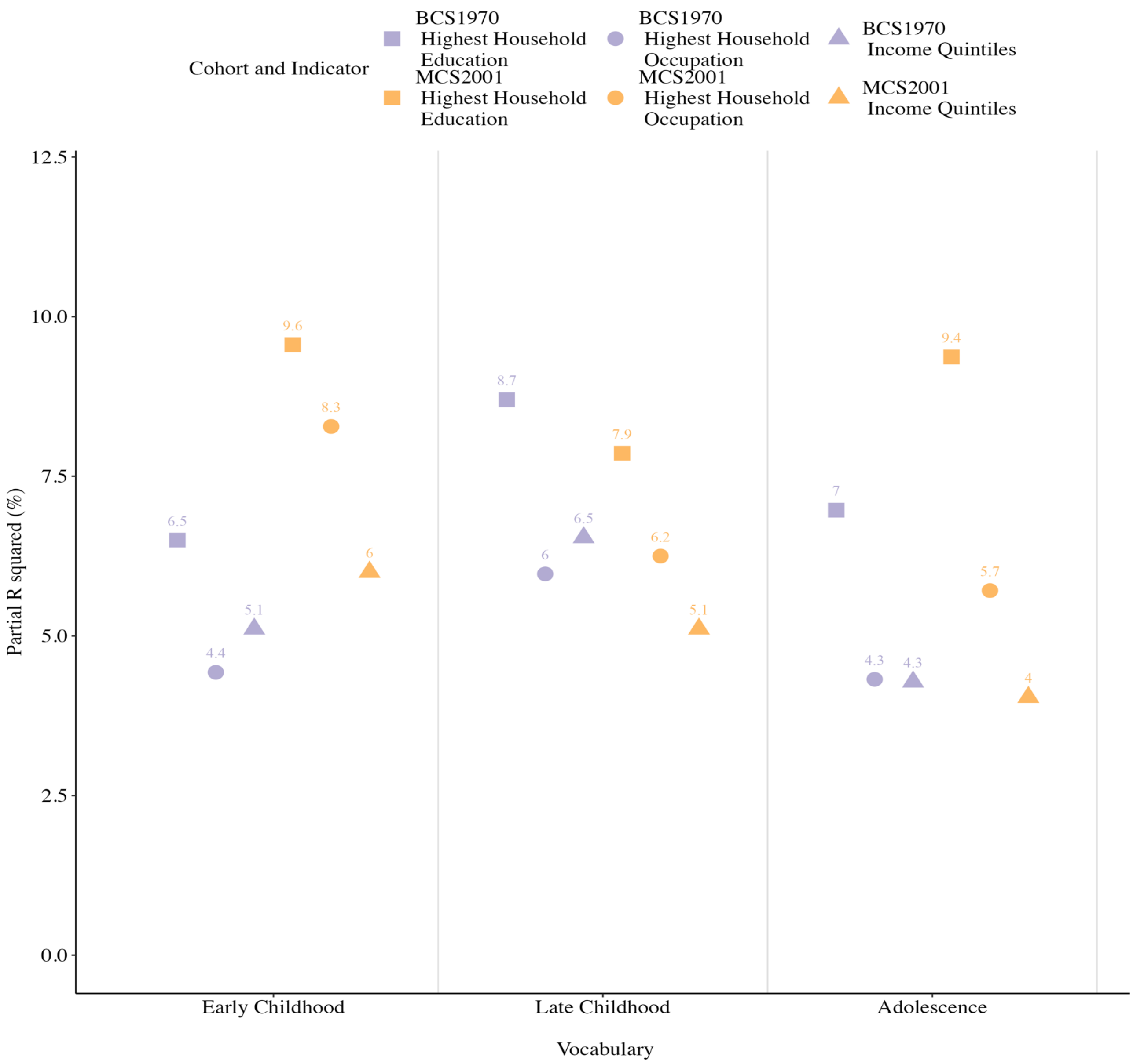


Figure 4: Variance in language explained by SEC indicators in the MCS2001 and BCS1970 cohort. Partial R^2 values (having adjusted for potential confounders of sex, ethnicity, English as additional language and age at time of vocabulary test) for highest household education and highest household

occupational status predicting vocabulary in early childhood, late childhood, and adolescence.

Discussion

Using two UK national birth cohorts, we analysed the relation between multiple SEC indicators and vocabulary across childhood and across generations, and found that (i) all SEC measures predict unique variance at most timepoints and there is generally a monotonic step up in child language for each step up on any given SEC measure. Parent education has the greatest predictive value (closely followed by income, wealth, and occupation) and neighbourhood deprivation the least; (ii) inequalities persist from ages 3 to 14 years, with SEC indicators explaining the most variance in vocabulary scores at 5 years, and an accelerated increase in vocabulary at the higher ends of the socio-economic scale at 14 years; and (iii) across three decades, observed inequalities have generally been stable, but the advantage associated with having parents with higher levels of education has increased.

Overall, the SEC predictor that explains the most variance in child vocabulary across development is caregiver education. However, income, wealth, and occupational status also uniquely predicted large amounts of variance. For all of these indicators, a step up from each level to the next was associated with a substantial step up in vocabulary. This pattern of monotonic increase occurred for all SEC indicators. Thus while most research exploring differences in child language, and in the quality and quantity in child directed speech, tends to compare higher and lower SEC groups (Fernald, Marchman, & Weisleder, 2013; Hart & Risley, 1995; Hirsh-Pasek, 2015; McGillion, Pine, Herbert, & Matthews, 2017; Rowe, 2012; Schwab & Lew-williams, 2016), our findings suggest differences exist across the range of the SEC measures, rather than just between those at the top and bottom of the distribution. Each of these SEC indicators deserve particular attention in the effort to unpick *why* SEC is related to child vocabulary so as to be able to find mechanisms for effective interventions. Caregiver education has been argued to be the most relevant SEC marker for child development (Hoff, 2013; Hoff et al., 2012) because it is associated with caregiver-child interactions and parent knowledge about development (Rowe, 2012, 2018). Parent vocabulary mediates the relation between parent education and child vocabulary ability (Sullivan, Moulton & Fitzsimons, 2021), as well as mediating the relationship between the home learning environment and vocabulary. For example, parents with strong language skills are more likely to participate in reading with their child and may also be more successful in engaging their children in such activities, compared to parents with poor language skills (Sullivan, Ketende & Joshi, 2013). The role of genetics should also be considered here, as language ability is observed to be partly heritable (Chow & Wong, 2021). Prising apart the relative influence of heredity and culture is challenging, given the interplay between the two (Scarr & Mccartney, 1983; Harden, 2021): caregivers and infants with different genetic profiles shape learning environments differently to one another. Unravelling this will require rich datasets that include information regarding interaction dynamics.

While income explained about 6% of unique variance in children's vocabulary, family wealth explained less (about 3-4%), particularly early in childhood. Income is often assumed to affect vocabulary outcomes through the provision of learning resources (Duncan et al., 2017; Washbrook & Waldfogel, 2011). Wealth is usually operationalised as total assets net of outstanding total debt (Killewald et al., 2017), and while one might assume this would act in a similar way to income, it may only become a predictor of outcomes in late adolescence-early adulthood, for example through access to quality secondary education in expensive neighbourhoods (Department for Education, 2017a; Machin, 2011), or financial assistance with higher education (Moulton, Goodman, Nasim, Ploubidis, & Gambaro, 2021; Pfeffer, 2018). Whereas in the UK, most wealth is concentrated in housing (with financial wealth only prominent at the top of the distribution), in the US, financial wealth is more common (Cowell, Karagiannaki, & McKnight, 2019; Office for National Statistics, 2019). International comparisons of the relative predictive value of different SEC indicators across many different countries, alongside qualitative studies, have the potential to shed light on the mechanisms via which these SEC indicators are likely affecting language acquisition and inequalities.

In the contemporary British cohort, inequalities in language ability widen between the ages of 3 and 5. This supports arguments for testing early interventions that seek to avoid inequalities becoming entrenched before children access formal schooling. There is also a clear advantage among 14-year-olds of having parents with a higher level of education. By this age, some adolescents may have vocabulary abilities exceeding those of their parents. Exposure to language occurs in increasingly diverse settings throughout the school years, including via interactions with peers, teachers, and written sources such as books and the internet (Sullivan et al., 2021). As children progress through school, vocabulary development (at least as measured by standardised tests) becomes more dependent on exposure to new words through reading, rather than oral language exposure (Elleman, Oslund, Griffin, & Myers, 2019). It is plausible that these sources of input are influenced by SEC. For example, the availability of and engagement with books and vocabulary-rich online content may be higher among higher SEC children (Maas, Emig & Seelmann, 2013). Children from disadvantaged backgrounds may require more support to acquire particular seams of vocabulary (Sullivan et al., 2021) and yet the type of school attended and the level of support available may differ based on SEC. For example, higher SEC children are more likely to attend private or higher quality schools than their lower SEC counterparts (Dearden, Ryan, & Sibieta, 2011), and parents of children at high performing schools are more likely to invest in educational materials and support, such as books and private tuition (Attanasio, Boneva, & Rauh, 2018). There are also SEC disparities in the amount of homework support adolescents receive at home, not only through tuition, but also in terms of additional hours spent on schoolwork (Jerrim, 2017). While universal education aims to address inequalities in educational opportunity in the UK, when it comes to vocabulary, disparities clearly persist throughout formal schooling. Further support across the lifespan and particularly in the early years and during adolescence is likely necessary to improve educational outcomes and open up employment opportunities (Deloitte, 2016).

Finally, cross-cohort comparisons suggest that inequalities in childhood language are generally similar across generations, despite decades of policy to reduce these inequalities. Nonetheless, there were some differences between the two cohorts: occupational status is becoming less valuable as a predictor, while parental university level qualifications are more clearly associated with better early child and adolescent language in contemporary society. Family income appears to be a slightly stronger predictor of early childhood language in the MCS2001 cohort, but a stronger predictor of late childhood and adolescent language in the BCS1970 cohort. It is possible that these measures are changing in the extent to which they are reliable indicators of the proximal causal factors that explain language learning (such as the caregiving / cultural environment and genetic factors). For example, the move to a more hour-glass shaped economy might mean that occupational status no longer differentiates households' social milieu as well as it once did. Likewise, while many once left the educational system even when they had the academic potential to go on, now with more opportunity to stay in education longer, this measure might better differentiate families along the lines of cognitive ability and educational aspiration. Finally, in the US, financial investments in children increased at the top of the income distribution with the rise of income inequality between 1970 and 2000 (Kornrich & Furstenberg, 2013); it is possible that corresponding increases in parental investments in children have also occurred in the UK, perhaps increasing the importance of income as a predictor of early childhood vocabulary in the MCS2001 cohort compared to the BCS1970 cohort. Alternatively, it might be that the relative importance of the various proximal causal mechanisms themselves is changing with time.

Limitations and strengths.

There are some limitations to our analyses that should be kept in mind when interpreting our results. First, although our cross-cohort comparison has provided insight into socioeconomic inequalities in vocabulary across historical time, and despite extensive efforts to harmonise our variables, historical and societal changes, particularly regarding occupational status and parent education, make it difficult to definitively compare results across the two cohorts, and such differences should be kept in mind when interpreting results. Nonetheless, when we conducted a sensitivity analysis to address this, using Ridit scores as a means of standardising SEC indicators, this revealed a similar pattern of results.

Second, it should be recognised that the vocabulary measures used at each age were necessarily different, meaning we could not assess within-child change in vocabulary scores throughout childhood. However, our focus was on the *extent* of inequalities at each age, and by using a standardised score, we were able to make comparisons that reflect population distributions in these language outcomes.

Third, while vocabulary is the most commonly used measure of language ability in research, especially with regards to inequalities, and is highly correlated with other aspects of language ability (Fenson et al., 1994; Fricke et al., 2017; Hulme, Snowling, West, Lervåg & Melby-Lervåg, 2020), a drawback of the exclusive use of standardised vocabulary tests is that they are potentially inherently biased against children experiencing social disadvantage, because the items included are more likely to occur in higher SEC settings. There has long been debate about how to separate out children's 'inherent potential' for learning language from the language ability they have in virtue of experience. Traditionally, this has been of interest to clinical researchers of speech, language, and communication pathologies, who have been interested in whether a child has a language delay due to relative lack of exposure to accessible linguistic interactions and/or due to an underlying difficulty with learning and/or processing language (e.g., Campbell et al., 1997) - since a clinician's therapeutic response may differ according to aetiology. Calls for the development and adoption of language measures that are sensitive to cultural variation in language experiences continue in the context of debates about how the observed relation between socio-economic disadvantage and language development plays out (Pace et al., 2017).

To demonstrate this limitation, we might outline three (not mutually exclusive) possible scenarios under which a child might perform poorly on a vocabulary test. First, we might consider a child who struggles to learn and process language (and who, in the absence of other known causes, may have a diagnosis of Developmental Language Disorder, with subsequent specialist speech and language therapy and educational support adapted to the specific challenges they face). Second, we might consider a child who has substantial linguistic experience and ability, but whose vocabulary has less overlap with that assessed by a standardised tests than children in the norming sample, due to cultural or socioeconomic differences. Despite having some linguistic strengths, this lack of overlap may still have a functional impact on the child, since this difference may play through in the educational system, making it harder to achieve grades that open doors to future social and economic opportunities. For example, it has long been argued that children from lower SEC backgrounds may have strengths in terms of their discourse skills, compared to middle class children, which are not captured by standardised tests (Heath, 1983; Hoff, 2013; Rogoff et al., 2017). Finally, we might consider a child who has had relatively little accessible linguistic experience, and as a consequence has lower language ability, but this difference is not associated with a skew in the types of language items being assessed on a standardised measure (as was the case for the second child). This is also likely to have a functional impact on the child, but one that cannot be as easily addressed in terms of changing the way standardised tests are normed (or indeed in terms of changes to curriculum, teaching methods, or educational assessment).

The standardised vocabulary measures employed as a proxy for general language ability cannot distinguish between these three types of children (or the more messy reality of several interacting factors contributing to differences in vocabulary assessment outcomes). However, whatever the source of a child's relative difficulty on a standardised test of vocabulary, these tests reflect skills that (rightly or wrongly) are likely important for accessing education (and are known to predict educational outcomes), thus understanding the relation between vocabulary measures and SEC remains important.

Finally, as with any longitudinal analysis, missing data had to be accounted for. Less advantaged individuals tend to be underrepresented in subsequent sweeps of cohort studies (Elliott & Shepherd, 2006; Mostafa & Wiggins, 2014). Further, a teachers strike in 1986 resulted in large amounts of missing data for the adolescent vocabulary measure in the BCS1970 (63.92%). To address this, our analyses were attrition weighted and we used multiple imputations with a rich set of auxiliary indicators to account for missing data, which is considered to be the best approach for appropriately dealing with such missingness (Little & Rubin, 2002). Despite these limitations, the strengths of this research lie in the use of large, nationally representative birth cohort studies with rich information on childhood SEC and researcher-collected, gold standard language measures throughout childhood. Although findings are generalisable to the United Kingdom and hold relatively stable across generations, they may not be generalisable beyond the UK.

Implications

The current findings have several important implications. First parent education level, income, wealth, and occupational status all explain substantial unique variance in child language. This suggests it is well worth testing the causal effects of supporting caregiver education (through lifelong learning) and/or caregiver understanding, motivation, and confidence in supporting child language development (through parenting support). Equally, it is worth testing the effect of reducing poverty – defined as low income relative to a norm (see the Baby's First Years project in the US for a move in this direction: Baby's First Years, 2018). Despite efforts to reduce poverty in the UK, it is ever-present: 22% of the UK population and 30% of children were living in relative poverty (after housing costs) in 2018-19 (Francis-Devine, 2020). Beyond political choices regarding wealth redistribution, educational attainment is claimed to be the key factor causing poor children to become poor adults (DWP, 2014). Since language is the foundation for reading ability and success in education (Public Health England, 2020), and our cross-cohort comparison revealed inequalities in vocabulary are persistently wide across time, targeting these sustained inequalities is assumed to be important in reducing the intergenerational transmission of poverty (Joseph Rowntree Foundation, 2016).

Second, since inequalities in vocabulary widen markedly between the ages of 3 and 5, it remains important to target this age group. A two-pronged approach is likely necessary, whereby family support is provided at the same time as increasing the quality of provision in early years settings (Department for Education, 2017b; Gambaro, Stewart, & Waldfogel, 2015). Regarding the first prong, we need to test ways of creating sustained support for families that leads to lasting cognitive benefits (e.g. testing the BBC's UK-wide Tiny Happy People programme; Tiny Happy People, 2021; Matthews et al., 2023). For the second prong, we need to test ways of improving the consistency and quality of pre-school education to help inequalities becoming entrenched before entry to formal schooling. Quality pre-school provision benefits language development (Becker, 2011; Schmerse, 2020) and is an important factor in supporting later educational attainment, particularly for disadvantaged SEC children (Department for Education, 2015). The introduction in the UK of the National Childcare Strategy in 1998 has made early years education a focus of policy making, particularly with respect to the availability, affordability and quality of education (Department for Education, 2017c). However, quality is inconsistent across different early years settings (Gambaro et al., 2015), such that it is now included in the Ofsted Education Inspection Framework (Ofsted, 2019).

Third, inequalities in vocabulary remain wide throughout childhood and the relative advantage of having parents with higher levels of education accelerates in adolescence as children near the point of being able to leave the education system. However, most language assessments and interventions do not go beyond the early years (Bercow, 2018). Since language skill is important for accessing many employment opportunities, not to mention taking part in wider activities and accessing services, seeking out effective ways to support adolescent language development is important (Bercow, 2018; Spencer et al., 2012).

Fourth, the fact that inequalities generally persist over historical time might be taken to support proposals that interventions to lift the language skills of more disadvantaged children need to be ambitious and scaled up considerably while remaining acceptable to those they are intended to support (Greenwood, Schnitz, Carta, Wallisch, & Irvin, 2020; List, Pernaudet, & Suskind, 2021; Wake et al., 2012). One cause for optimism on this front is that a recent large-scale evaluation has found that the Nuffield Early Language Intervention (NELI) is effective in promoting language skills of children entering formal education in England (West et al., 2021). However, another recent evaluation of a prominent UK intervention, Sure Start, suggests it benefitted child physical health (for example, reduced hospitalisations) - and did so most for those living in disadvantaged areas (Cattan, Conti, Ginja, & Farquharson, 2019) - but the benefits for cognitive outcomes are less clear (Melhuish, Belsky, & Leyland, 2010), perhaps because of a struggle to reach populations who stood to derive the maximum benefit (Law, Parkin, & Lewis, 2012). The current analyses suggest that to have a chance of making a difference, we would need to test a multi-pronged approach, implemented at a meaningful scale, for the long term and in a manner acceptable to children and their families, so as to reap sustained benefits and see the next generation of children reach their potential.

Conclusion

To sum up, the substantial individual differences we observe in child and adolescent language are explained by several SEC indicators each making their own unique contribution, most notably caregiver education, income, wealth, and occupational status. Inequalities are generally stable over developmental and historical time, and are monotonic, with each step up in SEC predicting a step up in language. The current evidence suggests a need to focus on the widening of inequalities as children enter compulsory education and as they prepare to leave it. This supports calls to test the effects of reducing poverty, increasing caregiver lifelong learning, improving early parenting support, improving quality of preschool education, and sustaining educational support throughout adolescence. Tests would need to provide evidence of both causal efficacy and acceptability to those they are intended to help. To succeed on both these fronts, the current evidence suggests we need to be ambitious.

References

Attanasio, O., Boneva, T., & Rauh, C. (2018). Parental Beliefs about Returns to Different Types of Investments in School Children. *Journal of Human Resources,* 0719-10299R1. https://doi.org/10.3368/jhr.58.2.0719-10299r1

Baby's First Years. (2018). https://www.babysfirstyears.com

Bann, D., Johnson, W., Li, L., Kuh, D., & Hardy, R. (2018). Socioeconomic inequalities in childhood and adolescent body-mass index, weight, and height from 1953 to 2015: an analysis of four longitudinal, observational, British birth cohort studies. *The Lancet Public Health*. https://doi.org/10.1016/S2468-2667(18)30045-8

Becker, B. (2011). Social disparities in children's vocabulary in early childhood. Does pre-school education help to close the gap? *British Journal of Sociology, 62*(1), 69–88. https://doi.org/10.1111/j.1468-4446.2010.01345.x

Beddington, J., Cooper, C. L., Goswami, U., Huppert, F. A., Jenkins, R., Jones, H. S., Tom, B. L., Sahakian, B. J., & Thomas, M. (2008). The mental wealth of nations. *Nature*, *455*, 1057–1060.

Bennetts, S. K., Love, J., Bennett, C., Burgemeister, F., Westrupp, E. M., Hackworth, N. J., Mensah, F. K., Levickis, P., & Nicholson, J. M. (2022). Do neighbourhoods influence how parents and children interact? Direct observations of parent-child interactions within a large Australian study. *Children and Youth Services Review*, 106704. https://doi.org/10.1016/j.childyouth.2022.106704

Bercow, J. (2018). Bercow : Ten Years On. *Royal College of Speech and Language Therapists, March*, 1–25. https://www.bercow10yearson.com/wp-content/uploads/2018/04/Bercow-Ten-Years-On-Summary-Report-.pdf

Bohlman, E. (2018). *ridittools: Useful Functions for Ridit Analysis*. https://cran.r-project.org/package=ridittools

Bolton, P. (2012). Education: historical statistics. *House of Commons Library, SN/SG/4252*(November), 20. http://www.parliament.uk/briefing-papers/SN00620.pdf

Brimer, M. A., & Dunn, L. M. (1962). *English Picture Vocabulary Test: Educational Evaluation Enterprises. English version of the Peabody Picture Vocabulary Test (PPVT; Dunn, 1959).*

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press.

Butler, N., Despotidou, S., & Shepherd, P. (1981). *1970 British Cohort Study: Ten Year Follow-Up*.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing Bias in Language Assessment. *Journal of Speech, Language, and Hearing Research, 40*(3), 519-525. https://doi.org/doi:10.1044/jslhr.4003.519

Carmines, E. G., & McIver, J. P. (1981). Analysing Models with Unobserved Variables: Analysis of Covariance Structures. In G. W. Bohrnstedt & E. F. Borgatta (Eds.), *Social Measurement: Current Issues* (pp. 65–115). Sage, Beverly Hills.

Cattan, S., Conti, G., Ginja, R., & Farquharson, C. (2019). *The health effects of Sure Start*. Institute for Fiscal Studies. http://www.ifs.org.uk

Chaplin Gray, J., Gatenby, R., & Simmonds, N. (2009). *Millennium Cohort Study Sweep 3 Technical Report.*

Chow, B. W., & Wong, S. W. L. (2021). What does genetic research tell us about the origins of language and literacy development? A reflection on Verhoef et al. (2020). *Journal of Child Psychology and Psychiatry, 62*(6), 739–741. https://doi.org/10.1111/jcpp.13399

Closs, S. J. (1986). *APU vocabulary test (multiple choice format, 1986)*. Kent: Hodder and Stoughton Educational Ltd.

Conger, R. D., & Donnellan, M. B. (2007). An Interactionist Perspective on the Socioeconomic Context of Human Development. *Annual Review of Psychology, 58*(1), 175–199. https://doi.org/10.1146/annurev.psych.58.110405.085551

Connelly, R., & Platt, L. (2014). Cohort profile: UK Millennium Cohort Study (MCS). *International Journal of Epidemiology, 43*(6), 1719–1725. https://doi.org/10.1093/ije/dyu001

Cowell, F., Karagiannaki, E., & McKnight, A. (2019). The changing distribution of wealth in the pre-crisis US and UK: The role of socio-economic factors. *Oxford Economic Papers, 71*(1), 1–24. https://doi.org/10.1093/oep/gpy047

Dearden, L., Ryan, C., & Sibieta, L. (2011). What Determines Private School Choice? A Comparison between the United Kingdom and Australia. *Australian Economic Review, 44*(3), 308–320. https://doi.org/10.1111/j.1467-8462.2011.00650.x

Deloitte. (2016). *Talent for survival: Essential skills for humans working in the machine age.*

Department for Education. (2015). *How pre-school influences children and young people's attainment and developmental outcomes over time (DFE-RB455). June,* 1–50. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/455670/RB455_Effective_pre-school_primary_and_secondary_education_project.pdf.pdf

Department for Education. (2017a). *House prices and schools: do houses close to the best-performing schools cost more?* (Issue March). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/600623/House_prices_and_schools.pdf

Department for Education. (2017b). *Statutory framework for the early years foundation stage.*

Department for Education. (2017c). *Study of Early Education: Good Practice in Early Education. January.*

Donaldson, G. W. (1998). Ridit scores for analysis and interpretation of ordinal pain data. *European Journal of Pain, 2*(3), 221–227. https://doi.org/10.1016/S1090-3801(98)90018-0

Dorling, D., Rigby, J., Wheeler, B., Ballas, D., Thomas, B., Fahmy, E., Gordon, D., & Lupton, R. (2007). *Poverty, wealth and place in Britain, 1968 to 2005.*

Doyle, O., Harmon, C. P., Heckman, J. J., & Tremblay, R. E. (2009). Investing in early human development: Timing and economic efficiency. *Economics and Human Biology, 7*(1), 1–6. https://doi.org/10.1016/j.ehb.2009.01.002

Duncan, G. J., & Magnuson, K. (2012). Socioeconomic status and cognitive functioning: Moving from correlation to causation. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(3), 377–386. https://doi.org/10.1002/wcs.1176

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2017). Moving beyond Correlations in Assessing the Consequences of Poverty. *Annual Review of Psychology, 68,* 413–434. https://doi.org/10.1146/annurev-psych-010416-044224

Dunn, L. M., Dunn, L. M., Bulheller, S., & Häcker, H. (1965). *Peabody Picture Vocabulary Test.* Circle Pines, MN: American Guidance Service.

DWP. (2014). *An evidence review of the drivers of child poverty for families in poverty now and for poor children growing up to be poor adults.*

Elleman, A. M., Oslund, E. L., Griffin, N. M., & Myers, K. E. (2019). A review of middle school vocabulary interventions: Five research-based recommendations for practice. *Language, Speech, and Hearing Services in Schools, 50*(4), 477–492. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0145

Elliott, C. D., Murray, D. J., & Pearson, L. S. (1979). *British Ability Scales.* Slough: National Foundation for Educational Research.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales Second Edition (BAS II) Early Years*. NFER-Nelson.

Elliott, J., & Shepherd, P. (2006). Cohort profile: 1970 British Birth Cohort (BCS70). *International Journal of Epidemiology, 35*(4), 836–843. https://doi.org/10.1093/ije/dyl174

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development, 59*(5).

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental science, 16*(2), 234–248. https://doi.org/10.1111/desc.12019

Fitzsimons, E., Agalioti-Sgompou, V., Calderwood, L., Gilbert, E., Haselden, L., & Johnson, J. (2017). *Millennium Cohort Study Sixth Survey 2015-2016 User Guide (First Edition). February.*

Francis-Devine, B. (2020). *Poverty in the UK: statistics - briefing paper 7096* (Issue 7096).

Fricke, S., Burgoyne, K., Bowyer-Crane, C., Kyriacou, M., Zosimidou, A., Maxwell, L., Lervåg, A., Snowling, M. J., & 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*(10), 1141–1151. https://doi.org/10.1111/jcpp.12737

Gambaro, L., Stewart, K., & Waldfogel, J. (2015). A question of quality: Do children from disadvantaged backgrounds receive lower quality early childhood education and care? *British Educational Research Journal, 41*(4), 553–574. https://doi.org/10.1002/berj.3161

Goodman, A., & Butler, N. (1986). *BCS70 - The 1970 British Cohort Study : The Sixteen-*

year Follow-up.

Greenwood, C. R., Schnitz, A. G., Carta, J. J., Wallisch, A., & Irvin, D. W. (2020). A systematic review of language intervention research with low-income families: A word gap prevention perspective. *Early Childhood Research Quarterly, 50,* 230–245. https://doi.org/10.1016/j.ecresq.2019.04.001

Griffiths, D., Lambert, P. S., & Tranmer, M. (2011). Multilevel modelling of social networks and occupational structure. *Applications of Social Network Analysis (ASNA), University of Zurich,* 4–7.

Grinstein-Weiss, M., Williams Shanks, T. R., & Beverly, S. G. (2014). Family assets and child outcomes: Evidence and directions. *Future of Children, 24*(1), 147–170. https://doi.org/10.1353/foc.2014.0002

Hansen, K. (2014). *Millennium Cohort Study: A Guide to the Datasets (Eighth Edition) First, Second, Third, Fourth and Fifth Surveys*. *February,* 1–102.

Harden, K. P. (2021). The genetic lottery: why DNA matters for social equality. Princeton University Press.

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

Heath, S. B. (1983). *Ways with words*. Cambridge University Press.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015). The Contribution of Early Communication Quality to Low-Income Children's Language Success. *Psychological Science, 26*(7), 1071–1083. https://doi.org/10.1177/0956797615581493

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*(1), 4–14. https://doi.org/10.1037/a0027238

Hoff, E., Laursen, B., & Bridges, K. (2012). Measurement and Model Building in Studying the Influence of Socioeconomic Status on Child Development. *The Cambridge Handbook of Environment in Human Development*, 590–606. https://doi.org/10.1017/cbo9781139016827.033

Holmes, C., & Mayhew, K. (2012). The Changing Shape of the UK Job Market and its Implications for the Bottom Half of Earners. In *Resolution Foundation* (Issue March).

https://www.isc.co.uk/media/3179/isc_census_2016_final.pdf

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hulme, C., Snowling, M. J., West, G., Lervåg, A., & Melby-Lervåg, M. (2020). Children's Language Skills Can Be Improved: Lessons From Psychological Science for Educational Policy. *Current Directions in Psychological Science, 29*(4), 372–377. https://doi.org/10.1177/0963721420923684

Institute of Child Health. (1975). *The 1970 birth cohort: 5 year follow up maternal self completion questionnaire.*

Jerrim, J. (2017). Private tuition and out of school study, new international evidence. In *The Sutton Trust* (Issue September).

Joseph Rowntree Foundation. (2016). *We can solve poverty in the UK.*

Killewald, A., Pfeffer, F. T., & Schachner, J. N. (2017). Wealth inequality and accumulation. *Annual Review of Sociology, 43,* 379–404. https://doi.org/10.1146/annurev-soc-060116-053331

Kornrich, S., & Furstenberg, F. (2013). Investing in Children: Changes in Parental Spending on Children, 1972–2007. *Demography, 50*(1), 1–23. https://doi.org/10.1007/s13524-012-0146-4

Law, C., Parkin, C., & Lewis, H. (2012). Policies to tackle inequalities in child health: Why haven't they worked (better)? *Archives of Disease in Childhood, 97*(4), 301–303. https://doi.org/10.1136/archdischild-2011-300827

Law, J., Charlton, J., & Asmussen, K. (2017). Language As a Child Wellbeing Indicator. In *Early intervention foundation.* Early Intervention Foundation. https://www.eif.org.uk/report/language-as-a-child-wellbeing-indicator\

Law, J., Charlton, J., Dockrell, J., Gascoingne, M., McKean, C., & Theakston, A. (2017). Early language development : Needs , provision , and intervention for preschool children from socio- economically disadvantaged backgrounds a report for the Education Endowment Foundation. In *Public Health England.* https://educationendowmentfoundation.org.uk/public/files/Law_et_al_Early_Language_Development_final.pdf

List, J. A., Pernaudet, J., & Suskind, D. L. (2021). Shifting parental beliefs about child development to foster parental investments and improve school readiness outcomes. *Nature Communications, 12*(1), 1–10. https://doi.org/10.1038/s41467-021-25964-y

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.

Maas, J. F., Ehmig, S. C., & Seelmann, C. (2013). Prepare for life: raising awareness for early literacy education; results and implications of the international conference of experts. Prepare for Life: International Conference of Experts, Leipzig, Germany. Leipzig: Stiftung Lesen.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics, 18*(6), 723–729. https://doi.org/10.1016/j.labeco.2011.05.005

Matthews, D., Bannard, C., Fricke, S., Levickis, P., Salter, G., Solaiman, K., Thornton, E., & Pine J. (2023) *Tiny Happy People Evaluation: Final Results.* BBC Education.

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*(10), 1122–1131. https://doi.org/10.1111/jcpp.12725

Mclennan, D., Noble, S., Noble, M., Plunkett, E., Wright, G., & Gutacker, N. (2019). The English Indices of Deprivation 2019 - technical report. In *Ministry of Housing, Communities and Local Government.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/833951/IoD2019_Technical_Report.pdf

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology, 27,* 415–444.

Melhuish, E., Belsky, J., & Leyland, A. (2010). The impact of Sure Start Local Programmes on five year olds and their families. In *DfE.*

Meng, X. L., & Rubin, D. B. (1992). Performing likelihood ratio tests with multiply-imputed data sets. *Biometrika, 79*(1), 103–111. https://doi.org/10.1093/biomet/79.1.103

Mostafa, T., & Wiggins, R. D. (2014). Handling attrition and non-response in the 1970 British Cohort Study. *CLS Working Paper 2014/2, June*.

Moulton, V., Goodman, A., Nasim, B., Ploubidis, G. B., & Gambaro, L. (2021). Parental Wealth and Children's Cognitive Ability, Mental, and Physical Health: Evidence From the UK Millennium Cohort Study. *Child Development*, *92*(1), 115–123. https://doi.org/10.1111/cdev.13413

Moulton, V., McElroy, E., Richards, M., Fitzsimons, E., Northstone, K., Conti, G., Ploubidis, G. B. G. B., Sullivan, A., & O'Neill, D. (2020). *A guide to the cognitive measures in five British birth cohort studies.* (Issue August). London, UK: CLOSER. https://www.closer.ac.uk/cognitive-measures-guide

Neuman, S. B., Kaefer, T., & Pinkham, A. M. (2018). Double dose of disadvantage: Language experiences for low-income children in home and school. *Journal of Educational Psychology, 110*(1), 102–118. https://doi.org/10.1037/edu0000201

O'Neill, D., Kaye, N., & Hardy, R. (2020). *Data harmonisation.* CLOSER Learning Hub, London, UK: CLOSER.

Office for National Statistics. (2019). *Total wealth in Great Britain April 2016 to March 2018* (Issue April 2016).

Ofsted. (2019). *Early years inspection handbook for Ofsted registered provion*. www.gov.uk/government/publications/ofsted-

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.* Routledge & Kegan Paul.

Oxford University Press. (2018). *Why Closing the Word Gap Matters: Oxford Language Report*.

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying Pathways Between Socioeconomic Status and Language Development. *Annual Review of Linguistics, 3*(1), 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226

Parsons, S. (2014). *Childhood cognition in the 1970 British Cohort Study* (Issue November). Centre for Longitudinal Studies.

Perkins, S. C., Finegood, E. D., & Swain, J. E. (2013). Poverty and language development: Roles of parenting and stress. *Innovations in Clinical Neuroscience, 10*(4), 10–19.

Pfeffer, F. T. (2018). Growing Wealth Gaps in Education. *Demography, 55*(3), 1033–1068. https://doi.org/10.1007/s13524-018-0666-7

Public Health England. (2020). *Best start in speech , language and communication : Guidance to support local commissioners and service leads.*

Regidor, E. (2004). Measures of health inequalities: Part 2. *Journal of Epidemiology and Community Health, 58*(11), 900–903. https://doi.org/10.1136/jech.2004.023036

Renard, F., Devleesschauwer, B., Speybroeck, N., & Deboosere, P. (2019). Monitoring health inequalities when the socio-economic composition changes: Are the slope and relative indices of inequality appropriate? Results of a simulation study. *BMC Public Health, 19*(1), 1–9. https://doi.org/10.1186/s12889-019-6980-1

Rogoff, B., Coppens, A. D., Alcalá, L., Aceves-Azuara, I., Ruvalcaba, O., López, A., & Dayton, A. (2017). Noticing Learners' Strengths Through Cultural Research. *Perspectives on Psychological Science, 12*(5), 876–888. https://doi.org/10.1177/1745691617718355

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2). https://doi.org/10.18637/jss.v048.i02

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development, 83*(5), 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*(2), 122–127. https://doi.org/10.1111/cdep.12271

Rowe, M. L., & Weisleder, A. (2020). Language Development in Context. *Annual Review of Developmental Psychology, 2*(1), 201–223. https://doi.org/10.1146/annurev-devpsych-042220-121816

Rubin, D. B. (1984). *Multiple imputation for nonresponse in surveys.* New York, NY: John Wiley & Sons.

Scarr, S., & Mccartney, K. (1983). How People Make Their Own Environments : A Theory of Genotype → Environment Effects. *Child Development, 54*(2), 424–435.

Schmerse, D. (2020). Preschool Quality Effects on Learning Behavior and Later Achievement in Germany: Moderation by Socioeconomic Status. *Child Development, 91*(6), 2237–2254. https://doi.org/10.1111/cdev.13357

Schomaker, M., & Heumann, C. (2014). Model selection and model averaging after multiple imputation. *Computational Statistics and Data Analysis, 71*, 758–770. https://doi.org/10.1016/j.csda.2013.02.017

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *Wiley interdisciplinary reviews. Cognitive science, 7*(4), 264–275. https://doi.org/10.1002/wcs.1393

Spencer, S., Clegg, J., & Stackhouse, J. (2012). Language and disadvantage: A comparison of the language abilities of adolescents from two different socioeconomic areas. *International Journal of Language and Communication Disorders, 47*(3), 274–284. https://doi.org/10.1111/j.1460-6984.2011.00104.x

Sullivan, A. (2007). Cultural capital, cultural knowledge and ability. *Sociological Research Online, 12*(6), 1–14. https://doi.org/10.5153/sro.1596

Sullivan, A., & Brown, M. (2015). Reading for pleasure and progress in vocabulary and mathematics. *British Educational Research Journal, 41*(6), 971–991. https://doi.org/10.1002/berj.3180

Sullivan, A., Ketende, S., & Joshi, H. (2013). Social Class and Inequalities in Early Cognitive Scores. *Sociology, 47*(6), 1187–1206. https://doi.org/10.1177/0038038512461861

Sullivan, A., Moulton, V., & Fitzsimons, E. (2021). The intergenerational transmission of language skill. *The British Journal of Sociology, 72*(2), 207–232. https://doi.org/10.1111/1468-4446.12780

Tiny Happy People. (2021). https://www.bbc.co.uk/tiny-happy-people

Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using Multivariate Statistics* (4th ed., pp. 653–771). Allyn & Bacon: Needham Heights, MA, USA.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wake, M., Levickis, P., Tobin, S., Zens, N., Law, J., Gold, L., Ukoumunne, O. C., Goldfeld, S., Le, H. N. D., Skeat, J., & Reilly, S. (2012). Improving outcomes of preschool language delay in the community: Protocol for the Language for Learning randomised controlled trial. *BMC Pediatrics, 12*. https://doi.org/10.1186/1471-2431-12-96

Washbrook, E., & Waldfogel, J. (2011). *On your marks: Measuring the school readiness of children in low-to-middle income families. December*. http://www.resolutionfoundation.org/app/uploads/2014/08/On-your-marks.pdf

West, G., Snowling, M. J., Lervåg, A., Buchanan-Worster, E., Duta, M., Hall, A., McLachlan, H., & Hulme, C. (2021). Early language screening and intervention can be delivered successfully at scale: evidence from a cluster randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.13415

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377–399. https://doi.org/10.1002/sim.4067

WHO. (2013). Health Inequality Monitoring. *World Health Organization*.

Data, code and materials availability statement

Data: The data used in this paper are held by the UK data service (UKDS; https://ukdataservice.ac.uk). Users of these datasets must agree to an End User License before accessing the data. The datasets can be accessed by creating an account, setting up a project, adding the relevant datasets to the project and agreeing to the End User License for each dataset to be downloaded. Because of restrictions put in place by the UKDS, we are unable to provide a direct link to the data. However, we have provided a note on the datasets used in the GitHub repository for this project (https://github.com/emmathornton/inequalities-vocabulary), which details each dataset required.
Code: All code for this paper can be found on GitHub: https://github.com/emmathornton/inequalities-vocabulary

Ethics statement

Analyses in this paper consist entirely of secondary data analysis of previously collected publicly-available data, therefore no ethical approval was sought.

Authorship and Contributorship Statement

ET, DM, PP and CB conceptualised the study and designed the analysis plan. ET and CB analysed the data, and ET produced data visualisations. ET wrote the first draft of the manuscript, and ET, DM, PP and CB revised the manuscript. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

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 Dr Vanessa Moulton (CLS, UCL) for their extremely helpful comments and advice on the inclusion of income in our cross-cohort comparison analyses, and on the conceptualisation of the wealth variable.

Appendix A

Appendix A contains the details of the vocabulary measures used.

MCS2001 cohort only analyses.

British Ability Scales II (BAS II): Naming vocabulary. Ages 3 & 5 (Elliott, Smith & McCulloch, 1996). This test consists of 36 items of coloured pictures of objects. Cohort members were asked to name each item. Progression through this test depends on performance, and poor performance may result in a different, easier set of items being administered. Cohort members were born over a 1.5 year period (September 2000-January 2002) and assessed over a range of months, so age at the time of testing may differ between cohort members. Therefore, we used t-scores (as published in the data), which are adjusted for item difficulty and age on BAS II age normed data. These were converted to z scores for analyses.

Age 3: At the age of 3, cohort members start the test at item 1. The test ended if the cohort member made five sequential errors. Item 16 was a “decision point” based on performance so far: if the cohort member had got 3 or more items wrong prior to item 16, the test was terminated. If not, the test continued to item 30, the next decision point, where the test was terminated if the cohort member had got 3 or more items wrong. If not, the test continued until item 36 (the end of the test)(Moulton, 2020).

Age 5: The assessment started from picture 12, as this is where children aged 5 start the test. Progression depended on the answers given by the cohort member and the test ended when the child made five sequential errors. However, if at the beginning of the test, the child has made five sequential errors and had less than three correct items, the assessment restarted at an earlier stage with easier items and more teaching items (Chaplin Gray, Gatenby, & Simmonds, 2009; Moulton, 2020). Therefore, MCS cohort members did not complete the same items, as progression through the test depends on their performance and poor performance may result in administration of an easier set of items.

British Ability Scales II (BAS II): Verbal similarities. Age 11. (Elliott, Smith & McCulloch, 1996). This is a measure of verbal reasoning and verbal knowledge. There were 37 items in total (although the first was a practice item and not counted in the final score). Three words were read out to the cohort member, usually by the interviewer, and cohort members had to name the category to which the three words belong (Moulton, 2020; see Figure S1 for examples). Cohort members started the test at age 16, as this is where children aged 11 start the test, and completed up to item 28 (the decision point, based on performance so far). At this point, if there are less than

3 incorrect answers, cohort members continue to item 33. If there are less than 3 correct answers, cohort members are rerouted to an earlier stage, and instead complete items 8-15. If there are five sequential errors and less than three correct items, the cohort members are rerouted to an earlier stage and again complete items 8-15. However, if these items are also too difficult, the test starts again from item 1(Hansen, 2014; Moulton, 2020).

Progression through this test depends on performance, and poor performance may result in a different, easier set of items being administered. Cohort members were born over a 1.5 year period (September 2000-January 2002) and assessed over a range of months, so age at the time of testing may differ between cohort members. Therefore, we used t-scores (as published in the data), which are adjusted for item difficulty and age on BAS II age normed data. These were converted to z scores for analyses.

Word Activity Task. Age 14 (Closs, 1986). This test is a measure of vocabulary and also assessed the understanding of meanings of words and word knowledge. Items were a subset of the items from the Applied Psychology Unit (APU) Vocabulary Test(Closs, 1986). Cohort members were given a list of 20 target words, each presented alongside 5 other words. Cohort members had to choose the word which meant the same, or nearly the same as the target word, from the 5 options. Items increased in difficulty throughout the test (Fitzsimons et al., 2017; Moulton, 2020). See Figure S2 for examples of items.

Figure S1. ***Example items from BAS II Verbal Similarities***

	(a)	*(b)*	*(c)*	*(d)* *(cohort member's answer)*
FIRST 3 ITEMS (from item 16)				
	Syrup	Toffee	Cake	_____________
	Water	Oil	Blood	_____________
	Jar	Bag	Box	_____________
LAST 3 ITEMS (items 26-28)				
	Fraud	Lie	Forgery	_____________
	Hurricane	Draught	Blizzard	_____________
	Siren	Beacon	Horn	_____________

Figure S2. ***Example Items from Word Activity Task***

	(a)	(b)	(c)	(d)	(e)
FIRST 5 WORDS					
QUICK	always	best	neat	sick	fast
TIDINGS	steps	reason	jetty	mountains	news
CONCEAL	advise	hide	gather	freeze	conciliate
UNIQUE	several	matchless	simple	ancient	absurd
DUBIOUS	tawny	obstinate	gloomy	muddy	doubtful
LAST 5 WORDS					
OBSOLETE	execrable	secret	innocuous	rigid	redundant
ERUDITE	learned	spasmodic	superfluous	pathetic	spurious
PROSAIC	commonplace	flowery	laudable	poetical	spacious
ASCETIC	artistic	dissolute	austere	antipathetic	charlatan
PUSILLANIMOUS	loud	living	timid	averse	correct

Cross Cohort Comparison

Early Childhood Language Ability

MCS2001 Age 5: BAS II Naming vocabulary (Elliott, Smith & McCulloch, 1997). Details of this test can be found above. The difference here is that in order to aid comparability to BCS1970 data, we here used the ability scores, which are just adjusted for item difficulty and account for the items that the cohort member completed. We adjusted for age in months at the time of the test, instead of using the t-scores available in the data, which are adjusted for age based on BAS II age norms.

BCS1970 Age 5: English Picture Vocabulary Test (EPVT; Brimer & Dunn, 1962). This test is a UK version of the Peabody Picture Vocabulary Test (Dunn, Dunn, Bullheller & Häcker, 1965). Cohort members were shown 56 sets of four diverse images and a specific word associated with each set of four images. They were asked to select one picture that matched the presented word and were awarded one point for every correct response. The items became increasingly difficult as the test progressed, and the test stopped when the child made five errors in a set of eight items (Parsons, 2014); the 5th wrong answer in a set of 8 sequential items was the ceiling item. Each cohort member's score was the number of correct responses reached before the ceiling item, or (for cohort members who completed the final item of the test without making 5 mistakes in 8 consecutive items), the number of correct responses at the end of the test. Some children did not have a base item, meaning they did not correctly answer 5 of the first 8 items; these children were given a score of 0. Details on the scoring of this vocabulary measure and the SPSS syntax used can be found in appendix 3 of "Childhood Cognition in the 1970 British Cohort Study" (Parsons, 2014).

Scores in the current sample ranged from 0- 56, with higher scores indicating a better language ability. The EPVT has been reported to have a reliability coefficient of .96 (Osborn, Butler, & Morris, 1984). The BCS data does not contain item level responses for the EPVT, only the raw total score, therefore we cannot report the alphas for our analysis sample. However, the items administered in this test were obtained from the British Library to ensure that the procedure and items administered were comparable to other vocabulary tests. Target words can be found in Figure S3 (which are taken from the Age 5 Test Booklet, see here:
https://cls.ucl.ac.uk/wp-content/uploads/2017/07/BCS70_age5_test_booklet.pdf). An example of the 4 pictures administered to cohort members could be a drawing of a spider, whale (target), bird and giraffe.

Figure S3. ***English Picture Vocabulary Test Items***
Late Childhood Language Ability

MCS2001 Age 11: BAS II verbal similarities (Elliott, Smith & McCulloch, 1997). Details of this test can be found above. The difference here is that in order to aid comparability to BCS1970 data, we here used the ability scores, which are just adjusted for item difficulty and account for the items that the cohort member completed. We adjusted for age in months at the time of the test, instead of using the t-scores available in the data, which are adjusted for age based on BAS II age norms.

BCS1970 Age 10: BAS word similarities (Elliot, Murray & Pearson, 1979). This test was made up of 21 items, each of which consisted of three words. The teacher read these sets of items out loud and cohort members had to a) name another word that was consistent with the three words in the item and b) state how the words were related. In order to receive a point, cohort members had to correctly answer both parts of the question (Moulton, 2020; Parsons, 2014). If they only answered one part correctly, cohort members received a score of 0 for that item. When the cohort member failed to give the correct group name and an example for four sequential items,

English Picture Vocabulary Test Score Sheet
(Survey Version)

Introductory word (Page P) — P ☐ ☐ ball

Practice words (Pages A, B & C) — A ☐ ☐ spoon; B ☐ ☐ chair; C ☐ ☐ car

Test words (Pages 1 to 56)

1	☐ ☐ drum	15	☐ ☐ goat	29	☐ ☐ barber	43	☐ ☐ sole
2	☐ ☐ time	16	☐ ☐ peeping	30	☐ ☐ wasp	44	☐ ☐ walrus
3	☐ ☐ fence	17	☐ ☐ temperature	31	☐ ☐ yawning	45	☐ ☐ weapon
4	☐ ☐ skiing	18	☐ ☐ signal	32	☐ ☐ captain	46	☐ ☐ sentry
5	☐ ☐ chicken	19	☐ ☐ river	33	☐ ☐ trunk	47	☐ ☐ wailing
6	☐ ☐ climbing	20	☐ ☐ badge	34	☐ ☐ argument	48	☐ ☐ globe
7	☐ ☐ leaf	21	☐ ☐ hook	35	☐ ☐ coin	49	☐ ☐ valve
8	☐ ☐ digging	22	☐ ☐ whale	36	☐ ☐ hive	50	☐ ☐ plumage
9	☐ ☐ teacher	23	☐ ☐ acrobat	37	☐ ☐ chemist	51	☐ ☐ assistance
10	☐ ☐ sewing	24	☐ ☐ tweezers	38	☐ ☐ funnel	52	☐ ☐ carpenter
11	☐ ☐ nest	25	☐ ☐ submarine	39	☐ ☐ insect	53	☐ ☐ destruction
12	☐ ☐ arrow	26	☐ ☐ balancing	40	☐ ☐ cutlery	54	☐ ☐ spire
13	☐ ☐ parachute	27	☐ ☐ binocular	41	☐ ☐ shears	55	☐ ☐ reel
14	☐ ☐ cobweb	28	☐ ☐ ornament	42	☐ ☐ exhausted	56	☐ ☐ coast

the test was terminated. Items became progressively harder throughout the test. Details on the scoring of this vocabulary measure and the SPSS syntax used can be found in appendix 3 of "Childhood Cognition in the 1970 British Cohort Study" (Parsons, 2014).

Adolescent language ability

MCS2001 Age 14: Word activity task (Closs, 1986). Details of this test can be found above. We adjusted for age in months at the time of the test, to account for the fact that cohort members were different ages in the MCS2001 and BCS1970 cohorts at the adolescent time point. Items from this test were a subset of the test administered to BCS1970 cohort members when they were aged 16.

BCS1970 Age 16: Vocabulary test (Closs, 1986). This test consisted of 75 items: an item consisted of a target word, presented with a multiple-choice list, from which cohort members had to select a word that meant the same as the target word (Moulton, 2020; Parsons, 2014). Items were progressively harder throughout the test (see Figure S4 for examples). Details on the scoring of this vocabulary measure and the SPSS syntax used can be found in appendix 3 of "Childhood Cognition in the 1970 British Cohort Study"(Parsons, 2014).

Figure S4. ***Example Items from the Vocabulary Test***

	(a)	*(b)*	*(c)*	*(d)*	*(e)*
FIRST 5 WORDS					
BEGIN	ask	start	plain	over	away
AID	help	contrive	assent	manage	hurry
FOREST	grass	wood	sleep	grind	judge
QUICK	always	best	neat	sick	fast
REWARD	notice	golden	prize	stable	Marine
LAST 5 WORDS					
UBIQUITOUS	omnipresent	perdition	adduce	muddy	viscous

PROSAIC	commonplace	flowery	laudable	poetical	spacious
ASCETIC	artistic	dissolute	austere	antipathetic	charlatan
APOSTATE	insufferable	monastic	exegesis	renegade	vicious
PUSILLANIMOUS	loud	living	timid	averse	correct

Figure adapted from Childhood Cognition in the 1970 British Cohort Study, page 29 (Parsons, 2014). Full list of items can be found in the age 16 guide to BCS1970 data(Goodman & Butler, 1986)

Appendix B

Appendix B contains the methods and results for the preliminary analysis of the parent education variable, to determine whether maternal education or highest household education should be used.

Rationale

Previous research often uses maternal education as an indicator of parent education. We consider household SES for all of our other indicators. We therefore conducted a preliminary analysis to determine which measure of parent education predicted the most variance in our outcomes (language at ages 3, 5, 11 and 14). We stated in our preregistration (https://osf.io/482zw/) that we would use the measure of parent education that predicted the most variance in our outcome variables in our main analyses.

Method

Measures

Language ability. At ages 3 and 5, cohort members completed the naming vocabulary subscale of the BAS II. At age 11, cohort members completed the verbal similarities subscale of the BAS II. At age 14, cohort members completed a Word Activity Task. Please refer to the main manuscript and Appendix A for details.

NVQ. When cohort members were aged 3, highest NVQ level was used (both academic and vocational qualifications derived into NVQ levels 1-5, with level 5 equating to higher qualifications). Highest household NVQ was derived from mother and fathers NVQ levels. We considered highest household, mother's and father's NVQ levels as separate predictors.

Analysis plan

Following multiple imputation (see manuscript), we conducted a series of multiple linear regressions: we predicted language at each age with 3 separate regression models, with highest household NVQ level, mother's NVQ level and father's NVQ level as predictors in separate models, in turn. We controlled for gender, ethnicity and whether English was spoken as an additional language in the home.

Results

Table S1 shows results for separate models (one with highest household NVQ level, one with mother's NVQ level and one with father's NVQ level) predicting language at ages 3, 5, 11 and 14. As can be seen from Table S1, highest household NVQ consistently predicted the most variance in language at each age. Therefore, we use a measure of highest household NVQ as an indicator of parent's education in our analyses.

Table S1. ***Partial R2 values for NVQ variables***

	Partial R^2 (%)			
	Age 3	Age 5	Age 11	Age 14
Highest household NVQ	6.81	8.53	6.45	7.16
Mother's NVQ	6.71	8.38	5.83	6.84
Father's NVQ	4.8	6.22	5.22	5.9

Appendix C

Appendix C contains plots showing the extent of missing data in each of our analyses.

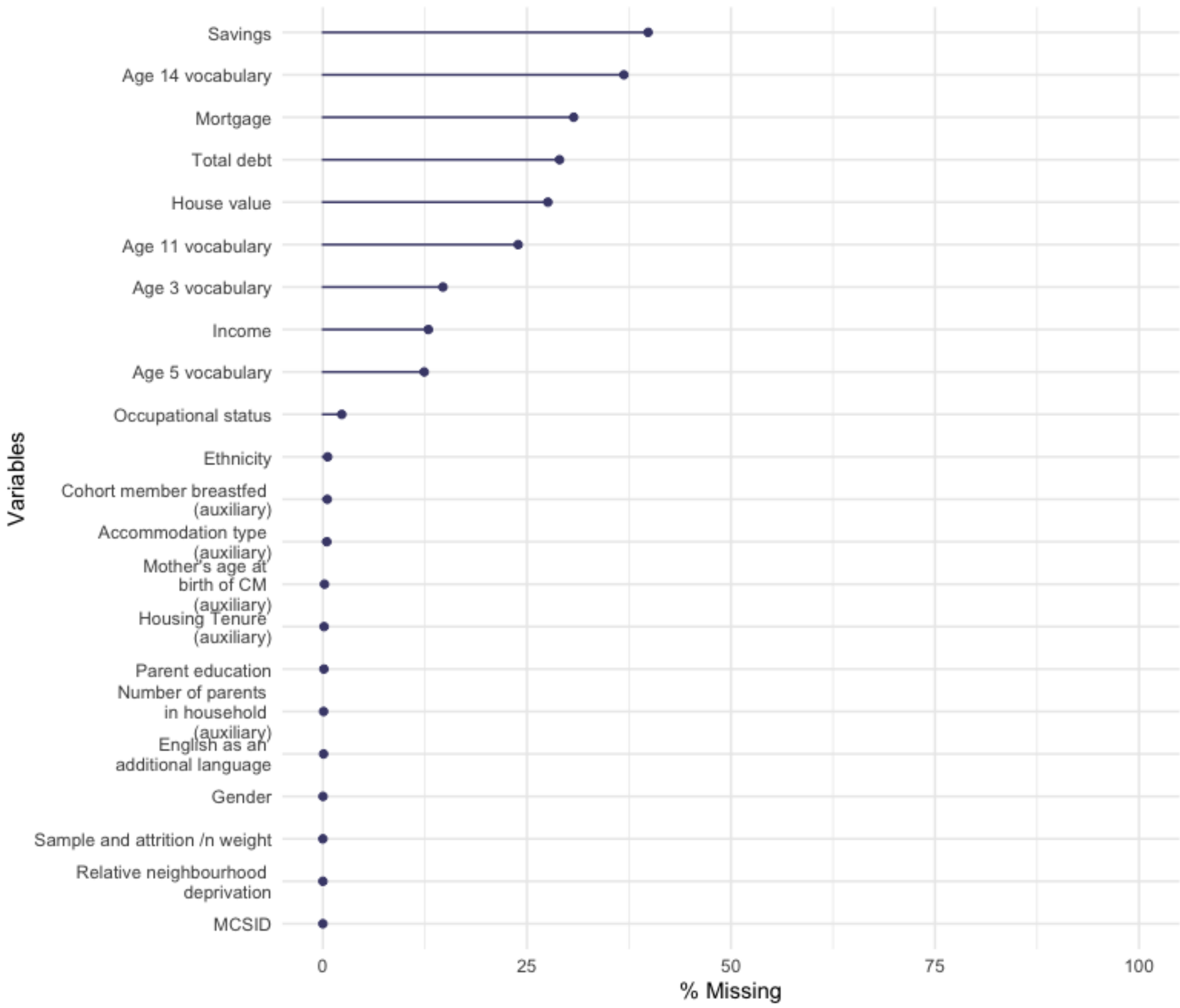


Figure S5. ***Proportion of missing data in the analytical sample used in RQ1-3 (MCS2001, N = 17,070)***

Figure S6. Proportion of missing data in the analytical sample used in the cross cohort comparison (MCS2001, N = 16,020)

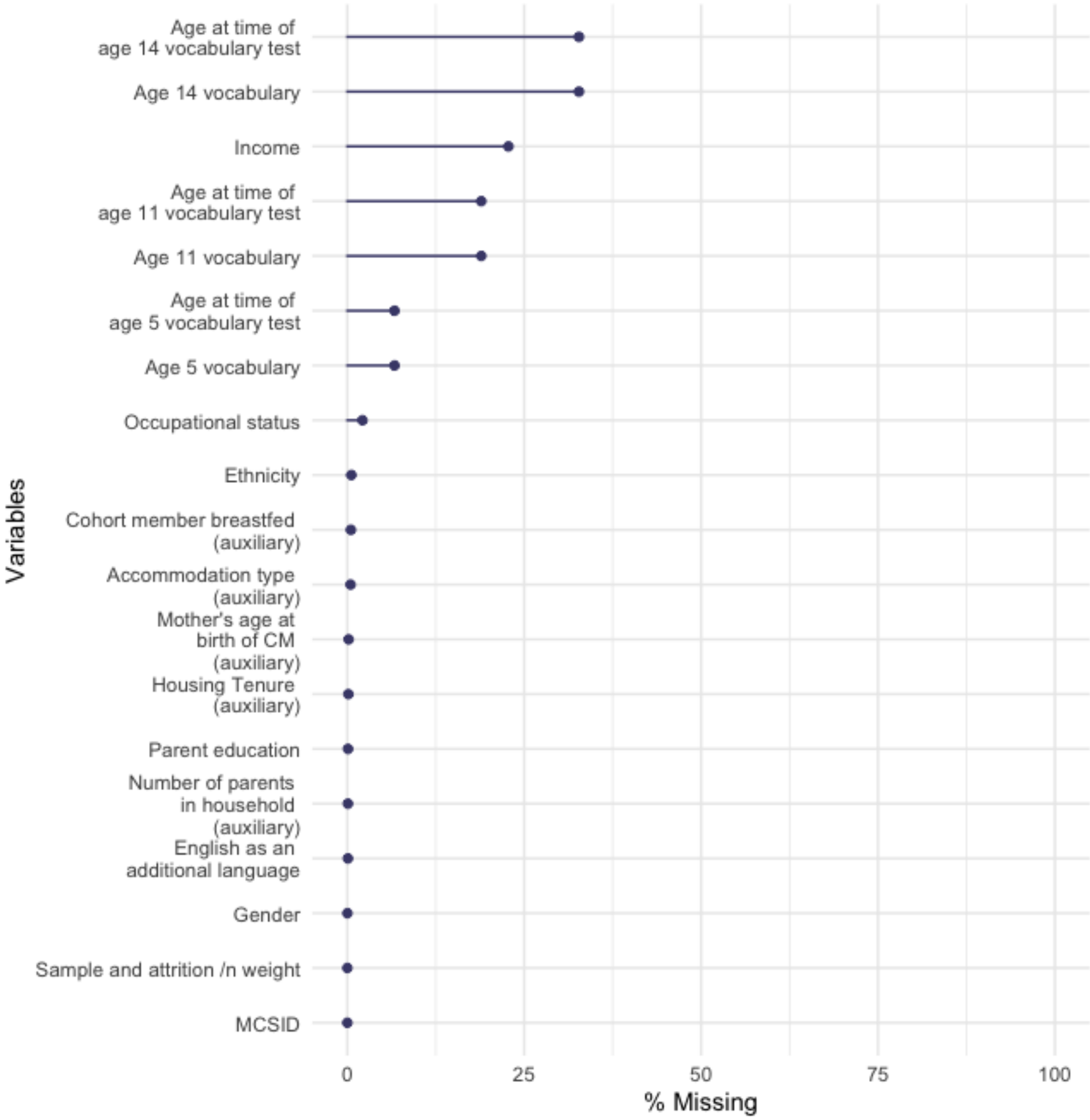
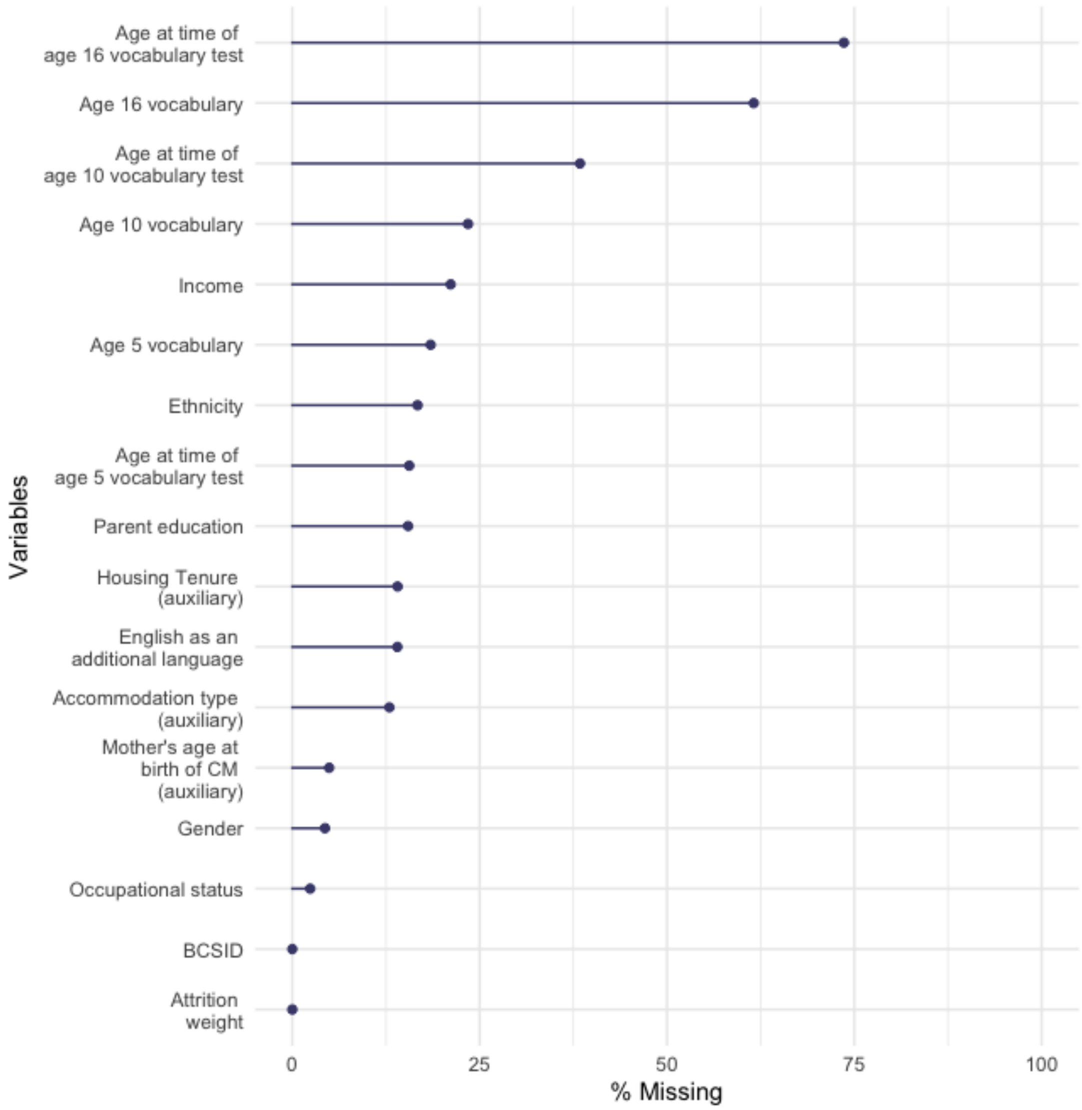


Figure S7. Proportion of missing data in the analytical sample used in the cross cohort comparison (BCS1970, N = 14,851)



Appendix D

Appendix D contains the details for the creation of the attrition weight in the BCS1970.

Procedure

1. Generate a response variable, whereby 1=response and 0=missing
2. Compile predictor variables (detailed below). Where data was missing for these, single imputation was used (random imputation, where impute random values sampled from the non-missing values of the variable)
3. Logistic regression, where response variable is the outcome, and predictor variables are variables deemed to predict missingness (detailed below)
4. Obtain predicted probabilities from the logistic regression
5. The weight variable is the inverse of these probabilities (ie predicted value/1
6. Apply a constant to the weight (weight/1.38)

A weight was created for those who were missing at age 5, those who were missing at age 10 and those who were missing at age 16. This is because although some people may have been missing at age 5, they could have returned by age 10, or they may have participated at age 5, but not age 10. These three weights were then combined into one weight variable, where the weight for age 5 response was used, if this was missing, the age 10 weight was used and if both of these were missing, the age 16 weight was used.

The mean of the final weight variable was 0.9, with a standard deviation of 0.16. The range was 0.83 to 3.82.

Predictor variables

The decision on which variables to include as predictors of response were made following the guides to the BCS datasets (Butler, Despotidou, & Shepherd, 1981; Goodman, 1986; Institute of Child Health, 1975).

Variables predicting response at the age 5 sweep:
From the birth data:

- Whether the cohort member was born to a teenage mother
- Whether the mother had high parity (defined as ≥5 pregnancies of ≥ 20 weeks of gestation)
- Whether the mother was a heavy smoker (defined as ≥15 a day)
- Marital status of mother at birth of cohort member (0=married, 1=single)
- Gender
- Father’s social class

- Mother's social class

Variables predicting response at the age 10 sweep:
From the birth data:

- Gender
- Parents born outside of Britain
- Age mother and father left full time education
- Whether the cohort member was born to a teenage mother
- Whether the mother was a single mother at birth
- Father unemployed
- Whether the cohort member was a twin
- Mother aged 40+ at child's birth

From the age 5 data:

- Child's ethnic group
- Parents with no qualifications
- Separation of mother and cohort member as a baby for 1 month or more
- Father's social class
- Low birthweight (<5lb)
- Family moved 3 or more times since 1970
- Crowded accommodation (>1 person per room = crowded)
- Whether living in private rented accommodation
- Social rating of the neighbourhood (1=poor, 0=not poor)

Variables predicting response at the age 16 sweep:

- Gender
- Father's social class
- Region

Appendix E

Appendix E contains the comparisons of the analytical sample with the full cohort samples.

Table S2. ***Full cohort sample vs analytical sample: RQ 1-3, ~2001 born cohort sample only***

	Proportion (%) or Mean(±SD) [95% CIs]	
Variable	*Whole Cohort (N= 19243)*	*Analytical Sample (N=17,070)*
Vocabulary		
Age 3 (Naming Vocabulary Score)	49.9(±11.13) [49.72;50.08]	49.33(±11.38) [49.16;49.5]
Age 5 (Naming Vocabulary Score)	54.67(±10.97) [54.5;54.85]	54.38(±11.05) [54.21;54.54]
Age 11 (Word Similarities Score)	58.8(±9.76) [58.64;58.97]	58.55(±9.88) [58.4;58.7]
Age 14 (Word Activity Task Score)	7.15(±2.63) [7.1;7.2]	7.01(±2.61) [6.97;7.05]
Demographics		
Sex (Male)	50.95	50.95
Sex (Female)	49.05	49.05
Ethnicity (White)	85.98	85.97
Ethnicity (mixed)	3.33	3.33
Ethnicity (Indian)	1.91	1.91
Ethnicity (Pakistani & Bangladeshi)	4.47	4.48
Ethnicity (Black/ Black British)	3.05	3.05

Ethnicity (other incl. Chinese)	1.27	1.26
EAL (English only)	88.5	88.49
EAL (English and another language)	9.01	9.02
EAL (only another language)	2.49	2.49
Socioeconomic Circumstances		
Parent Education (NVQ1)	5.75	5.75
Parent Education (NVQ2)	25.3	10.23
Parent Education (NVQ3)	15.97	25.3
Parent Education (NVQ4)	35.38	15.97
Parent Education (NVQ5)	7.37	35.37
Parent Education (None of these/overseas qualifications)	10.23	7.37
Income Quintile 1	20	21.28
Income Quintile 2	24.46	25
Income Quintile 3	21.53	21.13
Income Quintile 4	20.79	19.97
Income Quintile 5	13.22	12.62
Occupational Status (routine)	22.16	26.73
Occupational Status (intermediate)	18.99	12.26
Occupational Status (higher managerial)	39.04	17.91

Occupational Status (unemployed)	19.81	20.02
Wealth Quintile 1		23.08
Wealth Quintile 2		22.43
Wealth Quintile 3		19.08
Wealth Quintile 4		38.69
Wealth Quintile 5		19.8
Relative Neighbourhood Deprivation (most deprived decile)	12.96	12.96
Relative Neighbourhood Deprivation (10 - <20%)	10.84	10.84
Relative Neighbourhood Deprivation (20 - <30%)	10.32	10.32
Relative Neighbourhood Deprivation (30 - <40%)	9.11	9.11
Relative Neighbourhood Deprivation (40 - <50%)	9.73	9.73
Relative Neighbourhood Deprivation (50 - <60%)	9.73	9.73
Relative Neighbourhood Deprivation (60 - <70%)	8.77	8.77
Relative Neighbourhood Deprivation (70 - <80%)	9.02	9.02
Relative Neighbourhood Deprivation (80 - <90%)	9.55	9.55
Relative Neighbourhood	9.96	9.96

Deprivation (least deprived decile)	

Note: wealth variable compiled after imputation of house value, mortgage, savings and debts and then split to quintiles, therefore cannot calculate proportions before imputation for full sample. Means (±SD) and proportions for analytical sample are pooled across 25 imputed datasets. All descriptives are sample and attrition weighted.

Table S3. ***Full sample vs analytical sample comparisons for RQ4: cross-cohort comparison***

	BCS1970 cohort		MCS2001 cohort	
	Full Cohort Sample (N=17,196)	Analytical Sample (14,851)	Full Cohort Sample (N=19,243)	Analytical Sample (N=16,020)
Language				
Early childhood	35.3(±10.81) [35.11;35.49]	34.74(±11.19) [34.56;34.92]	108.42(±15.89) [108.17;108.68]	107.98(±16.09) [107.73;108.23]
Late childhood	12.06(±2.61) [12.01;12.11]	12.03(±2.64) [11.99;12.07]	120.64(±16.52) [120.36;120.93]	120.18(±16.83) [119.92;120.44]
Adolescence	42.49(±12.65) [42.16;42.82]	41.51(±13.23) [41.3;41.72]	7.13(±2.63) [7.08;7.18]	7(±2.6) [6.96;7.04]
Potential confounders				
Sex (male)	51.42	50.55	51.33	51.33
Sex (female)	48.58	49.45	48.67	48.67
Ethnicity (white)	95.83	93.52	86.03	86.03
Ethnicity (minority)	4.17	6.48	13.97	13.97
English as an additional language (no)	96.86	94.97	88.64	88.64

English as an additional language (yes)	3.14	5.03	11.36	11.36
SES predictors				
Parent education (no/low level)	54.13	54.49	21.13	21.14
Parent education (O-levels/GCSEs grades A*-C)	21.08	20.23	32.1	32.1
Parent education(post-16 quals)	7.76	7.66	21.85	21.85
Parent education (university level quals)	17.03	17.62	24.93	24.92
Income Quintile 1	20.94	21.31	18.09	19.67
Income Quintile 2	19.56	19.81	18.65	19.58
Income Quintile 3	21.28	20.84	20.32	20.44
Income Quintile 4	21.11	20.68	21.04	20.07
Income Quintile 5	17.11	17.36	21.9	20.24
Occupational status (routine)	0.55	14.32	19.78	22.47
Occupational status (intermediate)	15.21	1.16	22.27	19.78
Occupational status (higher managerial)	53.71	50.88	18.94	18.98

Occupational status (unemployed)	30.52	33.63	39.01	38.76

Means (±SD) and proportions for analytical sample are pooled across 25 imputed datasets. All descriptives are sample and attrition weighted (MCS2001 cohort only).

Appendix F

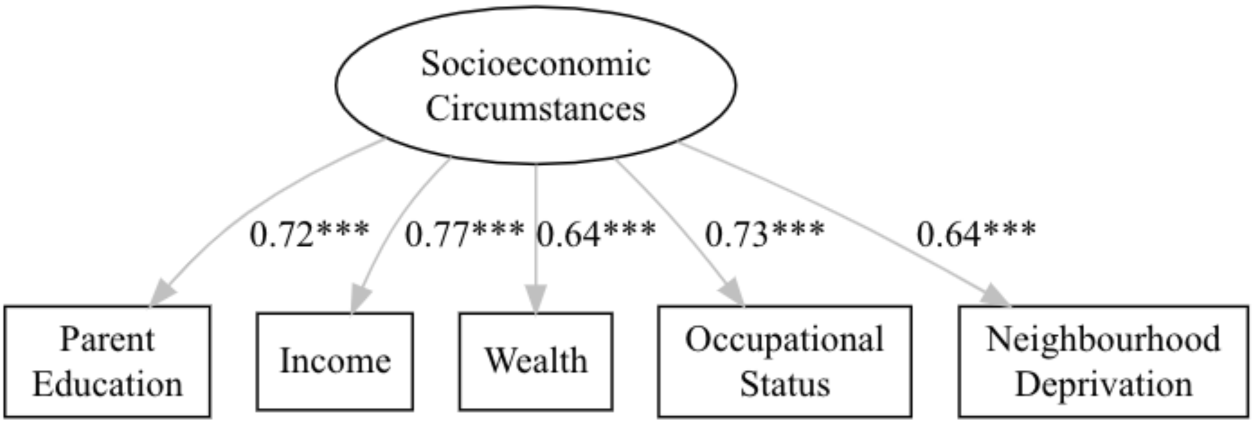
Appendix F contains the details for the confirmatory factor analysis of socioeconomic variables.

Using the lavaan package in R (Rosseel, 2012), a CFA was conducted to create a latent variable of SES. A robust weighted least squares estimator (WLSMV in the lavaan package) was used. This was due to the fact that maximum likelihood estimators are not currently supported for ordered data in the package. A latent variable factor score was then created for each individual imputed dataset, and regression models, where the factor score was the main predictor, were ran for each imputed dataset. The results of the regression models were then pooled. This procedure was conducted on separate regression models, where vocabulary at ages 3, 5, 11 and 14 were the outcome variables.

The latent variable was made up of highest household education, income, wealth, occupational status and relative neighbourhood deprivation. These variables were added to the CFA model in this order. Factor loadings can be found in Figure S1. Model fit was examined with the normed χ^2 (χ^2/df) statistic (Ullman, 2001), Comparative Fit Index (CFI) (Hu & Bentler, 1999), Root Mean Square Error of Approximation (RMSEA)(MacCallum, Browne, & Sugawara, 1996), Standardized Root Mean Square Residual (SRMR)(Hu, 1999) and Tucker Lewis Index (TLI)(Hu, 1999). Normed χ^2 statistics between 1 and 2 suggest a good model fit, and between 2 and 3 suggest an acceptable model fit (Carmines & McIver, 1981). CFI and TLI values of >.9 indicate an acceptable fit and >.95 indicate a good model fit (Hu, 1999). RMSEA values of 0.01 indicate an excellent model fit, 0.05 indicates a good fit and 0.08 indicates an acceptable model fit (MacCallum, 1996). Finally, SRMR values <.08 are indicative of a good fit (Hu, 1999). Robust fit indices are reported.

The model converged on 25 imputed datasets. Estimates were pooled across the 25 imputed datasets, using Rubin's rules (Rubin, 1984). The normed χ^2 statistic indicated a poor model fit (normed χ^2 (χ^2/5)) = 20.39. The remaining fit indices indicated the model was a good fit to the data (RMSEA = 0.034; SRMR = 0.023; CFI = 0.996; TLI= 0.993). Standardised factor loadings indicate that all variables loaded onto the latent construct (see Figure S8).

Figure S8. ***Factor Loadings for CFA***



Appendix G

Appendix G contains the model comparisons for the main analysis.

Model comparisons were conducted to determine whether each SES predictor contributed unique variance in language ability at each age; a model with all indicators included simultaneously was compared to a model with each removed in turn). If the five-predictor model was a better fit to the data than the four-predictor model following the removal of an SES indicator, then the SES variable that was dropped can be said to account for significant variance in language ability at that age.

Age 3. Parent education (Dm(5, 4519.02)= 47.08, p<.001), income (Dm(4, 2541.26)= 14.62, p<.001), wealth (Dm(4, 415.26) = 5.16, p <.001), occupational status (Dm(3, 1421.67)= 17.07, p<.001) and relative neighbourhood deprivation (Dm(9, 8022.27)= 2.42, p=.009) all accounted for significant variance in language ability at age 3.

Age 5. Parent education (Dm(5, 3051.86)= 51.42, p<.001), income (Dm(4, 1458.42)= 10.01, p<.001), wealth (Dm(4, 481.19) = 4.39, p = .002), occupational status (Dm(3, 2602.84)= 35.08, p<.001) and relative neighbourhood deprivation (Dm(9, 7731.82)= 3.63, p<.001) all accounted for significant variance in language ability at age 5.

Age 11. Parent education (Dm(5, 1308.32)= 30.99, p<.001), income (Dm(4, 861.01)= 7.33, p<.001), wealth (Dm(4, 352.28) = 8.57, p <.001), occupational status (Dm(3, 473.5)= 11.99, p<.001) and relative neighbourhood deprivation (Dm(9, 2628.53)= 2.97, p = .002) all accounted for significant variance in language ability at age 11.

Age 14. Parent education Dm(5, 690.38)= 41.28, p<.001), income (Dm(4, 494.82)= 4.05, p = .003), wealth (Dm(4, 316.61)= 4.08, p=.003) , occupational status (Dm(3, 382.10)= 9.02, p<.001) all accounted for significant variance in language ability at age 14. Relative neighbourhood deprivation did not account for significant variance (Dm(9, 1702.14)= .83, p=.589).

Appendix H

Appendix H contains the AIC values for the main analysis.

Method

AIC values were used to determine whether a model that condenses multiple SES indicators into a single composite factor is a better fit to the data than a model that includes all of these predictors simultaneously. This was to assess how a composite measure of overall socioeconomic position performs relative to individual measures and all indicators included simultaneously. The model with the lowest AIC value is the “best model” and the ΔAIC is the difference between the AIC of each of the remaining models and the AIC of the best model. The ΔAIC values are used to infer the level of support for each remaining model (Fabozzi, Focardi, Rachev & Arshanapalli, 2014). The rules of thumb for interpreting the ΔAIC values are: <2 indicates that the candidate model is almost as good as the best model; values 4-7 indicate considerably less support for the candidate model and >10 indicates that there is no support for this model being the best fit to the data (Fabozzi et al, 2014; Burnham & Anderson, 2002). AIC values are needed here as the models are not nested, therefore the drop one analyses previously used are not applicable. There are also differing numbers of predictors between the composite model and a model containing all predictors simultaneously; AIC values take account of model complexity.

Results

Regardless of age, a model that included each SES indicator as separate predictors was the “best model” (indicated by the smallest AIC values) and the ΔAIC values for the composite model at all ages were greater than 10, lending no support for the composite factor being as good a fit to the data as the ‘all predictors separately’ model (see Table S4). Thus, it is better to include SES indicators separately when predicting language ability, even when the greater model complexity is taken account of, and there may be a reduction in the predictive accuracy of the model if we reduce the indicators to a composite measure. Compared to individual measures, however, the composite factor was a better fit to the data at all ages (see Table S5). Therefore, compared to individual indicators of SES a composite measure is better than any one measure, but including all as separate indicators provides the best fit to the data.

Table S4. ***AIC and ΔAIC values Individual SES predictors compared to composite factor***

	Mean AIC [95% CIs]							
Indicator	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC
Parent Education	47295.78[47256.69;47334.88]	261.77	47721.07[47686.83;47755.32]	373.56	50165.43[50112.69;50218.17]	273.82	51195.47[51120.75;51270.19]	56.95
Income	47576.52[47538.05;47614.99]	542.51	48155.47[48119.85;48191.09]	807.96	50479.66[50423.63;50535.68]	588.05	51705.46[51629.54;51781.37]	566.94
Wealth	47921.55[47885.86;47957.25]	887.54	48533.14[48494.77;48571.52]	1185.63	50560.94[50502.53;50619.35]	669.33	51860.72[51783.21;51938.23]	722.2
Occupational Status	47432.03[47391.75;47472.31]	398.02	47805.53[47771.99;47839.08]	458.02	50294.69[50239.6;50349.78]	403.08	51533.4[51453.66;51613.14]	394.88
Neighbourhood Deprivation	48017.16[47976.83;48057.5]	983.15	48574.37[48539.55;48609.18]	1226.86	50768.7[50713.05;50824.35]	877.09	52014.12[51933.42;52094.82]	875.6
Composite	47034.01[46992.18;47075.84]	AIC*	47347.51[47310.99;47384.04]	AIC*	49891.61[49833.8;49949.43]	AIC*	51138.52[51062.96;51214.07]	AIC*

AIC* = best model; Values are the mean AIC values across 25 imputed datasets; All models adjusted for gender, ethnicity and EAL.

Table S5. ***AIC and ΔAIC values for a model containing all predictors simultaneously vs a composite factor.***

	Mean AIC [95% CIs]							
	Age 3 Language (AIC)	ΔAIC	Age 5 Language (AIC)	ΔAIC	Age 11 Language (AIC)	ΔAIC	Age 14 Language (AIC)	ΔAIC
Composite Factor	47034.01[46992.18;47075.84]	189.46	47347.51[47310.99;47384.04]	179.9	49891.61[49833.8;49949.43]	108.93	51138.52[51062.96;51214.07]	166.25
All predictors (simultaneous)	46844.55[46804.36;46884.75]	AIC*	47167.61[47132.05;47203.17]	AIC*	49782.68[49726.64;49838.71]	AIC*	50972.27[50896.97;51047.57]	AIC*

AIC* = best model
Values are the mean AIC values across 25 imputed datasets
All models adjusted for gender, ethnicity and EAL

Appendix I

Appendix I contains the coefficients for the associations between SEC indicators and vocabulary in the MCS2001 cohort.

Table S6: ***Associations between SEC indicators and vocabulary at ages 3, 5, 11 and 14 in the MCS2001 cohort.***

		β [95% CIs] *p value*			
	Indicator	Age 3	Age 5	Age 11	Age 14
Parent Education	NVQ2	.20[.13;.26] * * * p<.001	.24[.17;.30] * * * p<.001	.18[.10;.26] * * * p<.001	.15[.06;.24] * * * p<.001
	NVQ3	.34[.28;.41] * * * p<.001	.36[.29;.44] * * * p<.001	.32[.24;.41] * * * p<.001	.26[.17;.35] * * * p<.001
	NVQ4	.58[.52;.65] * * * p<.001	.66[.60;.73] * * * p<.001	.55[.48;.63] * * * p<.001	.55[.47;.64] * * * p<.001
	NVQ5	.74[.66;.82] * * * p<.001	.90[.82;.97] * * * p<.001	.80[.71;.89] * * * p<.001	.93[.82;1.03] * * * p<.001
	None of these/overseas qualifications	-.11[-.19;-.04] * * * p<.001	-.09[-.17;-.01] * p= .020	-.06[-.15;.03] p= .190	-.04[-.15;.06] p= .410
Income	Income Quintile 1	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Income Quintile	.17[.13;.21] * * * p<.001	.16[.11;.21] * * * p<.001	.12[.07;.18] * * * p<.001	.15[.09;.21] * * * p<.001
	Income Quintile 3	.43[.39;.48] * * * p<.001	.43[.38;.48] * * * p<.001	.31[.26;.36] * * * p<.001	.30[.24;.36] * * * p<.001
	Income Quintile 4	.55[.50;.59] * * * p<.001	.56[.51;.60] * * * p<.001	.44[.39;.49] * * * p<.001	.47[.40;.53] * * * p<.001

	Income Quintile 5	.64[.59;.70] * * * p<.001	.74[.68;.79] * * * p<.001	.66[.60;.72] * * * p<.001	.65[.57;.72] * * * p<.001
Wealth	Wealth Quintile 1	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Wealth Quintile 2	.04[-.03;.12] p= .220	.02[-.05;.09] p= .580	.05[-.02;.13] p= .170	-.01[-.08;.06] p= .730
	Wealth Quintile 3	.26[.21;.32] * * * p<.001	.24[.19;.29] * * * p<.001	.26[.21;.32] * * * p<.001	.16[.10;.22] * * * p<.001
	Wealth Quintile 4	.35[.31;.40] * * * p<.001	.35[.30;.40] * * * p<.001	.35[.29;.41] * * * p<.001	.25[.18;.32] * * * p<.001
	Wealth Quintile 5	.48[.43;.53] * * * p<.001	.54[.49;.58] * * * p<.001	.52[.47;.58] * * * p<.001	.48[.41;.55] * * * p<.001
Occupational Status	Occupational Status (routine)	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Occupational Status (unemployed)	-.24[-.29;-.19] * * * p<.001	-.26[-.30;-.21] * * * p<.001	-.18[-.24;-.12] * * * p<.001	-.14[-.20;-.08] * * * p<.001
	Occupational Status (intermediate)	.22[.17;.26] * * * p<.001	.20[.15;.24] * * * p<.001	.18[.13;.23] * * * p<.001	.12[.06;.17] * * * p<.001
	Occupational Status (higher managerial)	.39[.36;.43] * * * p<.001	.47[.44;.51] * * * p<.001	.42[.38;.46] * * * p<.001	.44[.39;.49] * * * p<.001
Neighbourhood Deprivation	Relative Neighbourhood Deprivation (most deprived decile)	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Relative Neighbourhood Deprivation (10 - <20%)	.11[.05;.16] * * * p<.001	.10[.04;.16] * * * p<.001	.19[.12;.26] * * * p<.001	.12[.04;.19] * * * p<.001
	Relative Neighbourhood Deprivation (20 - <30%)	.18[.12;.24] * * * p<.001	.23[.16;.29] * * * p<.001	.24[.16;.31] * * * p<.001	.17[.09;.24] * * * p<.001
	Relative Neighbourhood Deprivation (30 - <40%)	.27[.21;.34] * * * p<.001	.28[.22;.34] * * * p<.001	.31[.24;.38] * * * p<.001	.22[.14;.30] * * * p<.001

	Relative Neighbourhood Deprivation (40 - <50%)	.29[.23;.36] * * * p<.001	.33[.26;.39] * * * p<.001	.31[.24;.39] * * * p<.001	.25[.17;.32] * * * p<.001
	Relative Neighbourhood Deprivation (50 - <60%)	.36[.30;.42] * * * p<.001	.41[.34;.47] * * * p<.001	.36[.29;.44] * * * p<.001	.28[.20;.36] * * * p<.001
	Relative Neighbourhood Deprivation (60 - <70%)	.43[.37;.50] * * * p<.001	.47[.40;.53] * * * p<.001	.49[.41;.56] * * * p<.001	.36[.28;.44] * * * p<.001
	Relative Neighbourhood Deprivation (70 - <80%)	.49[.42;.55] * * * p<.001	.57[.50;.63] * * * p<.001	.49[.41;.56] * * * p<.001	.46[.38;.54] * * * p<.001
	Relative Neighbourhood Deprivation (80 - <90%)	.48[.42;.55] * * * p<.001	.56[.50;.63] * * * p<.001	.49[.41;.56] * * * p<.001	.45[.37;.53] * * * p<.001
	Relative Neighbourhood Deprivation (least deprived decile)	.60[.54;.66] * * * p<.001	.68[.62;.75] * * * p<.001	.62[.54;.69] * * * p<.001	.55[.48;.63] * * * p<.001
Composite	Composite SEC	.28[.26;.29] * * * p<.001	.32[.30;.33] * * * p<.001	.28[.26;.29] * * * p<.001	.28[.26;.30] * * * p<.001

All coefficients taken from models adjusted for gender, ethnicity and English as an additional language (EAL).
*p<.05 ; ** = p<.01; *** p<.001.

Appendix J

Appendix J contains the methods and results for the sensitivity analysis whereby age 14 SEC predictor variables were used to predict age 14 vocabulary.

Rationale

Our main analysis used SES indicators taken at age 3. We found that the strongest associations were with age 5 language ability. We conducted a sensitivity analysis with age 14 SES indicators, to check whether this result was due to the proximity of the SES exposure to the age 5 language outcome. We therefore predicted age 14 language with age 14 SES indicators, using the same methodology as the main analyses (see methods in main manuscript)

Method

Vocabulary measures.

Ages 3 & 5. Cohort members completed the Naming Vocabulary BAS II subscale, as a measure of expressive vocabulary. Cohort members were shown a series of images and were asked to name each item in the image(Moulton, 2020).

Age 11. Cohort members completed the Verbal Similarities BAS II subscale. This is a measure of verbal reasoning and verbal knowledge. Three words were read out to the cohort member, usually by the interviewer, and cohort members had to say how the words were related to each other(Moulton, 2020).

Age 14. Word Activity task. This test was a subset of items from the Applied Psychology Unit (APU) Vocabulary Test(Closs, 1986). Cohort members were given a list of 20 target words, each presented alongside 5 other words. Cohort members had to choose the word which meant the same, or nearly the same as the target word, from the 5 options(Fitzsimons, 2017; Moulton, 2020). Total scores out of 20 were converted into z scores for analyses.

Measures of Socioeconomic Circumstance

Five indicators of family SEC were used: parent education, family income, wealth, occupational status and relative neighbourhood deprivation. Operationalisation of these variables is discussed below. These were taken from the age 14 sweep of the MCS2001 cohort.

Parent education. As a measure of parent's education, highest household NVQ level was used (both academic and vocational qualifications derived into NVQ levels 1-5, with level 5 equating to higher qualifications).

Family income. UK OECD weighted income quintiles were used (an indication of household income 1=lowest, 5=highest, accounting for family size).

Wealth. A measure of total net wealth, taken from the age 14 sweep of the MCS2001 cohort. This measure was derived from 4 variables: amount outstanding on all mortgages, house value, amount of investments and assets, and amount of debts owed. Outstanding mortgages were subtracted from the house value, to give a measure of housing wealth. Debts owed were taken from the amount of investments and assets, to give a measure of financial wealth. Housing wealth and financial wealth were then summed to give an overall measure of total net wealth.

Occupational status. Highest household occupational status (National Statistics Socioeconomic Classification (NS-SEC) 3 categories: higher managerial; intermediate; routine, with a fourth category for those who were unemployed) at 14 years.

Relative neighbourhood deprivation. Indices of multiple deprivation (IMD) are the government official measure of relative deprivation (Mclennan, 2019). We used IMD deciles at age 14 (with 1= most deprived and 10=least deprived) as a measure of relative neighbourhood deprivation.

Analyses.

Language scores at age 14 were considered as the outcome variable. Separate models were conducted for SEC measure when the cohort members were aged 3, and when they were aged 14. Drop-one analyses were used to assess the unique contribution of each SEC predictor; a model with all 5 SEC predictors was compared to models with each predictor removed in turn (see main manuscript). A composite factor was included as the predictor variable, adjusting for the potential confounding variables. Results for models considering age 3 SEC predictors of age 14 language ability were compared to that of models considering age 14 SEC predictors of age 14 language ability.

Results

Partial R^2 values for age 3 SEC indicators predicting age 14 vocabulary, compared to age 14 SEC indicators predicting age 14 vocabulary, can be found in Table S7 and Figure S9. With the exception of parent education and occupational status, individual indicators measured at age 14 contributed more variance to age 14 vocabulary. Regression coefficients can be found in Table S8 and are plotted in Figures S10 and S11. Figure S10 displays the regression coefficients for age 3 SEC indicators predicting age 14 vocabulary, compared to age 14 SEC indicators predicting age 14 vocabulary, whilst Figure S11 shows the age 14 SEC coefficients plotted against the main analysis results for all ages. As can be seen from Figure S10, the slopes are similar in steepness, regardless of which age SEC indicators were measured, although the age 14 SEC measures indicate wider inequalities than the age 3 measures. However, when compared to vocabulary at other ages, the main pattern of results remains (see Figure S11): inequalities are widest at the age of 5 and remain persistently wide throughout childhood and into adolescence. Proximity of the SEC measure to age 14 vocabulary does not appear to affect the main pattern of results.

Model comparisons were conducted to determine whether each SES predictor contributed unique variance in vocabulary; a model with all indicators included simultaneously was compared to a model with each removed in turn. All age 14 SES indicators predict unique variance in age 14 vocabulary: Compared to a model without parent education, a model with all SEC predictors was a significantly better fit (Dm(5, 9215)= 26.86, p<.001). Compared to a model without income, a model with all SEC predictors was a significantly better fit (Dm(4, 2560.91)= 5.02, p<.001). Compared to a model without wealth, a model with all SEC predictors was a significantly better fit to the data (Dm(4, 2188.22)= 11.12, p <.001). Compared to a model without occupational status, a model with all SEC predictors was a significantly better fit to the data (Dm(3,8985.4)= 7.98, p<.001). Finally, compared to a model without relative neighbourhood deprivation, a model with all SEC predictors was a significantly better fit to the data (Dm(9, 10706.37)= 5.11, p<.001).

These findings are in line with that of the main analysis, with the exception of relative neighbourhood deprivation. When measured at the age of 3, relative neighbourhood deprivation did not contribute unique variance in age 14 vocabulary. This perhaps indicates that the proximity of neighbourhood deprivation is important regarding age 14 vocabulary.

Table S7. ***Model R^2 for age 3 SEC predictors and age 14 SEC predictors predicting age 14 language***

Indicator	Age 3 SEC measures	Age 14 SEC measures
Parent Education	7.1	5.5
Income	4.3	4.4
Wealth	3.4	4.6
Occupation	5.3	3.1
Relative Neighbourhood Deprivation	2.6	3.9
SEC composite	7.4	7.7
All predictors simultaneously	8.54	8.74

R^2 of models adjusted for gender, ethnicity and English as an additional language.

Figure S9. ***Partial R^2 Values for SEC indicators (Ages 3 & 14) for predicting Age 14 Vocabulary***

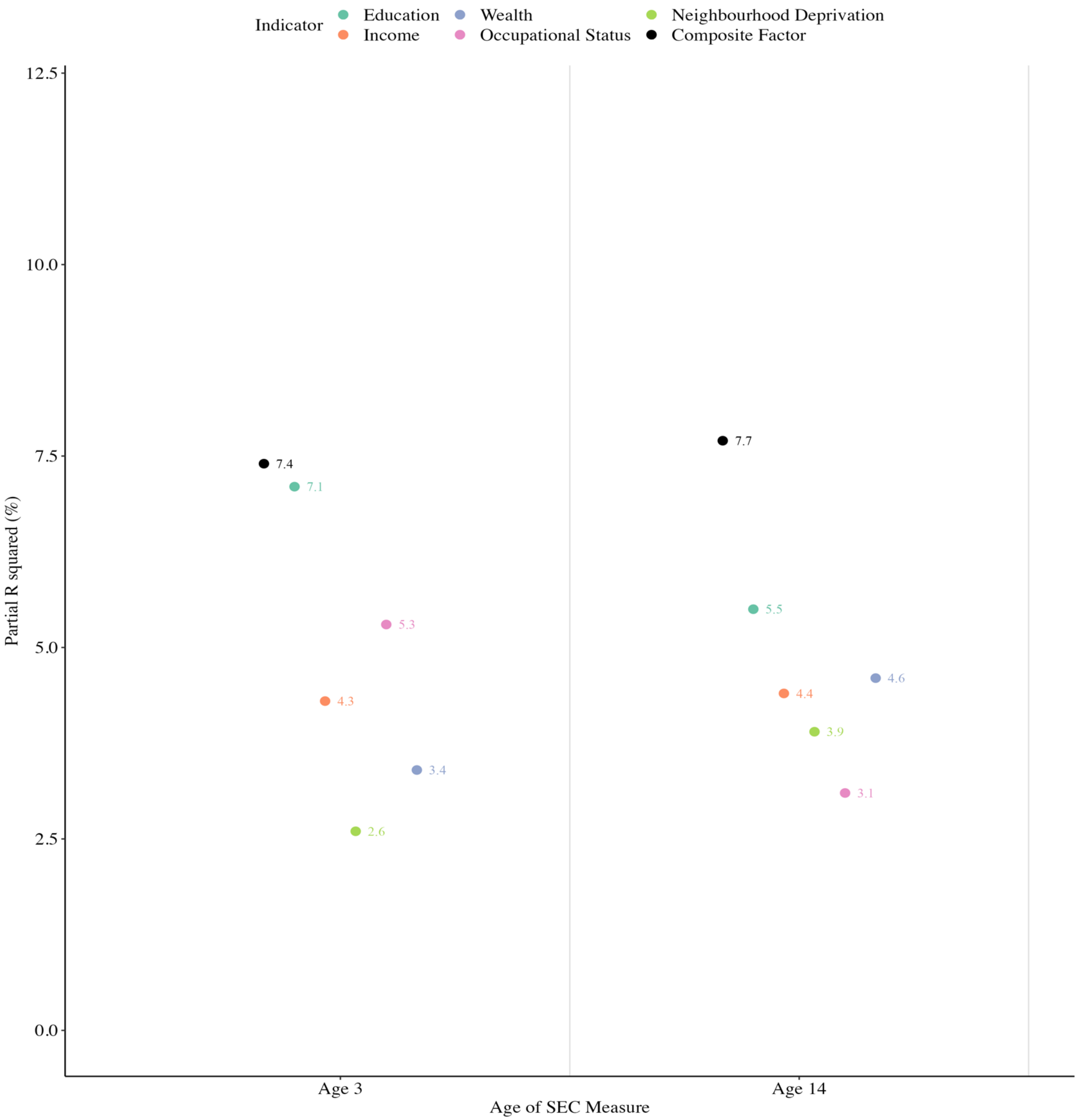


Table S8: Age 14 sensitivity check with Age 14 SES measures: [95% CIs]

		β [95% CIs] *p value*	
	Indicator	Age 14 Vocabulary	
Parent Education	NVQ1	REFERENCE	
	None of these/ over-sees qualifications	-.04[-.14;.05] p= .360	
	NVQ2	.18[.10;.27] * * * p<.001	
	NVQ3	.29[.20;.38] * * * p<.001	
	NVQ4	.51[.43;.59] * * * p<.001	
	NVQ5	.68[.59;.77] * * * p<.001	
Income	Income Quintile 1		REFERENCE
	Income Quintile		.15[.09;.21] * * * p<.001
	Income Quintile 3		.24[.18;.30] * * * p<.001
	Income Quintile 4		.42[.36;.48] * * * p<.001

	Income Quintile 5	.63[.57;.70] * * * p<.001			
Wealth	Wealth Quintile 1		REFERENCE		
	Wealth Quintile 2		.15[.09;.21] * * * p<.001		
	Wealth Quintile 3		.26[.19;.32] * * * p<.001		
	Wealth Quintile 4		.39[.33;.45] * * * p<.001		
	Wealth Quintile 5		.60[.54;.66] * * * p<.001		
Occupational Status	Routine			REFERENCE	
	Unemployed			-.07[-.12;-.02] * * * p<.001	
	Intermediate			.18[.13;.24] * * * p<.001	
	Higher managerial			.41[.36;.47] * * * p<.001	
Relative neighbour-	Most deprived decile				REFERENCE
	10 - <20%				.10[.02;.17] * * p= .010

	20 - <30%	.14[.06;.22] * * * p<.001	
	30 - <40%	.23[.16;.31] * * * p<.001	
	40 - <50%	.29[.21;.36] * * * p<.001	
	50 - <60%	.36[.28;.44] * * * p<.001	
	60 - <70%	.45[.37;.53] * * * p<.001	
	70 - <80%	.43[.35;.51] * * * p<.001	
	80 - <90%	.50[.42;.58] * * * p<.001	
	Least deprived decile	.66[.58;.74] * * * p<.001	
Composite	Composite SEC		.28[.26;.30] * * * p<.001

Figure S10. ***Relationships between SEC indicators (Ages 3 & 14) and Vocabulary (Age 14)***

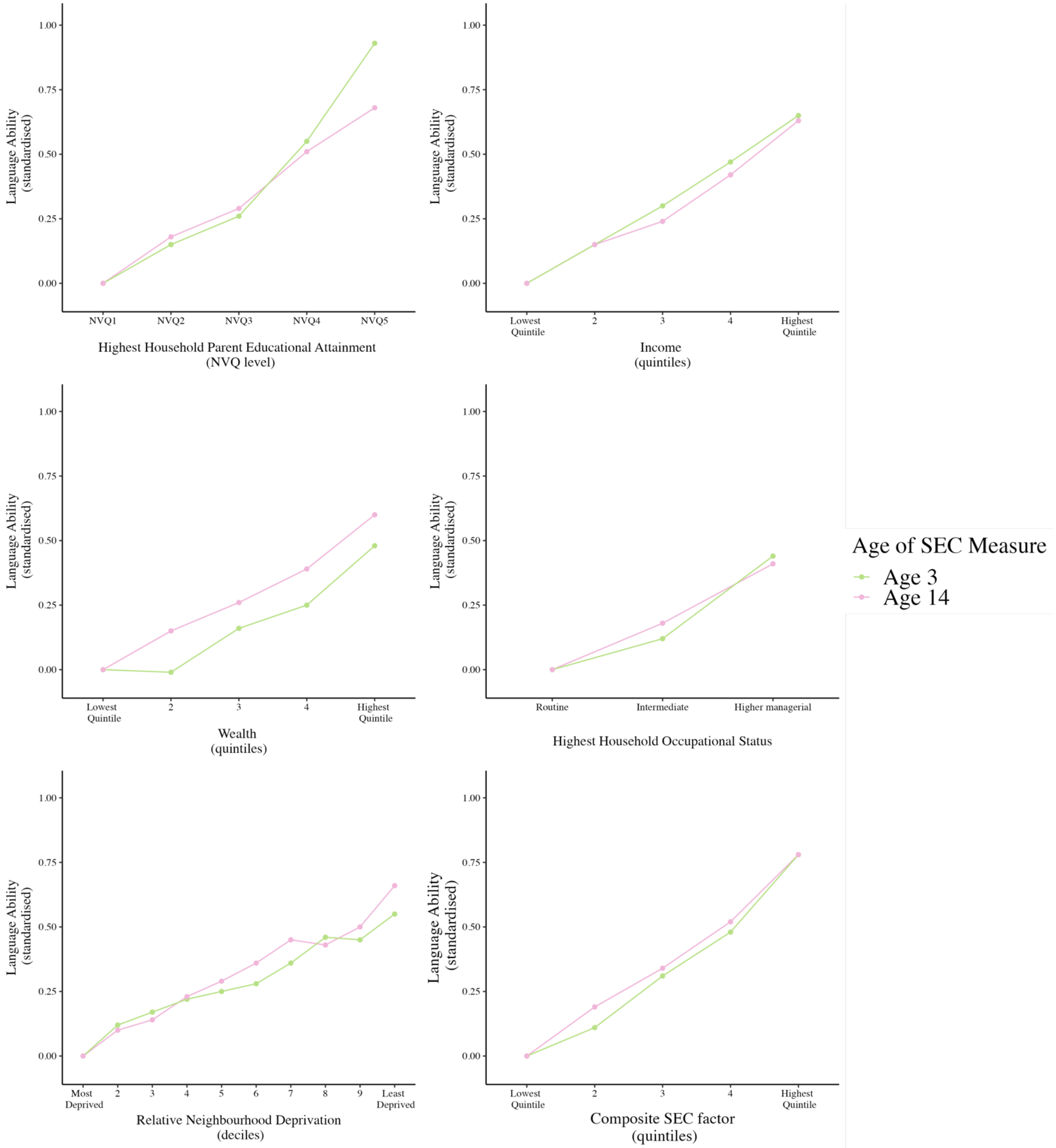
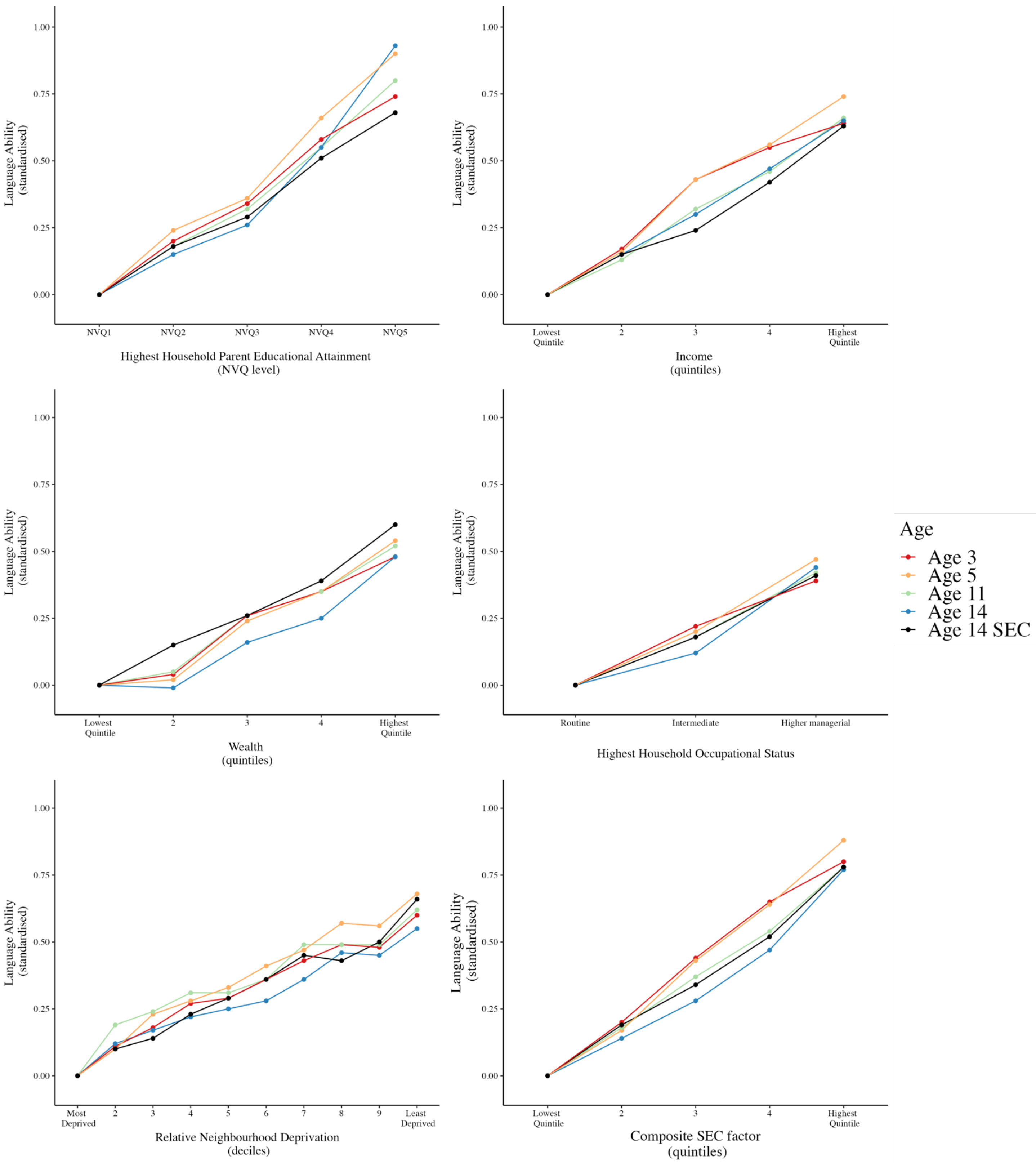


Figure S11. ***Relationships between SEC indicators and Vocabulary at ages 3, 5, 11 and 14. Age 14 vocab predicted by age 3 and age 14 SEC indicators***



Appendix K

Appendix K contains the descriptive statistics and regression coefficients for the cross-cohort comparison.

Table S9. *Descriptive statistics for language measure by SEC group in each cohort.*

	BCS1970			MCS2001		
	Early childhood language *Range=0-56*	Late childhood language *Range=0-20*	Adolescent language *Range=0-74*	Early childhood language *Range=10-170*	Late childhood language *Range= 10-179*	Adolescent language *0-20*
Parent Education: No /low level qualifications	32.12(11.36), [31.87;32.37]	11.37(2.6),[11.32;11.43]	38.66(13.24), [38.38;38.95]	99.49(16.97) [98.93;100.04]	113.52(18.58), [112.91;114.13]	6.08(2.35), [6;6.16]
Parent Education:O levels/GCSEs grades A*-C	36.49(9.93), [36.14;36.84]	12.32(2.46), [12.23;12.41]	42.31(12.55), [41.86;42.75]	106.66(14.97), [106.24;107.07]	118.12(16.77), [117.66;118.59]	6.57(2.32),[6.51;6.63]
Parent Education: Post 16 education	37.9(10.22), [37.31;38.49]	12.8(2.42), [12.66;12.94]	44.63(12.06), [43.93;45.32]	110.26(14.29), [109.78;110.73]	122.22(14.93), [121.72;122.71]	7.07(2.41),[6.99;7.15]
Parent Education: University level	39.45(10.1), [39.06;39.84]	13.41(2.38), [13.31;13.5]	48.05(11.62), [47.6;48.5]	114.88(14.56) [114.43;115.34]	126.7(14.04), [126.26;127.14]	8.28(2.81), [8.19;8.37]
Income (Quintile 1 - Lowest)	30.38(11.64), [29.98;30.79]	10.99(2.68),[10.9;11.09]	37.65(13.43), [37.18;38.12]	101.32(17.37), [100.75;101.89]	115.78(18.38), [115.18;116.38]	6.4(2.4), [6.32;6.48]
Income (Quintile 2)	33.57(11.28), [33.16;33.98]	11.78(2.6),[11.69;11.87]	40.12(13.32), [39.63;40.6]	104.47(15.76) [103.93;105.01]	116.73(16.95), [116.15;117.31]	6.53(2.37), [6.45;6.61]
Income (Quintile 3)	35.07(10.73), [34.69;35.45]	11.99(2.54), [11.9;12.07]	41.18(12.83), [40.73;41.63]	107.99(15.45), [107.46;108.52]	120.11(16.5), [119.55;120.68]	6.84(2.53), [6.75;6.92]
Income (Quintile 4)	36.52(10.26), [36.16;36.88]	12.46(2.49), [12.37;12.55]	43.17(12.59), [42.73;43.62]	111.81(14.32), [111.3;112.32]	122.4(15.61), [121.84;122.95]	7.28(2.65), [7.19;7.38]
Income (Quintile 5 - Highest)	38.88(10.03), [38.49;39.27]	13.14(2.4), [13.05;13.23]	46.26(12.32), [45.78;46.74]	114.03(13.99), [113.52;114.54]	125.68(14.54), [125.14;126.21]	7.93(2.75), [7.83;8.03]

Occupational Status (Unemployed)	30.95(11.85), [28.7;33.19]	11.56(3.07), [10.98;12.14]	39.31(13.24), [36.81;41.81]	100.2(16.8), [99.63;100.78]	114.16(18.55), [113.53;114.79]	6.19(2.4), [6.11;6.27]
Occupational Status (Routine)	30.59(11.7), [30.1;31.08]	11.01(2.63), [10.9;11.12]	37.43(13.19), [36.88;37.98]	104.92(15.73), [104.42;105.42]	117.12(17.5), [116.57;117.67]	6.56(2.36), [6.48;6.63]
Occupational Status (Intermediate)	34.15(10.78), [33.91;34.39]	11.78(2.55), [11.73;11.84]	40.34(13.11), [40.05;40.63]	108.3(15.63), [107.74;108.86]	120.56(15.85),[119.99;121.13]	6.86(2.45), [6.77;6.95]
Occupational Status (higher managerial)	37.52(10.84), [37.21;37.83]	12.86(2.53), [12.78;12.93]	45.1(12.58), [44.74;45.47]	113.56(13.92), [113.2;113.92]	124.85(14.46), [124.48;125.22]	7.74(2.72), [7.68;7.81]

Table S10. ***Associations between SEC and language ability in the MCS2001 and BCS1970 cohorts in early childhood, late childhood and adolescence***

	Indictor	Early Childhood Vocabulary (BCS)	Late Childhood Vocabulary (BCS)	Adolescent Vocabulary (BCS)	Early Childhood Vocabulary (MCS)	Late Childhood Vocabulary (MCS)	Adolescent Vocabulary (MCS)
Parent Education	No/low level qualifications	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	O levels/ GCSEs grades A*-C	.35[.30;.40] * * * p<.001	.35[.30;.39] * * * p<.001	.27[.21;.33] * * * p<.001	.30[.26;.35] * * * p<.001	.25[.20;.30] * * * p<.001	.18[.13;.23] * * * p<.001
	Post 16 education	.48[.41;.54] * * * p<.001	.53[.45;.60] * * * p<.001	.43[.35;.52] * * * p<.001	.54[.50;.59] * * * p<.001	.50[.45;.54] * * * p<.001	.37[.32;.43] * * * p<.001
	University level qualifications	.65[.60;.69] * * * p<.001	.76[.71;.82] * * * p<.001	.70[.63;.77] * * * p<.001	.85[.80;.89] * * * p<.001	.76[.71;.81] * * * p<.001	.84[.79;.89] * * * p<.001
Occupa-	Routine	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Intermediate	.28[.23;.34] * *	.28[.23;.33] * * *	.21[.15;.27] * * *	.31[.27;.36] * * *	.25[.20;.30] * * *	.13[.08;.19] * * *

		* p<.001	p<.001	p<.001	p<.001	p<.001	p<.001
	Higher managerial	.62[.56;.68] * * * p<.001	.70[.64;.76] * * * p<.001	.58[.51;.65] * * * p<.001	.44[.39;.49] * * * p<.001	.42[.37;.48] * * * p<.001	.31[.25;.36] * * * p<.001
	Unemployed	.09[-.16;.35] p= .480	.25[.02;.48] * p= .030	.19[-.05;.43] p= .120	.06[.01;.11] * p= .030	.03[-.03;.09] p= .370	-.02[-.08;.03] p= .410
Income Quintiles	Quintile 1 (Most Deprived)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Quintile 2	.25[.19;.32] * * * p<.001	.29[.23;.34] * * * p<.001	.18[.10;.25] * * * p<.001	.63[.58;.68] * * * p<.001	.56[.51;.61] * * * p<.001	.51[.46;.57] * * * p<.001
	Quintile 3	.37[.31;.43] * * * p<.001	.36[.30;.41] * * * p<.001	.26[.19;.33] * * * p<.001	-.26[-.31;-.20] * * * p<.001	-.17[-.22;-.12] * * * p<.001	-.14[-.19;-.08] * * * p<.001
	Quintile 4	.48[.42;.54] * * * p<.001	.53[.47;.59] * * * p<.001	.40[.33;.46] * * * p<.001	.20[.15;.25] * * * p<.001	.20[.15;.25] * * * p<.001	.12[.06;.18] * * * p<.001
	Quintile 5 (Least Deprived)	.70[.64;.76] * * * p<.001	.79[.74;.85] * * * p<.001	.63[.55;.72] * * * p<.001	.48[.44;.52] * * * p<.001	.45[.40;.49] * * * p<.001	.46[.41;.51] * * * p<.001

All coefficients taken from models adjusted for gender, ethnicity, English as an additional language (EAL) and age of cohort member at the time of the language test.
*p<.05
** = p<.01
*** p<.001

Appendix L

Appendix L contains the methods and results for the sensitivity analysis that used Ridit scores in the cross cohort comparison.

Rationale

The education system and occupational structure of the UK has changed over the period that separates the BCS1970 and MCS2001 cohorts, leading to changes in the composition of these two SEC indicators. We therefore conducted a supplementary analysis to our cross-cohort comparison, whereby highest household occupational status and highest household educational attainment were converted to Ridit scores to aid comparability across cohorts (Donaldson, 1998). Ridit scores put ordered categories onto a scale of 0-1, based on the distribution of the categories within any dataset. The resulting coefficients of regression models with SEC Ridit scores as the predictor provide the slope index of inequality (SII). The SII represents the estimated absolute inequalities in an outcome (here, vocabulary) between the highest and lowest SEC groups (Bann, Johnson, Li, Kuh, & Hardy, 2018; Renard, Devleesschauwer, Speybroeck, & Deboosere, 2019; WHO, 2013) and accounts for the changes in the composition of the SEC indicator (Regidor, 2004; WHO, 2013). Therefore, this method allows us to compare inequalities in vocabulary in two cohorts, despite the underlying distributions of SEC variables differing across cohorts. However, as this is an absolute measure of inequalities, this method is not able to discern gradients within the distribution and so hence this method forms our supplementary analysis.

Method

Highest household educational attainment and highest household occupational status were converted to Ridit scores separately in each cohort. The *toridit()* function from the ridittools package in R was used (Bohlman, 2018). Ridit scores were calculated for each imputed dataset and regression models, where the Ridit score was the main predictor, were ran for each imputed dataset. The results of the regression models were then pooled. This procedure was conducted on separate regression models, where early childhood vocabulary, late childhood vocabulary and adolescent vocabulary in each cohort were the outcome variables. This results in 9 separate regression models in each cohort (see Table S11). All models controlled for gender, ethnicity and English as an additional language (EAL).

Results

Regression coefficients can be found in Table S11. Because our Ridit scores rank occupation and education from the lowest SEC to the highest SEC, positive coefficients are indicative of higher vocabulary abilities among the highest SEC group (WHO, 2013). Coefficients indicate better vocabulary scores in the most advantaged group. This is the case for all ages and in both cohorts.

The results from this supplementary analysis confirm the results of the main cross cohort comparison analysis. As can be seen from Table S11, inequalities based on highest household education are largest in the MCS2001 cohort for early childhood and adolescent vocabulary, but in the BCS1970 cohort for late childhood language ability. Turning to highest household occupational status, inequalities are largest for vocabulary at all ages in the BCS1970 cohort, indicated by the bigger coefficients for this cohort. A comparison of the partial R^2 values for the main analysis and Ridit score analysis can be found in Table S12 and Figure S12. These are similar across both analyses.

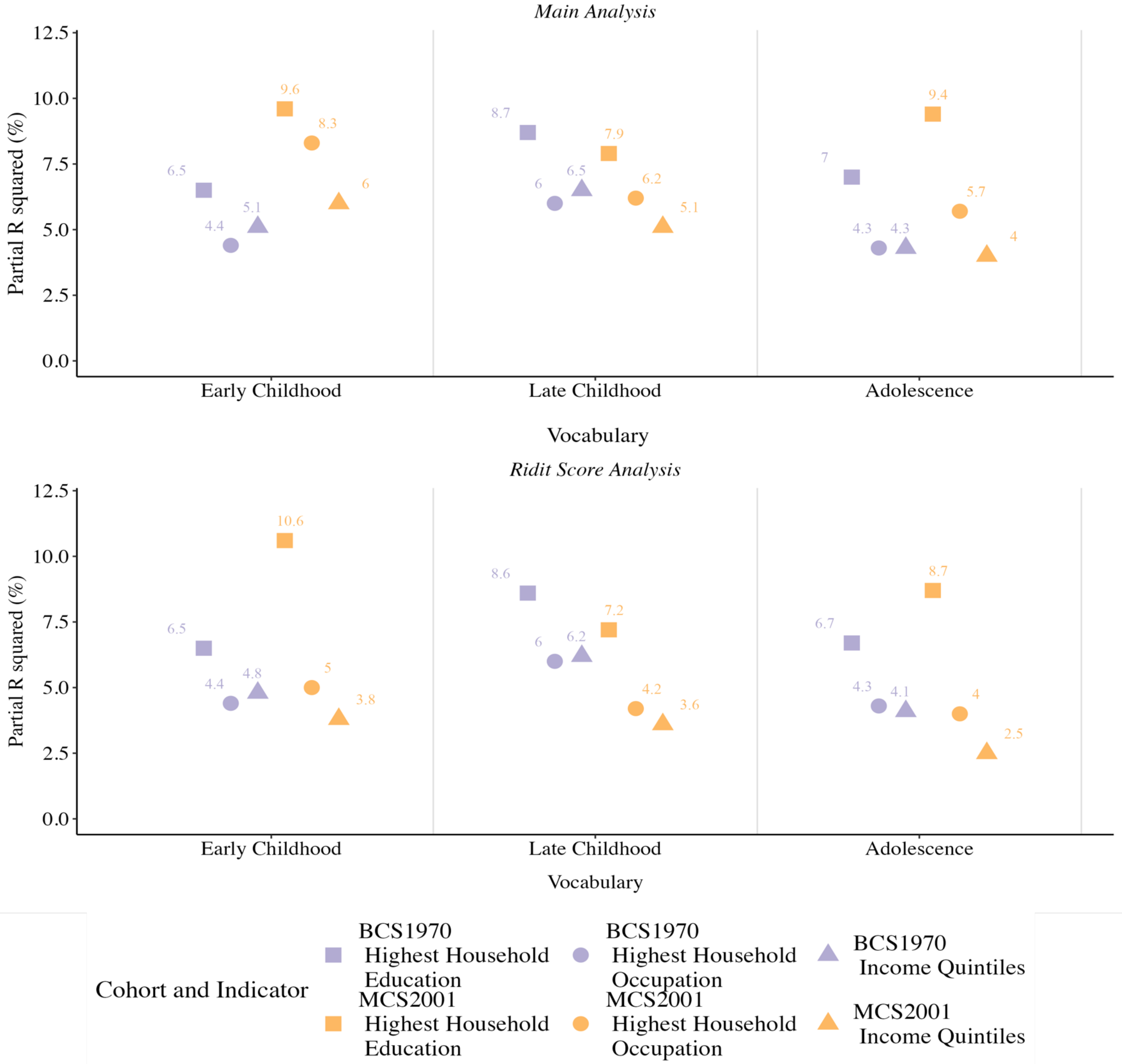
Table S11. ***Regression coefficients for models predicting vocabulary using SEC Ridit scores***

	β [95% CIs] *p value*					
	Highest Household Education (Ridit score)		Highest Household Occupation (Ridit score)		Income (Ridit score)	
Vocabulary	1970	2001	1970	2001	1970	2001
Early Childhood	.98[.92;1.05]*** *p<.001*	1.08[1.03;1.14]*** *p<.001*	.81[.73;.88]*** *p<.001*	.78[.71;.84]*** *p<.001*	.76[.47;1.06]*** *p<.001*	.64[.59;.70]*** *p*<.001
Late Childhood	1.12[1.05;1.19]*** *p<.001*	.98[.93;1.04]*** *p<.001*	.94[.86;1.01]*** *p<.001*	.72[.65;.78]*** *p<.001*	.86[.53;1.19]*** *p<.001*	.62[.56.69]*** *p<.001*
Adolescent	.99[.89;1.08]*** *p<.001*	1.07[1.01; 1.13]*** *p<.001*	.79[.70;.88]*** *p<.001*	.76[.69;.84]*** *p<.001*	.70[.42;.97]*** *p<.001*	.57[.49; .65]*** *p<.001*

Table S12. ***Partial R^2 values for Ridit scores predicting vocabulary throughout childhood in two cohorts***

	Partial R^2 (%)					
	Highest Household Education (Ridit score)		Highest Household Occupation (Ridit score)		Income (Ridit score)	
	1970	2001	1970	2001	1970	2001
Early Childhood Vocabulary	6.5	10.6	4.4	5	4.8	3.8
Late Childhood Vocabulary	8.6	7.2	6	4.2	6.2	3.6
Adolescent Vocabulary	6.7	8.7	4.3	4	4.1	2.5

Figure S12. ***Partial R^2 Values for SEC indicators in each cohort: Comparison of Main and Ridit Score Analyses***



Appendix M

Appendix M contains the sensitivity analysis for the cross-cohort comparison which included only those of a White ethnicity.

Method

Vocabulary measures (cross-cohort comparison).

Early language ability. For the BCS1970 cohort, receptive vocabulary was measured at age 5 using the English Picture Vocabulary Test (EPVT), a UK version of the Peabody Picture Vocabulary Test (Brimer, 1962; Dunn, 1965). Cohort members were shown 56 sets of four diverse images and heard a specific word associated with each set of four images. They were asked to select one picture that matched the presented word and were awarded one point for every correct response(Moulton, 2020; Parsons, 2014). For the MCS2001cohort, expressive vocabulary was measured using the naming vocabulary sub-test of the BAS II (Colin D. Elliott, 1996). We adjusted for age in months at the time of the test in both cohorts. All scores and ages were converted to z scores for analyses.

Late childhood language ability. When the BCS1970 cohort members were aged 10, they completed the BAS word similarities subscale (Elliott, 1979). The test was made up of 21 items, each of which consisted of three words. The teacher read these sets of items out loud and cohort members had to a) name another word that was consistent with the three words in the item and b) state how the words were related. In order to receive a point, cohort members had to correctly answer both parts of the question (Moulton, 2020; Parsons, 2014). Details on the scoring of this vocabulary measure and the SPSS syntax used can be found in appendix 3 of "Childhood Cognition in the 1970 British Cohort Study" (Parsons, 2014). When MCS2001 cohort members were aged 11, they completed the BAS II verbal similarities subscale (detailed above). As already mentioned, test scores for the MCS2001 cohort were adjusted for item difficulty. In both cohorts, we controlled for age at the time of the test and converted all scores to z scores.

Adolescent language ability. When aged 16, BCS1970 cohort members completed the APU Vocabulary Test (Closs, 1986). This consisted of 75 items: an item consisted of a target word, presented with a multiple-choice list, from which cohort members had to select a word that meant the same as the target word(Moulton, 2020; Parsons, 2014). These items got progressively harder throughout the test. Details on the scoring of this vocabulary test can be found in appendix 3(Parsons, 2014). When MCS2001cohort members were aged 14, they completed the Word Activity Task (detailed above). Words used in the Word Activity Task were a subset of the words used in the BCS1970 cohort Vocabulary Test, which cohort members completed aged 16(Fitzsimons, 2017). Scores were adjusted for age and converted to *z* scores for analyses.

Indicators of socioeconomic circumstance.

Harmonised measures of the following two indicators were used as measures of SEC:

Parent education. The highest academic qualification achieved in the household when the cohort member was aged 5. Where this information is missing, information from previous sweeps was used.

Occupational status. Highest household occupational status at age 5. For the BCS1970 cohort, this was ascertained with the Registrar General's classification. For the MCS2001cohort, the NS-SEC classification system was used. Where this information is missing, information from previous sweeps was used.

Analysis plan.

We had 3 separate outcome variables in each cohort (early childhood language ability, late childhood language ability and adolescent language ability). We built two regression models per outcome, one with occupational status as the predictor variable and the other with parent education as the predictor variable. Because our measures of language ability were standardised within each cohort, we were able to directly compare coefficients between cohorts and establish the rate of inequality in language ability at each age in the two cohorts.

Results

Partial R^2 values can be found in Table S13, and regression coefficients can be found in Table S14. The results from this sensitivity analysis confirm the results of the main cross cohort comparison analysis. As can be seen from Table S14, inequalities based on highest household education are largest in the MCS2001 cohort for early childhood and adolescent vocabulary, but in the BCS1970 cohort for late childhood language ability. Turning to highest household occupational status, inequalities are largest for vocabulary at all ages in the BCS1970 cohort, indicated by the bigger coefficients for this cohort. Thus, the ethnic composition of the two cohorts do not appear to be driving the results of our cross-cohort comparison.

Table S13. ***Partial R2 Values for cross-cohort comparison (%)***

	Partial R^2 (%)					
	Highest Household Education		Highest Household Occupation		Income	
	1970	2001	1970	2001	1970	2001
Early Childhood Vocabulary	7.6	10.2	5.4	9.1	6	6.4
Late Childhood Vocabulary	9.3	7.9	6.5	6.5	6.5	5.1
Adolescent Vocabulary	7.1	9.6	4.7	5.9	4.3	4

Table S14. ***β[95% CIs] for SEC predicting vocabulary in MCS2001 and BCS1970 Cohorts***

		BCS1970 Cohort				MCS2001 Cohort	
		Indicator	Early Childhood Vocabulary	Late Childhood Vocabulary	Adolescent Vocabulary	Early Childhood Vocabulary	Late Childhood Vocabulary
Parent Education	No/low level qualifications	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	O levels/ GCSEs grades A*-C	.37[.32;.41] * * * p<.001	.35[.30;.40] * * * p<.001	.27[.21;.33] * * * p<.001	.30[.26;.35] * * * p<.001	.26[.21;.32] * * * p<.001	.19[.13;.25] * * * p<.001
	Post 16 education	.49[.43;.56] * * * p<.001	.54[.47;.61] * * * p<.001	.44[.36;.53] * * * p<.001	.53[.48;.58] * * * p<.001	.50[.45;.56] * * * p<.001	.39[.32;.45] * * * p<.001
	University level qualifications	.66[.61;.70] * * * p<.001	.78[.73;.83] * * * p<.001	.70[.62;.77] * * * p<.001	.82[.77;.87] * * * p<.001	.76[.70;.81] * * * p<.001	.86[.80;.92] * * * p<.001
Occupational Status	Routine	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Unemployed	.06[-.19;.32] p= .630	.19[-.09;.47] p= .190	.15[-.23;.54] p= .440	.03[-.02;.08] p= .280	.01[-.06;.08] p= .770	-.03[-.09;.04] p= .400
	Intermediate	.30[.25;.35] * * * p<.001	.28[.23;.34] * * * p<.001	.23[.16;.30] * * * p<.001	.29[.24;.34] * * * p<.001	.24[.19;.30] * * * p<.001	.13[.07;.19] * * * p<.001

	Higher managerial	.66[.60;.71] * * * p<.001	.72[.66;.79] * * * p<.001	.61[.53;.69] * * * p<.001	.42[.37;.47] * * * p<.001	.41[.35;.47] * * * p<.001	.30[.24;.36] * * * p<.001
	Quintile 1 (Most Deprived)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
	Quintile 2	.27[.22;.33] * * * p<.001	.28[.22;.33] * * * p<.001	.17[.10;.24] * * * p<.001	.58[.53;.63] * * * p<.001	.53[.48;.59] * * * p<.001	.50[.44;.56] * * * p<.001
Income Quintiles	Quintile 3	.39[.33;.45] * * * p<.001	.34[.28;.40] * * * p<.001	.26[.19;.33] * * * p<.001	-.28[-.34;-.22] * * * p<.001	-.20[-.26;-.15] * * * p<.001	-.16[-.22;-.10] * * * p<.001
	Quintile 4	.51[.45;.57] * * * p<.001	.52[.46;.59] * * * p<.001	.39[.32;.46] * * * p<.001	.19[.14;.23] * * * p<.001	.20[.15;.26] * * * p<.001	.12[.05;.19] * * * p<.001
	Quintile 5 (Least Deprived)	.72[.66;.78] * * * p<.001	.78[.72;.85] * * * p<.001	.63[.54;.72] * * * p<.001	.44[.40;.48] * * * p<.001	.42[.38;.47] * * * p<.001	.46[.40;.51] * * * p<.001