UNIVERSITY OF LIVERPOOL

Socioeconomic inequalities in vocabulary and their implications for educational attainment and mental health in adolescence

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Emma Ruth Thornton

April 2022

Department of Psychology

Institute of Population Health

University of Liverpool

UNIVERSITY OF LIVERPOOL

INSTITUTE OF POPULATION HEALTH

DECLARATION

This thesis is the result of my own work. The material contained in the thesis has not been presented, nor is currently being presented, either wholly or in part for any other degree or qualification.

Signed ..... (Candidate)

14.04.2022

Signed ..... (Candidate)

Acknowledgements

I would first like to thank my incredible supervisory team, Colin Bannard, Praveetha Patalay and Danielle Matthews. Thank you for all the interesting, insightful discussions we've had during our supervision meetings – I always look forward to them. Thank you for support and encouragement, and for helping me become a better researcher and persevere, particularly when things got challenging. Colin – thank you for seeing my potential during the third year of my undergraduate degree and approaching me to apply for this PhD. Thank you for your help, guidance, patience, and support with learning R – I never would have thought at the start of this PhD that I would actually end up enjoying coding! Praveetha, thank you for instilling in me your wisdom and knowledge, especially about the MCS data, which was very intimidating when I first downloaded it. Thank you for organising the online boardgame evenings and work-along sessions over the last couple of years - they made doing a PhD in lockdown much more fun -, for our mutual love and enthusiasm for SushiGo and pretty graphs, and for making me click submit when I tried to put it off even longer (I would probably still be sat on drafts of things to submit now otherwise!). Danielle, thank you for all the extra meetings and long chats to make sure each draft was the highest possible standard, and for giving me the opportunity to work with you on the Tiny Happy People project alongside my PhD. This was such an invaluable and enjoyable experience and has equipped me with additional research skills. I have thoroughly enjoyed working with you all.

I would like to extend my thanks to my PhD pals. Hannah, thank you for reading every draft of my work and for being honest and thorough in your feedback, and most importantly, for being a truly great friend. To Becky, thank you so much for your endless support, encouragement and therapy sessions, for all of our work along zoom calls over the last couple of years, and geeking out with me over MCS data, this experience wouldn't have been the same without you. Leanne, Sarah and Harri, thank you for all the virtual movie nights and making working from home during the pandemic more bearable – I can't wait to blow off steam with you when we finally get to Paris! Amber – I'm so lucky to have met you, thank you for always being on the end of the phone for a good natter, for being such a supportive and genuine, caring friend, and for all the great memories – I can't wait to make many more. Best of luck with the final push for your PhD, I know you will be amazing. Rachel, Alice, and Sophie, thank you for all the much-needed coffee breaks, wholesome Sunday afternoons and roadtrips, and for putting up with me and my weirdness (sorry, and thank you Rachel, for bearing the brunt of this in the office, and being the Scully to my

Hitchcock, the Ant to my Dec). Finally, to the rest of my office buddies, Marie, Ellie, and Patrick – thank you for all the hilariously fun times we had pre-pandemic, you made coming into the office so fun and I miss you all.

To my oldest friends, Becka, Beth, Sarah, and Coles, thank you for being the best group of gals I've ever met and putting up with me since we were 12 – my life would not be the same without you all. Thank you for being weird with me, always looking after me, making me laugh until I can't breathe (or until I aspirate my tea!), or crying together when we've needed to, and for making me take much-needed breaks. Thank you for always listening to me throughout this PhD, even if you've not always known what I'm on about, I appreciate it all the same. Becka – thank you for always providing me with a place to stay and being my tour guide during various work trips to London. Jess, thank you for being an extra set of ears, particularly when I had questions about GCSEs and BTECs, and for always being up for watching Grey's Anatomy on repeat with me!

Josh, thank you for looking at all of my graphs and pretending to be excited with me about the tables I've managed to generate with code, and for being my sounding board over various interpretations of results. Thank you for being you, for sharing my love and enthusiasm for every single dog we pass in the street, for always sending me dog videos and for being my adventure buddy, you have made sure I had a work-life balance whilst doing this PhD.

Last, but by no means least, thank you to my sister Lara for always being interested and asking the world's supply of (sometimes annoying) questions about what I'm up to, and to my sister Olivia, thank you for supplying me with endless cups of tea and biscuits during the write up of this thesis. To my favourite little companion, Leo, thank you for sitting on my feet under my desk to keep them warm, for joining in my supervision meetings, and forcing me outside to take a break, I know you have no idea what I sit and do all day instead of playing with you, I promise I'll pay you more attention now. Finally, thank you Mum and Dad for always supporting me through everything, for always showing an interest in what I'm up to and for being my biggest cheerleaders, I'm lucky to have you. Thank you for letting me live at home whilst I complete this PhD, and for the endless supplies of Diet Coke and ice cream whilst I did, but don't worry, I'll finally fly the nest soon – don't miss me too much.

I am grateful to the cohort members of the 1970 British Cohort Study, the Millennium Cohort Study, and their families. I am 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. Finally, This PhD would not have been possible without the studentship awarded by the ESRC NWSSDTP – thank you.

Abstract

This thesis examines the extent to which childhood vocabulary can be used to predict adolescent mental health and educational outcomes, and to which differences in early vocabulary are associated with socioeconomic circumstances (SEC). These questions arise out of an increased interest in the relation between SEC and vocabulary, and efforts to develop interventions to promote vocabulary prior to school entry, which have so far been met with mixed success. Since vocabulary is thought to predict later mental health and education outcomes, and there are social inequalities in language, education and mental health, this thesis sought to provide insight into whether early vocabulary interventions are likely to benefit children in the longer term, through secondary data analysis of two large, nationally representative UK datasets: 1) The 1970 British Cohort Study (BCS1970), and 2) The Millennium Cohort Study (MCS2001).

I first examined whether early child vocabulary is related to adolescent mental health, finding small effects, which importantly differed as a function of reporter: when adolescent self-report was considered, better childhood vocabulary skill predicted *poorer* adolescent mental health outcomes.

I next looked more closely at the relation between SEC and vocabulary, by investigating socioeconomic inequalities in vocabulary throughout childhood and into adolescence. I found multiple SEC indicators (most notably parent education, income and occupation) uniquely predicted child and adolescent vocabulary. Inequalities persist from ages 3 to 14 years, with effects being most pronounced at the start and end of formal schooling.

I finally investigated whether vocabulary at school entry predicted educational outcomes at the end of secondary school. Here, unlike for mental health, there was a clear relation: better childhood vocabulary predicted better educational outcomes on GCSE or equivalent examinations. This effect was substantially moderated by SEC. Thus, not all children benefit from strong early vocabulary skills in the same way.

Overall, the findings of this thesis suggest that good vocabulary skill, as measured by standardised tests, is important for educational attainment but not internalising mental health. Effective interventions are potentially well placed to improve educational outcomes. However, moderation effects suggest such early interventions alone may not suffice, since the educational benefits of good early vocabulary do not appear stable over SEC strata. In sum, although early language interventions are well placed to improve educational outcomes,

in order to improve wider functioning in adolescence, we also need to directly target internalising mental health and structural inequalities.

Table of Contents

List of Figures

List of Tables

Chapter 1 : General Introduction

The ability to effectively communicate with others is important for success in all walks of life, allowing us to build and maintain relationships, articulate thoughts, ideas, needs and desires to others, and develop an understanding of, and knowledge about, the world around us (Bercow, 2008; Law, Charlton, Dockrell, et al., 2017). Fundamental to communication skills is language development and oral language skills. Early language skills are often claimed to be an important causal factor in long-term life outcomes, such as mental health, education and employment (Field, 2010). Language and communication skills are also important for success in the economy: in a society that is becoming increasingly knowledge based and technological, there is an increased reliance on cognitive skills, including language ability (Beddington et al., 2008), and such skills are highly valued by, and are central to, the modern workforce (Deloitte, 2016).

There is wide variance in language skills by the time children start school. Some children are slow to develop language initially and early delays are subsequently resolved. However, other children experience persistent delays throughout childhood (Law, Charlton, Dockrell, et al., 2017). The prevalence of persistent delays is estimated to be around 10% of all children, and this is thought to increase in areas of social disadvantage (Norbury et al., 2016). There is a well-established link between socioeconomic circumstance (SEC) and language ability, with those from more advantaged backgrounds, on average, performing better on standardised measures of language than their disadvantaged counterparts (Pace et al., 2017). The risk of delayed or poor language is experienced disproportionately by children from low SEC backgrounds, and these initial inequalities in language ability tend to affect other areas of development (Bercow, 2018; Law, Charlton, & Asmussen, 2017). There are likely to be high societal costs that result from the impacts of language difficulties: for example, those with poor vocabulary in childhood (defined as vocabulary scores <2 standard deviations below the mean, with scores on a measure of non-verbal cognition >1 standard deviation below the mean) are at least twice as likely to be unemployed in adulthood (Law, Rush, Schoon, & Parsons, 2009), and there is thought to be an overrepresentation of people with poor language skills in the criminal justice system (Bryan, 2004), with links between language difficulties and behavioural difficulties, and comorbidities between the two being common (Bornstein et al., 2013; Chow & Wehby, 2018). Disordered language may also impact on friendships and bullying in childhood and adolescence (see Durkin & Conti-Ramsden, 2010, for a review), and can place individuals experiencing such impairments at an

educational disadvantage: for example, Johnson et al., (2010) found that those with a history of language disorders were less likely to finish secondary or tertiary education by the age of 25. Evidence from studies on individuals with language impairments can give insight into the possible outcomes for those who are slower to develop their language, but still fall within the typically developing range.

Due to the proposed link between early language skills and later life outcomes, arguments have been made that early language should be considered as a child wellbeing indicator, in a similar vain to obesity and diet (Law, Charlton, & Asmussen, 2017). Public Health England (PHE), before they were replaced in 2020, had an overarching aim of protecting and improving the health and wellbeing of people in England, and the reduction of health inequalities (Public Health England, 2020b). Early language ability is now viewed as a public health wellbeing indicator, with PHE devising a speech, language and communication pathway, aiming to improve language abilities of all children, and reduce the number of children who do not reach the expected level in communication, language and literacy at the end of their first year of school by 50% by the year 2028 (Public Health England, 2020a). Furthermore, the Marmot Review highlighted children's early years, including their language and communication skills, as fundamental to later health inequalities (Marmot, 2010) and child language ability is now seen as a key strategy for improving social mobility in England (Department for Education, 2017d). There are also many charities and initiatives in the UK promoting language development, such as ICAN, The Communication Trust, Early Intervention Foundation and Education Endowment Foundation, to name a few.

Despite this widespread attention to early language skills, it is thought that there is still a lack of understanding about their importance, especially after the early years, among decision makers and professionals. Interest in promoting language skills with the aim of improving educational attainment and wellbeing is widespread, given the assumed importance of language skills for these outcomes. Many studies report a relation between SEC and child language outcomes, and it is thought that such effects may result from the impact of SEC on caregiver input or the quality of early childhood education settings (see section 1.5 and Chapter 4). As a result, many parenting interventions that focus on increasing the quality of parent-child interactions have been developed and tested (McGillion et al., 2017; Ridge et al., 2015; Roberts & Kaiser, 2011, 2012; Suskind et al., 2016). However, existing research suggests that in order to be effective, particularly in the long term, such interventions need to be substantial, delivered at scale and potentially combined with school-level interventions (such as the Nuffield Early Language Intervention (NELI), which is

effective in promoting language skills of children entering formal education in England (West et al., 2021). Evidence also highlights the importance of training staff to deliver interventions effectively (Law, Charlton, Dockrell, et al., 2017), and thorough evaluations of interventions intending to be scaled up are needed (Wake et al., 2011). Effective interventions that have sustained benefits are likely to be unavoidably costly, both financially and in terms of the time it would take to train staff to implement them properly. A large amount of responsibility is also placed on caregivers and the wider family to put interventions and their messages into practice, beyond the study settings.

In order to justify large amounts of spending and intensity of effort in developing and implementing effective interventions, we need to be sure that such efforts have the intended effects on child outcomes beyond improved language ability, such as benefitting educational attainment and wellbeing. Parenting interventions can unintentionally widen social inequalities in language ability, as a result of socioeconomic discrepancies in intervention uptake, compliance and effectiveness (Marulis & Neuman, 2013; Mol & Bus, 2011; White, Adams, & Heywood, 2009), and efforts and resources may be wasted if delivered to those who would not benefit from them (Law, Charlton, Dockrell, et al., 2017). To try and minimise inequalities widening further as a result of interventions, we need to take a step back and revisit socioeconomic inequalities in child language ability, and investigate any subsequent effects early inequalities in language ability may have on later outcomes. Furthermore, to provide insight into whether well-intended parenting interventions will have the desired effects of improving wider functioning, we need to establish whether or not it is really the case that early language ability is related to positive educational and wellbeing outcomes. Research that has investigated the link between language and outcomes such as educational attainment and mental health have focussed on small, or clinical, samples, limiting knowledge and understanding about the extent of such relationships in the general population (Law, Charlton, & Asmussen, 2017). Furthermore, it is important to investigate any such relations in a cohort of children growing up in contemporary society, alongside the increased focus on the importance of language ability. This thesis takes advantage of the opportunity to investigate inequalities in vocabulary ability, and the impacts of early vocabulary on educational attainment and mental health in adolescence in a cohort of children born at the turn of this millennium: the Millennium Cohort Study (see Chapter 2).

In this thesis, I focus on two key outcomes in adolescence, that go on to have important effects in adulthood: educational attainment and mental health. Early language skill is thought to be linked to both outcomes, with previous research suggesting a relationship

between them. Despite research that has investigated the relation between language ability and education, and language ability and mental health (see Chapters 3 and 5), it is important to understand the nature of these relationships in the general population, and how they are influenced by socioeconomic circumstances. It is also important to establish whether there is a further benefit to having language skills at the top of the distribution, or if there is a minimum set of language skills that are sufficient for these outcomes (see Chapter 5). Large amounts of expensive research and national effort has been put into understanding and improving language skills in recent years, with the recent PHE speech, language and communication pathway emphasising the need for universal support for all children, and targeted and/or specialised support for those who need it (Law, Reilly, & Snow, 2013; Public Health England, 2020a). However, sufficient gaps in our understanding of the relation between early language and such outcomes remain, which I seek to address in this thesis.

1.2. The Unidimensionality of Oral Language

Language is a symbolic, arbitrary communication system, shared between specific cultures or groups in society, which relies on the combination of sounds, signs, or written words according to a systematic set of rules, to convey meaning. Language can be described as consisting of four components: 1) phonology, i.e., the contrasting relations between the sounds of a specific language, which convey different meanings, and rules for how these sounds can be combined to produce words; 2) semantics, which refers to the meaning of both single words and combinations of words (i.e., vocabulary); 3) grammar, which can be further split into two components of morphology and syntax (morphology refers to elements of a word that conveys meaning, such as plural endings (‘s’) or past tense markers (‘ed’); syntax refers to the rules that determine how words are combined to produce sentences that make sense); and 4) pragmatics, i.e., the conventions that determine how language is used in social situations, such as starting, regulating and ending conversations (Asmussen et al., 2018; Buckley, 2003). These components of language are interdependent and develop alongside each other rather than in isolation. For example, vocabulary develops in tandem with phonological development, as acquiring new vocabulary simultaneously increases the awareness of sounds (Vihman & Croft, 2007).

Traditionally, these components have been regarded as distinct from one another in language assessment. For example, test batteries that measure language ability have separate subscales, reflecting the different components of language ability, and assessments during

speech and language therapy tend to consist of semantics, grammar and phonology, with a diagnosis of developmental language disorder stemming from difficulty with vocabulary, grammar, discourse or pragmatics (American Psychiatric Association, 2013; Anthony et al., 2014). However, there is evidence to suggest that oral language can be usefully described as a single construct, particularly early in development, with components loading onto a single factor (see for example, Fricke et al., 2017, although this may be an artefact of the tests used to measure language ability, rather than the four components themselves being inseparable). Tests of vocabulary and grammar have been found to form a single factor in children aged 5 to 10 years old (Tomblin & Zhang, 2006), and in a multi-level factor analysis that accounted for classroom level factors, vocabulary and grammar were not distinct constructs in pre-school children (Anthony et al., 2014). Collectively, these studies provide further evidence for the unidimensionality of oral language skill in measurement.

There is some evidence to suggest that vocabulary and grammar may emerge as separate measurable constructs throughout development (Language and Reading Research Consortium, 2015; Tomblin & Zhang, 2006), and phonological and discourse skills (such as understanding narratives or the ability to make inferences) may be distinct from a general language factor made up of vocabulary and grammar (Anthony et al., 2014; Language and Reading Research Consortium, 2015). However, despite a wealth of research supporting the unidimensionality of the construct of language in assessments, one study found that measures of vocabulary and grammar considered separately were the best fit to the data, from preschool through to age 10, although the two constructs were highly correlated (Lonigan & Milburn, 2017). This study diverges from previous studies, which position vocabulary and grammar as emerging constructs that are unidimensional to start with, but become increasingly separable throughout childhood (as children age, measures of vocabulary and grammar are increasingly distinctive). However, the evidence for an underlying general language construct, particularly in the early years of development, indicate that although language acquisition is a complex process, language ability has been said to be best described as a singular construct when assessing language ability (Hulme et al., 2020). Together, this evidence suggests that test scores, for example on measures of vocabulary, likely reflect test scores on measures of grammar, and vice versa, as performance on such measures is likely driven by the same underlying factor, and separate measures of vocabulary and grammar may only be informative when used on older children (Tomblin & Zhang, 2006). Further, interventions that target vocabulary likely also benefit grammatical skills (Lonigan & Milburn, 2017). Thus, although the four components of language are cognitively separable

constructs, they are not necessarily separable from one another in language assessment, and current measures of language ability may not adequately distinguish between these components. Therefore, when measuring language ability, a single construct is acceptable in terms of predictive ability, and in many contexts is as good as a multi-dimensional approach.

1.3. Vocabulary

Vocabulary development refers to the learning of new words, and involves the ability to recognise and say the sounds that make up a word, learning what the word means, and how to represent and generalise the word correctly (for example, the word dog can be used to refer to different breeds of dog, but not other animals that look similar to dogs; Ambridge & Lieven, 2011; Law, Charlton, Dockrell, et al., 2017). Although not the only component of language development, vocabulary skill is thought to underlie oral language and the development of reading comprehension, and children need a good understanding of words in order to communicate effectively, both orally and in written language (Law, Charlton, Dockrell, et al., 2017). Although the ability to successfully combine words (which requires grammatical skill) is also important, vocabulary has been conceptualised as the "tree trunk" when considering language development as a tree, and grammatical ability as the branches (see Law, Charlton, & Asmussen, 2017); the emergence of grammar during a child's second year is highly contingent on vocabulary size (Bates & Goodman, 1997) and a minimum vocabulary size is essential for children to begin to combine words. Once children can say between 50-100 words (usually between the ages of 18 and 20 months), they begin to combine words to make two-word phrases (Bates et al., 2003). As vocabulary size and depth increases, children are able to form more word combinations as their grammatical development begins (Buckley, 2003). Furthermore, vocabulary is important for the development of literacy and numeracy (Lervåg et al., 2018; Moll et al., 2015; Slusser et al., 2019; Snowling & Hulme, 2020).

Receptive language, or comprehension, refers to what children understand, and expressive language refers to the use of language (both verbal and non-verbal) to convey one's thoughts and ideas. Vocabulary development is typically measured by assessing either receptive vocabulary (the number of words a child understands) or expressive vocabulary (the number of words a child can say), for example, using assessment batteries that examine how many words a child understands and/or says at a specific point in time, or through parent-report measures of the number of words a child can understand or understand and

says, such as the MacArthur-Bates Communication Development Inventory (CDI; Alcock, Meints, & Rowland, 2020; Fenson et al., 1994). Receptive vocabulary precedes expressive vocabulary, and it is thought that children typically understand around 80-100 words when they begin to say their first words (Fenson et al., 1994).

This thesis uses measures of receptive and expressive vocabulary as either the main predictor or the outcome in a series of analyses. Historically, the modality of language deficit has been considered to be important when diagnosing language impairment, with a clear distinction between Expressive Language Disorder or Mixed Receptive-Expressive Language Disorder being made in the DSM-IV-TR (American Psychiatric Association, 2000). Although receptive and expressive vocabulary are clearly separable phenomena, they may rely on the same underlying knowledge, to such an extent that only one needs to be measured to gain an understanding of both: research does not support a distinction between separate measures of receptive and expressive language, with findings instead indicating expressive and receptive vocabulary and grammar may be unidimensional in measurement (Anthony et al., 2014; Lonigan & Milburn, 2017; Tomblin & Zhang, 2006). Furthermore, the DSM-5 no longer distinguishes between receptive and expressive dimensions of language for the diagnosis of language impairment, and instead focusses on the persistent difficulties in language across these modalities (American Psychiatric Association, 2013). Therefore, there may be little distinction between measures of receptive and expressive language ability. In addition, receptive vocabulary and expressive vocabulary has been found to be highly correlated with each other, and early receptive vocabulary ability is predictive of later expressive vocabulary (Chiat & Roy, 2008; Conway et al., 2017; Jordan & Coulter, 2017).

1.4. Individual differences in Language Development

Some children begin to understand their first words at around the age of 6 months, with a key milestone in receptive vocabulary development being the ability to recognise their name, followed gradually by the names of common household items and family members (Bergelson & Swingley, 2012; Bortfeld et al., 2005). Expressive vocabulary develops later, with some children producing their first words around their first birthday. On average, 12-month-olds produce less than 10 words. Around 6 months later, most children experience a vocabulary spurt and by their second birthday, they can produce around 300 words (Asmussen et al., 2018; Fenson et al., 1994; Reznick & Goldfield, 1992). Vocabulary

continues to develop, and by the time children are aged around 6 years old, on average they know approximately 10,000 words (Anglin et al., 1993; Diesendruck, 2008).

There are individual differences in the onset of children's first words and the subsequent rate at which children increase their vocabularies, in addition to individual differences in other domains of language (such as grammar, phonology and pragmatics) (Fenson et al., 1994; Kidd et al., 2018; Kidd & Donnelly, 2020). For example, norming data from Wordbank (an open depository of CDI data from across the globe) indicate that 18-month-old children in the 10th percentile of vocabulary produce around 10 words, whilst those in the 90th percentile can produce just under 200 words. By the time children are aged 2 and a half years, there is a gap of around 400 words between these percentiles (Frank et al., 2017). A population-based survey based on children in Surrey, England (SCALES), which includes children with Developmental Language Disorder and a range of developmental conditions known to affect language, such as Autism and Down's Syndrome (Norbury et al., 2016) has also found large individual differences in vocabulary among primary school children in England, in both year 1 and year 3 (Norbury et al., 2017).

Individual differences have been found to be stable once children start school, meaning individuals tend to maintain their position in the language distribution relative to their peers, despite any individual gains in vocabulary, irrespective of biological and social risk factors, sex, non-verbal ability, birth order, and regardless of levels of language skill (Bornstein et al., 2014, 2016b, 2016a; Norbury et al., 2017). Individual differences are thought to be less stable before children start school, and even in the later school years when stability estimates are high, some variance in language is left unaccounted for, suggesting that language skill is modifiable and may be particularly sensitive to intervention in the early childhood years, when stability estimates are much smaller (Bornstein et al., 2014). For some children, the stability of individual differences fluctuates throughout childhood: for example, in a study of language trajectories between the ages of 4 and 11 in an Australian sample, of children who started school with poor language skills, 4% experienced further decreases in ability throughout childhood, and 2% experienced gains in ability, such that they subsequently had language scores that fell in the typical range. However, for most children (the remaining 94% of the sample), the relative position in the distribution regarding language ability was established by the time they started school (McKean, Wraith, et al., 2017). It is important to note that stability of individual differences in language skill does not mean that children will not make personal gains in their own language ability, rather they will maintain their position in the distribution of language skill, relative to their peers, and

individual differences displayed at one point in time will present in a similar way at a later point in time (Bornstein et al., 2016b).

Several factors have been put forward to explain individual differences in language ability. For example, in the early stages of development, girls are thought to acquire language faster than boys (Fenson et al., 1994; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991) and being male was a predictor of poor vocabulary skills at the age of 4 (Reilly et al., 2010). Birth order is also thought to be related to individual differences, with suggestions that parents use more complex language (such as longer sentences) to their first born children compared to later born children (Hoff-Ginsberg, 1998). However, differences in the speed of language acquisition based on birth order only appear early in development (Fenson et al., 1994). There is a heritable component of language ability, which may increase throughout childhood and into adolescence (Chow & Wong, 2021; Hayiou-Thomas et al., 2012). A family history of language impairment, low birthweight and being from a non-English-speaking background has also been linked to individual differences in language ability (Reilly et al., 2010).

Children differ in the language learning environments they experience, with some children being exposed to substantially more input than others (Hart & Risley, 1995a), as well as a differing quality of linguistic environments, both of which are important for young children's language development (Golinkoff et al., 2019; Rowe, 2018). Because language is a reciprocal interaction, individual differences in children's disposition and language ability may affect caregiver input, which in turn influences further individual differences in language development: caregivers are thought to adapt their input based on the language level of their child (Schwab & Lew-williams, 2016), and children who are more engaged and interested in interactions with their parents may in turn receive higher quality input, for example during shared book reading (Malin et al., 2014). The number of rare words a toddler hears is an important aspect of high-quality language input, which depends on the context of conversations. Rare words are those that occur infrequently in natural speech, with more rare words in child directed speech representing a more diverse input. Daily routines may provide limited opportunities for diverse language input. However, shared book reading is thought to increase the amount of new rare words toddlers hear, providing decontextualised opportunities to talk with them: in a simulation study, it was found that by adding 10 children's picture books a month to the household, the total number of rare word tokens heard would increase by 2%, and an additional 50 picture books a month would increase the number of rare word tokens by 9% (Montag et al., 2018). However, this study was merely a

simulation study, and evidence for whether increases in shared book reading *causally* benefits language development remains limited, with a recent randomised controlled trial of shared interactive book reading interventions indicating limited impacts on children's language skills (Noble et al., 2020).

Children also experience different home learning environments, such as the availability of books, educational toys, and exposure to linguistically enriching situations, such as visits to the library, museums, or art galleries (Rodriguez & Tamis-LeMonda, 2011; Roulstone et al., 2010; Tamis-LeMonda & Rodriguez, 2008). The quality of the home learning environment has been linked to language ability. For example, children in the Millennium Cohort Study who experienced a higher frequency of shared book reading with their caregivers at the age of three had higher vocabulary scores at the age of 11 (Law, Rush, King, Westrupp, & Reilly, 2017). There is also evidence to suggest that those with language skills at the bottom of the distribution may benefit most from increased shared book reading early in childhood (Law, Rush, et al., 2017). In a low income sample, a high-quality home learning environment from 15 months to the age of five was found to have positive effects on vocabulary and literacy skills at age 5, with those experiencing a consistently high quality environment benefitting the most in terms of their age 5 language skills, compared to those who had experienced less educationally supportive environments (Rodriguez & Tamis-LeMonda, 2011). Early childhood education experiences are also linked to individual differences, with high quality provision benefitting language development, particularly for those at risk of language difficulties, for example as a result of receiving less parental input at home (Becker, 2011; Schmerse, 2020; Vernon-Feagans & Bratsch-Hines, 2013). However quality is usually inconsistent across different early years settings (Gambaro et al., 2015). Importantly, the language learning environments that children experience differ as a function of family socioeconomic circumstances (Pace et al., 2017) .

1.5. Socioeconomic inequalities in language ability

Socioeconomic status (SES) is a theoretical, multi-faceted construct that encapsulates an individual's access to economic and social resources (such as education, healthcare and information), the position and prestige in social hierarchies resulting from access to such resources that enable some groups to flourish at the cost of others, and the reproduction of such prestige and position (APA, 2007). Differential access to resources creates socioeconomic inequalities in society. There is a lack of consensus about what SES actually

refers to, and this is reflected in the fact that there are many indicators used to determine one's SES (see Chapter 4), which are not necessarily highly correlated; for example, someone may have a high SES in terms of their level of education, but a low SES in terms of their household income. There are also many terms that are used in the literature to refer to the concept, such as socioeconomic position, social class, or social stratification (Galobardes et al., 2006a, 2006b, 2007). Throughout the analyses in this thesis, I adopt the term socioeconomic circumstances (SECs), because children do not necessarily have their own socioeconomic status, rather they experience a host of circumstances that reflect the socioeconomic status of their parents or caregivers, the household and neighbourhood in which they live (for example, educational qualifications, occupational status, household income, wealth and relative neighbourhood deprivation, all of which are considered as indicators of SEC in this thesis), and it is the effects that parental background in early childhood have on language development and later outcomes that I am interested in here.

The size of children's vocabulary varies greatly by the time they start school (Rowe et al., 2012) and a finding commonly reported in the literature is that children with parents of a higher SEC tend to have a larger vocabulary in their early childhood years compared to their lower SEC counterparts (Arriaga et al., 1998; Blanden & Machin, 2010; Clegg & Ginsborg, 2006; Dollaghan et al., 2000; Farkas & Beron, 2004; Hoff, 2003; Huttenlocher et al., 2010; Rowe, 2012; Rowe & Goldin-Meadow, 2009). In an oft-cited study, Hart and Risley reported a 32-million-word gap between lower and higher SEC backgrounds, suggesting that children from lower SEC families had heard fewer words between 0 and 48 months. Further, by the age of 3, children of high SEC parents produced more words than their lower SEC counterparts (Hart & Risley, 1995a). In a sample of 48 children in the US, identified as high or low SEC using a median split of scores on the Hollingshead Index (a composite measure of SEC, derived from the educational attainment and occupational status of both parents/caregivers; Hollingshead, 1975), high SEC children were reported to understand and say more words than their lower SEC counterparts at both 18 and 24 months; for example, at 24 months, high SEC children knew approximately 450 words, whereas low SEC children knew approximately 300 words (Fernald et al., 2013). Findings from this study suggest that socioeconomic differences in vocabulary size emerge by the age of 18 months in the United States (Fernald et al., 2013), and similar findings have been reported in the United Kingdom (McGillion, Pine, Herbert & Matthews, 2017). There is also evidence for socioeconomic disparities in language development before infants produce their first words. For example, high SEC parents were found to use gestures to convey a greater variety of meanings to their

14-month-old infants, compared to lower SEC parents, and these SEC differentials in parent gesture use were reflected in the gesture use of the 14-month-old infants. At school entry, there were SEC related differences in vocabulary, with higher SEC children having larger vocabularies, and this relation was partly explained by patterns of early gesture use (Rowe & Goldin-Meadow, 2009). Despite SEC differences in vocabulary, it is important to note that there is also variability within SEC groups, as well as between SEC groups; not all SEC children have small vocabularies, and not all high SEC children have large vocabularies (Fernald et al., 2013; Hoff, 2013a; Law, Todd, Clark, Mroz, & Carr, 2013). Regardless of this within-group heterogeneity, group differences in vocabulary as a function of SEC are clear.

The language learning environment, including child directed speech, is important in shaping children's language development. Children from high SEC families, on average, hear more words directed to them than children from lower SEC families, and children who are exposed to more words early in development often go on to have larger vocabularies (Hart & Risley, 1995; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Rowe, 2012). Some have argued that using child directed speech measures alone may lead to an underestimation of the amount of speech children from low SEC families hear (Sperry et al., 2019b); indeed, a recent meta-analysis concluded that whilst the quantity of child directed speech differed as a function of SEC, the amount of overall language input (regardless of whether or not it was directed towards the target child) did not differ across SEC groups (Dailey & Bergelson, 2021).

The quality of input appears to be important for language development above and beyond the quantity of input, which also differs based on SEC (Hirsh-Pasek et al., 2015; Pan et al., 2005; Schwab & Lew-williams, 2016). The use of a diverse vocabulary, a high frequency of rare words, decontextualised language and fluent, reciprocal turn taking between the caregiver and child are important markers of quality input (Hirsh-Pasek et al., 2015; Rowe, 2012). Rowe (2012) measured both the quality and quantity of language input when infants were aged 18, 30 and 42 months, and measured vocabulary when they were aged 30, 42 and 50 months. Three domains of quality were measured: diversity (number of different word types produced during interaction), sophistication (number of different rare words), and narrative (explanation, pretend and narrative utterances). More advantaged parents tended to produce a higher quantity of speech during the interaction, and an increased amount of diverse word types and rare words, compared to less advantaged parents. Quality of input was associated with child vocabulary skill a year later, over and above SEC and prior

vocabulary ability (Rowe, 2012). Asking children questions during interactions, and particularly wh-questions (who, what, where, when, why), may also be characteristic of a high quality language learning environment, and are challenging for children compared to simple yes-no questions (Rowe et al., 2017; Rowland et al., 2003). Reciprocal interactions are also important, and SEC-related differences in activity in Broca's area based on turn-taking have been reported: those who had experienced a higher frequency of turn taking with a caregiver, rather than just a greater quantity of input, had more activation when listening to a story, and this was found to explain the relation between parent education and vocabulary skill (Romeo et al., 2018). Differences in the quality of input children experience also exist within SEC groups, as well as between those from high and low SEC backgrounds. For example, Hirsh-Pasek et al, (2015) investigated the quality of the language environment in a low income sample, and found individual differences in the quality of interactions.

Caregiver responsiveness is another important aspect of high-quality input. Responsiveness is a facet of parenting that provides timely and developmentally appropriate reactions, that are contingent on children's behaviour, vocalisations or interests (Bornstein & Tamis-LeMonda, 1989; Eshel et al., 2006). Instead of directing their child's attention, a responsive parent is tuned into and follows the interests of their child (Barnett et al., 2021). Caregiver responsive behaviours predict language development (Hirsh-Pasek et al., 2015; Levickis et al., 2018) and have been identified as a possible indicator to identify toddlers who may benefit from parenting interventions (Barnett et al., 2021; Levickis et al., 2018). Further, responsive behaviours may differ as a function of SEC: from their first year, children from more deprived backgrounds hear less contingent talk (McGillion et al., 2017). Contingent talk occurs when caregivers are tuned into what their baby is interested in or paying attention to, and talks to them about it, and is a communication style known to facilitate word learning (McGillion et al., 2017; Tamis-LeMonda, Bornstein, & Baumwell, 2001; Tomasello & Farrar, 1986).

Difference or deficit?

The abilities of advantaged groups are often taken to be the norm, and deviations from this are considered deficits (Akhtar & Jaswal, 2013; Callanan & Waxman, 2013). However, the language skills of lower SEC children may simply be a difference from that of higher SEC groups, rather than a deficit. An important question that arises from this difference-or-deficit question is how much emphasis should be placed on interventions or treatments, rather than accepting such differences as being of equal importance compared to the norm (Akhtar

& Jaswal, 2013; Callanan & Waxman, 2013). Arguably, a difference should be considered a deficit when it cascades into negative outcomes (Callanan & Waxman, 2013). For example, Hoff has suggested a pragmatic approach whereby differences are considered deficits if they lead to negative consequences, and the causes of the differences can be addressed with interventions (Hoff, 2013). However, children from low SEC backgrounds are thought to have unique verbal strengths and capabilities, for example in terms of their narrative skills, compared to middle class children, yet these skills are not captured by the measures frequently used in research, such as standardised tests of language ability (Heath, 1983; Hoff, 2013; Rogoff et al., 2017). This point has also been made regarding the diagnosis and treatment of Developmental Language Disorder and Autism Spectrum Disorder among different cultures (Norbury & Sparks, 2013): the demographic of speech and language therapists in the UK are largely White, middle class, English-speaking females, and therapy can present discrepancies with cultural norms for some groups seeking intervention. Further, in Western countries, diagnosis of DLD largely depends on performance on standardised tests of English, as compared to norms that are usually derived from samples in which lower SEC groups are underrepresented, meaning such tests may not assess the skills of children from different backgrounds or cultures (Hoff, 2013; Norbury & Sparks, 2013).

Some children from lower SEC backgrounds tend to underperform on standardised tests measuring the skills that are important for success in education (such as phonological awareness, decoding, vocabulary and grammar). Despite the unique linguistic abilities of children from lower SEC backgrounds, this underperformance compared to their more advantaged counterparts places them at an educational disadvantage, and therefore their different language abilities constitute a deficit (Hoff, 2013). Hoff further argues that to embrace the different language skills of lower SEC children as simply being different and equally valid will not help close the achievement gap in schools, and ignores the possibility that all children may have to navigate the same society once they leave school (Hoff, 2013). Rather than the language skills of lower SEC groups being framed as a difference or a deficit, the extent to which a difference constitutes a deficit depends on the extent of negative outcomes that result. It is important to acknowledge this difference-deficit continuum when discussing the differences in language ability across the socioeconomic spectrum. With this information in hand, it will be possible to evaluate whether intervening to support early vocabulary development is likely to benefit children in terms of their educational outcomes and wellbeing. Before setting out the questions for the thesis, it is worth considering the historical debates in the area in more detail.

The difference-deficit debate has been longstanding in the literature, and an early critical debate is that between Bernstein and Labov. Bernstein proposed the existence of two codes of language use in speech, a restricted code and an elaborated code, which are functions of social relationships (Bernstein, 1970, 1971). Restricted codes are context specific and reliant on shared, common knowledge between speakers, removing the need for explicit meanings, and are tied to a specific context or social relationship. Elaborated codes, on the other hand, are context free and provide explicit meaning that is available to all (universalistic meaning). Whilst middle class children are said to have access to both elaborated and restricted codes, working class children only have access to restricted codes. It is not that working class children have inferior vocabulary skills compared to their middle class counterparts, rather they do not have access to the elaborated codes, which are common in the education system, meaning those who do not have access to elaborated codes are placed at an educational disadvantage as a result of a change in code from what they are used to (Bernstein, 1970, 1971). Bernstein emphasised that there is no superior code, but the importance and value of each code is determined by society, and while restricted codes do not present as a deficit, they may place those who only have access to restricted codes (the working class) at an educational disadvantage, since elaborated codes are characteristic of the education system, which is therefore discrepant with the language styles they are familiar with, despite normal language development (Bernstein, 1970).

Bernstein's theory of restricted and elaborated codes has been heavily criticised by Labov (Labov, 1969b), who argued that Bernstein was biased towards the middle class, painting them as superior to the working class, who have been suggested to experience verbal deprivation. Focussing on children from disadvantaged African-American communities in New York, Labov suggested that their language was not in fact deficient, despite the high rates of educational failure in these communities. Instead, Labov argued that the dialect of these children was skilled, complex and imaginative, despite not garnering social esteem (Labov, 1969b). Central to Labov's criticism was that before enforcing the language of middle-class children on lower class children from disadvantaged African-American communities, which seems to provide advantages in terms of education, the functional importance of such language should be separated from the features which are merely stylistic, and not important for success. In doing so, Labov suggested that the language use of the middle class is actually verbose, unclear and imprecise, and working-class children are more efficient users of language, avoiding the use of irrelevant detail in their speech, often suggesting complex arguments. However, because their use of English is non-standard,

which is not typical of formal situations such as education and are largely unknown to the middle class, they are placed at a disadvantage (Labov, 1969b).

Labov also argues that the methods used to assess language and ability of children immediately favour the middle class child and highlighted the fact that the same question posed to children from different backgrounds is likely interpreted differently, and will elicit different answers from the working and middle class (Labov, 1969b). However, since the dialect of working-class children is not the norm in formal situations such as school and the labour market, these children are at a disadvantage; rather than having deficient speech Labov claimed that working class African-American children underperform at school as a result of how schools respond to their speech, and a lack of opportunity for them to demonstrate their skilled language use. For example, teachers may mark lower class children (Labov specifically discusses children from working class African-American communities) as having poor reading skills as a result of a difference in pronunciation of words compared to standard English, despite reading the text and conveying the correct meaning (Labov, 1969a). Further, Labov suggested that all speakers engage in non-standard English, particularly in casual situations, and those who use such speech most in their casual speech are also the ones who are quickest to punish its presence in the classroom (Labov, 1969a). The ignorance towards non-standard English, Labov claimed, is why these children underperform at school.

In addition to the debate between Bernstein and Labov, a sociologist around the same time, Pierre Bourdieu, suggested the existence of a legitimate language, a type of language which lies at the top of the social hierarchy, resulting from social processes. He claimed language is not only a means of communicating, but also a mechanism of power, and the language one uses indicates their position in the social hierarchy (Bourdieu, 2000). He claimed there is a linguistic habitus (referring to a societal tendency to say certain things, competence, and the ability to use such competence in socially appropriate situations) and a linguistic market, which ultimately dictates what is or is not appropriate to say in different situations (Jenkins, 2006). The legitimate language, although not necessarily the type of language that is most common or dominant in society, is what is expected in formal situations such as school and the labour market, and is the norm against which all speakers of language are judged (Bourdieu, 2000). The legitimate language is used to teach and examine children's capabilities throughout the education system, and if, as a result of one's social background, children do not use the legitimate language, they are unable to progress through the education system.

Bourdieu emphasised the role of the education system in reproducing legitimate languages, since social judgements of the language one is using is often translated into educational judgements about one's ability, which enables the legitimate language to be maintained (Jenkins, 2006). Bourdieu argued that it is not necessarily about whether one's language is restricted or verbose. Rather it is about language use that results from the legitimate language of a given society, that is reflective of that culture as a whole, and importantly, valued in the education system (Bourdieu, 2000). The issues raised in these debates have implications for this thesis, and specifically for consideration of the extent to which individual differences in vocabulary predict educational attainment and mental health. For example, if early vocabulary is important for educational attainment, with those from higher SEC backgrounds feeling the most benefit, then it is possible that the dialect they possess is important in the education system.

1.6. Overview of thesis

The overall goals of this thesis are to revisit socioeconomic inequalities in child language ability and to investigate any subsequent effects of such differences on two important positive outcomes in adolescence: mental health and educational attainment. This will provide insight into whether intervening to support early vocabulary development is likely to benefit children in terms of their educational outcomes and wellbeing.

The aim of this chapter was to provide general background information pertaining to language and the presence of individual differences present during its development, providing the necessary context for the work presented in this thesis. I have discussed the separability of 1) language skill as a measurable construct, and 2) expressive and receptive language in measurement, since I focus solely on vocabulary skill, (both receptive and expressive) throughout the analyses in this thesis. The literature reviewed in this chapter allows us to be confident that our exclusive use of vocabulary measures offers a good proxy for broader language ability. While this chapter provides general context for this thesis, each empirical chapter (Chapters Three to Five) reviews chapter-specific literature, identifying important gaps in our knowledge of the relation between language ability and positive outcomes in adolescence, and in our understanding of socioeconomic inequalities in language ability.

This thesis focuses on outcomes in adolescence, since adolescence is seen as an important developmental period, with these adolescent outcomes having lifelong

consequences for adulthood functioning and outcomes (Viner et al., 2015). For example, adolescence is a critical period for susceptibility to internalising mental ill health (experiences of anxiety and depression; McLaughlin & King, 2016), experiences of which predict social exclusion, stigma, poor educational attainment, risky behaviours and poor physical health (Clayborne et al., 2019), and often persist into adulthood (Fergusson et al., 2005). For these reasons, one of the adolescent outcomes I focus on in this thesis is internalising symptoms in adolescence. The second outcome of interest is educational attainment at the end of secondary school, specifically performance in public examinations. Success at this stage of education is seen as the gateway to post-16 success, whether this be further education, an apprenticeship, or entry to the workforce (The Children's Comissioner, 2019; The Sutton Trust, 2019).

When investigating the relation between vocabulary and these two outcomes, I focus on age 5 vocabulary skill. This is because vocabulary ability relative to peers is likely to be established by this point, yet is still more amenable to change than at later ages (Bornstein et al., 2016b). Investigating the role of children's vocabulary skills at school entry will provide insight into whether parenting interventions in the first 5 years of life will provide long-term benefits in adolescence. Although research suggests socioeconomic differences in language ability are evident from as early as 18 months (McGillion et al., 2017), it can be difficult to identify those who are at risk of experiencing persistent difficulties at this early stage of development. This is due to the lower stability of language skills in the pre-school years (Bornstein et al., 2014): many children who are slow to talk resolve their language delays naturally without any intervention, and some children appear to have good language skills at first, and then begin to fall behind their peers at a later stage (Law, Charlton, Dockrell, et al., 2017).

It is hoped that this thesis will offer insight into the value of early language interventions for improving wider functioning. To be effective, particularly in the long term, language interventions need to be ambitious, which is both expensive and time-consuming for researchers and families alike. To justify large amounts of spending and intensity of effort in developing and implementing effective interventions, we need to be sure that such efforts have the intended effects on child outcomes beyond improved language ability, and do not unintentionally widen socioeconomic inequalities. Language interventions may be more effective for some groups than others, and it is likely that these will need to be tailored specifically for different groups. For example, parenting interventions that seek to narrow socioeconomic inequalities in language need to be relevant to lower SEC groups to allow

these individuals to engage with the intervention, particularly beyond the study period. Investigating the effect of vocabulary on outcomes in different socioeconomic groups will enable understanding of the relevance and effectiveness of parenting interventions for different groups. To avoid any unintended negative consequences of parenting interventions, a better understanding of socioeconomic inequalities in language ability is necessary, in addition to an investigation of any cascading effects early inequalities in language ability may have on later outcomes. This forms the basis for the goals of the thesis.

1.7. Structure of thesis

The protocols for each study were pre-registered and accompanying analysis plans and R code can be found on the Open Science Framework (OSF) or GitHub. Two large, longitudinal birth cohort studies are used throughout this thesis: The 1970 British Cohort Study (BCS1970) and The Millennium Cohort Study (MCS2001) (see Chapter 2). This thesis is structured in the following way:

Chapter 2: In this chapter, an overview of birth cohort studies is reported, and the two cohorts used throughout this thesis are described in detail. The sampling methods, number of cohort members at each sweep of each cohort in turn, and for the MCS2001, details on the sample design and weighting of the cohort are covered. This overview of the cohorts is followed by descriptions of the key variables used throughout the analyses in this thesis. Finally, the concept of missing data and multiple imputation are introduced.

Chapter 3: Chapter 3 describes the first empirical study and focusses on the adolescent outcome of internalising mental health. I investigated the relation between age 5 vocabulary and parent and self-reported internalising mental health in adolescence. To date, the majority of research on this relation has utilised parent-reported measures of internalising symptoms, despite low correlations between parent and self-report (Rescorla et al., 2013) and the recognition of child self-report as the recommended measure (The Children's Society, 2019). Comparing the relation based on reporter is important, as a different pattern as a function of reporter would have implications for the value of improving language ability to subsequently improve wellbeing. I also investigate whether the relation has changed over historical time, by comparing the BCS1970 and the MCS2001 cohorts.

Chapter 4: In Chapter 4, I revisit socioeconomic inequalities in vocabulary, with three main aims. These were: 1) to explore the explanatory value of different measures of SEC, including a composite measure, 2) to examine how this relationship changes with developmental time, and 3) to investigate whether this relationship has changed over historical time. MCS2001 data is analysed to explore the first two research questions, and I addressed research question 3 with a cross-cohort comparison of the BCS1970 and MCS2001.

Chapter 5: In the final empirical chapter, I study the role of age 5 vocabulary in predicting educational attainment at the end of secondary school, after adjusting for SEC and caregiver vocabulary factors. I also investigate the moderating role of SEC in the relation between age 5 vocabulary and educational attainment. This will tell us whether all children experience the benefits (if any) of having a strong vocabulary on their educational attainment, or is this benefit felt more strongly by some than others? This will have important implications for the value of intervening early to improve vocabulary skills to benefit educational attainment: will some groups benefit more than others? This chapter analyses data from the MCS2001 cohort only.

Chapter 6: Chapter 6 provides a general discussion. It discusses the key findings of the thesis in relation to one another, identifying three key themes. These are: 1) the impact of vocabulary on positive adolescent outcomes (Chapters 3 and 5); 2) changes in British society over recent decades, as suggested by cross-cohort comparisons (Chapters 3 and 4); and 3) the importance of vocabulary (and language skill more broadly) for social mobility (Chapters 4 and 5). Future directions and reflections of the research process, datasets and measures used are also discussed. The thesis concludes with a summary of each empirical chapter and the overall implications of our findings. Key findings show that whilst early language interventions are well placed to improve educational attainment to an extent, where associations between early language and adolescent mental health exist, they are negligible, making it unlikely that these interventions will have the desired effects when considering mental health in adolescence.

Each empirical chapter is accompanied by extensive sensitivity checks, which can be found in Appendices 1, 2 and 3. This was to assess the robustness of our findings against

different ways of dealing with missing data (i.e., complete case analyses instead of multiple imputation) and different ways of conceptualising key variables.

Chapter 2 : Data

This thesis consists of a series of secondary data analyses of two large, nationally representative British birth cohort studies: 1) The 1970 British Cohort Study (BCS1970), and 2) The Millennium Cohort Study (MCS2001). Birth cohort studies are a type of longitudinal cohort study, which track groups of people who were born within a specific period. Secondary data analysis of birth cohort studies, as demonstrated in this thesis, provide the opportunity to investigate the effect of early life experiences (such as SEC and vocabulary ability) on later life outcomes (such as mental health and educational attainment). Birth cohort studies are rich sources of data containing information about a wide variety of topics for large, nationally representative samples, which allow researchers to adjust for a wide range of potential confounding variables in analyses. There are multiple birth cohorts in the United Kingdom (see https://cls.ucl.ac.uk). The comparison of birth cohorts (cross-cohort comparison) allows us to investigate how relations may have changed over time, against a backdrop of economic and societal changes. Chapters 3 and 4 include a cross-cohort comparison between the BCS1970 and the MCS2001 to establish whether experiences between these two cohorts differ. Chapter 5 does not include a cross-cohort element for several reasons, primarily that the quality of educational attainment data available in the older BCS1970 cohort does not afford a detailed analysis in the way that the MCS2001 data does. This will be expanded on in Chapter 5.

The remainder of this chapter introduces the two birth cohort studies used, detailing the sampling methods, number of cohort members at each sweep of each cohort in turn, and for the MCS2001, details on the sample design and weighting of the cohort. This overview of the cohorts is followed by descriptions of the key variables used throughout the analyses in this thesis. Finally, missing data and multiple imputation are introduced.

2.1. 1970 British Cohort Study (BCS1970)

The BCS1970 began as the British Births Survey in 1970. To date, data has been collected at the birth of the cohort member, throughout childhood, adolescence, and into adulthood, with the most recent follow up taking place when cohort members were aged 51. This thesis focuses on data collected during the birth, age 5, 10, 16 and 34 follow up sweeps.

2.1.1. Sample

The BCS1970 sample includes all babies born with a gestation period >24 weeks, from 00:01 on 5th April 1970 to 24:00 on 11th April 1970 in England, Scotland, Wales, and Northern Ireland (this was not restricted to babies born in the NHS; Institute of Child Health, 1970). Data from 16,568 babies from England, Scotland, and Wales and 628 babies from Northern Ireland were collected during the birth sweep, resulting in a sample of 17,196 babies (16,815 singletons, 189 pairs of twins and 1 set of triplets). Cohort members from Northern Ireland were dropped after the initial birth sweep, and the sample was supplemented with children who were born abroad in the eligible week and had subsequently moved to Great Britain; there were an additional 79 new cohort members at age 5, 294 at age 10 and 65 at age 16 (CLS website: https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study/).

Data was collected during the age 5 follow up with the aim of reviewing and evaluating cohort members' experiences of pre-school health, care and education services in Britain, with a focus on physical and educational development (Institute of Child Health, 1975). Cohort members who were still alive and in Britain were tracked with NHS numbers, and health visitors invited families to participate. 16,284 children were identified as eligible for the 5 year follow up, and of this target sample, 13,135 were successfully tracked and interviewed around their 5th birthday (but not before). However, this procedure hinged on the data collected in the British Births Survey (i.e., the birth sweep) and therefore children who had entered Britain after the birth sweep could not be tracked, and are underrepresented in the age 5 sample (Institute of Child Health, 1975).

The age 10 follow up sweep had a focus on the physical and social development of cohort members, with the aim of investigating cohort members' health, care and education, and their social and family environments, in mid-childhood (Butler et al., 1981). The tracking of cohort members at age 10 was mainly performed via school registrations, and each school was asked for the name, address and birthday of each child born in the study recruitment week (5th-11th April 1970), meaning those who had since moved to the UK were also invited to participate (Butler et al., 1981). Data were collected from 14, 875 cohort members (out of a possible 16500) between 1980 and 1981, when they were aged 10-11.

The age 16 follow up was initially known as Youthscan and had the aim of investigating cohort member's health, care and education, and their social and family environments, in adolescence, given the critical developmental importance of this stage. In addition to the developmental changes faced by adolescents, for example in terms of

relationships, education and employment, BCS1970 cohort members went through adolescence during a period of economic change, with the rise of technology reducing the need to employ young people, leaving either government schemes or unemployment as the alternatives to work for those leaving education (Goodman & Butler, 1986). Data collection was originally planned to take place before cohort members turned 16, in ample time before they reached the statutory school leaving age. However, due to a teachers' strike in 1986, this was delayed and subsequently resulted in many questionnaires and tests being changed to self-completion. Similar tracking methods to those used in the 10 year follow up were employed, using Local Educational Authority and Regional Council (Scotland) secondary school registers. Other methods were also employed however, due to the number of children who had already left school before the follow up data collection began, such as through Family Practitioner Committees and Scottish Health Boards. 16,500 cohort members were found to be eligible for the age 16 follow up. Of these, 11,622 were productive.

The final relevant sweep for this thesis is the age 34 follow up. Data from this sweep was used in Chapter 3, in the form of a supplementary analysis. The age 34 sweep took place between February 2004 and June 2005, when cohort members were aged 34-35. By the time of the age 34 follow up, tracking of cohort members was achieved through address information obtained from a range of repositories, such as phone number databases, postcode databases, electoral registers, NHS records, and through sending an annual birthday card to cohort members (Dodgeon et al., 2006). A total of 13,107 cohort members took part in this follow up.

2.1.2. Survey Content

In the British Births Survey (now known as the birth sweep of the BCS1970), data was collected about the birth of the cohort member and social circumstances of the cohort member's family by means of questionnaires completed by the midwife at the birth, in addition to information taken from clinical records (Elliott & Shepherd, 2006). Information collected included that concerning smoking during pregnancy, the pregnancy and delivery, sex and physical measurements of the baby, antenatal care, parental occupational status, marital status and childcare (see Elliott and Shepherd, 2006 for a comprehensive list).

During the age 5 follow up wave of data collection, data were collected by means of maternal self-completion questionnaires (to reduce interviewer bias), an interview conducted by health visitors in the cohort member's home, a test booklet for cohort members (overseen by health visitors) and information from health records. These were collected over a 6-month

period in 1975 (Institute of Child Health, 1975). Information about the family's social background (such as parental educational attainment), maternal mental health, child's behaviour, cognitive ability and vocabulary ability were collected during this sweep (see Elliott and Shepherd, 2006, for a comprehensive list).

Data collection during the age 10 follow up included self-completion questionnaires (mother, cohort member, teacher), interviews (parents), a medical examination form (completed by community medical officer and the school nurse and including information from medical records) and educational assessments (Butler et al., 1981). Information about family socioeconomic background (such as parental education, occupation and accommodation types), child behaviour, academic ability, attitudes towards school, self-esteem, vocabulary and reading, as well as information about the use of health services, hospital admissions, disability and maternal health were collected during this sweep (see Elliott and Shepherd, 2006, for a comprehensive list).

In the age 16 follow up, data collection methods were comprised of 18 separate instruments, including self-completion questionnaires (cohort members, parents, and teachers), interviews (parents), assessments, diaries (such as dietary, leisure and activity diaries) and a medical examination form (Goodman & Butler, 1986). Information about cohort member and family health, chronic illnesses, exercise, self-esteem, diet, social experiences, family socioeconomic background, household composition, school performance, vocabulary ability, occupational interests and mental health were collected (see Elliott and Shepherd, 2006, for a comprehensive list).

In the adulthood sweeps of the BCS1970, data collection was obtained predominantly from the cohort members themselves, and in the age 34 sweep, data collection methods were mainly interview and self-completion questionnaires and assessments. The main interview comprised of a Computer Assisted Personal Interview (CAPI) and a Computer Assisted self-completion Interview (CASI), and lasted for approximately 50 minutes (Dodgeon et al, 2006). The CAPI covered topics such as household composition, relationships, education, employment, income, diet, exercise, and height and weight, whilst the CASI covered topics such as voting behaviour, attitudes, alcohol consumption, mental health and wellbeing and experiences of crime. Finally, the adult assessment was used to obtain data on literacy and numeracy skills and symptoms associated with dyslexia, and lasted for around 40 minutes (Dodgeon et al, 2006).

2.2. The Millennium Cohort Study (MCS2001)

The second longitudinal birth cohort used in this thesis is the Millennium Cohort Study (MCS2001), which follows 19,244 young people born across England, Scotland, Wales and Northern Ireland in 2000-02 (Connelly & Platt, 2014). The cohort began as a multidisciplinary study at the turn of the millennium, after a gap of 30 years since the last national birth cohort (the BCS1970). It reflected the interest in the early years, child inequalities and wellbeing of the New Labour government, elected in 1997 (Joshi & Fitzsimons, 2016). The original aims of the MCS2001 included comparison with the older birth cohorts, such as the BCS1970, to investigate how changes in social contexts may have influenced outcomes, the recruitment of cohort members from across the UK, with big enough samples for analyses within each UK country, and providing the opportunity to investigate the effects of early family circumstances, by collecting information from both parents, on child development and outcomes, initially throughout childhood and into adolescence, and later through adulthood (Centre for Longitudinal Studies, 2020; Joshi & Fitzsimons, 2016). The MCS2001 has rich information about family socioeconomic circumstances, such as parental education, occupational status, income, and wealth, as well as neighbourhood deprivation statistics. The MCS2001 also measured vocabulary from early childhood through to adolescence, as well as mental health, and later, cohort member educational attainment. The MCS2001 is therefore a valuable data source for the aims of this thesis, which include investigating the effects of vocabulary on outcomes, and how this has changed over recent decades.

2.2.1. Study design

For this cohort, a stratified clustered sample design was used, which specifically over-recruited subgroups of the population (ethnic minorities, those living in disadvantaged areas, and the smaller UK countries). Electoral wards were used to stratify the population. In England, there were three strata – 1) an ethnic minority stratum (children living in electoral wards where at least 30% of the population were either Black or Asian as reported in the 1991 census); 2) a disadvantaged stratum (those who were not in the ethnic minority stratum but who were in the poorest 25% of electoral wards in terms of the Child Poverty Index — the proportion of children younger than 16 years of age in any given electoral ward living in families that were receiving at least one of income support, jobseekers allowance, family credit or disability working allowance in 1998); and 3) an advantaged stratum (those who

were living in electoral wards other than those who were not in the ethnic minority or disadvantaged stratums). The ethnic minority stratum was not included in the stratification of Wales, Scotland and Northern Ireland, due to the small proportions of ethnic minority groups in these countries. Thus, stratification for the devolved nations was into either the disadvantaged or advantaged stratum (Plewis, 2007b). The sample was geographically clustered due to the use of electoral wards for stratification.

A total of 398 wards were selected across the UK: 200 of these were in England, 73 were in Wales, 62 were in Scotland and 63 were in Northern Ireland. Once the wards were selected, the sample was selected within each stratum in each country separately. In England and Scotland, ordering by region and subsequently by ward size within each region (in descending order) was applied to the populations; in Wales and Northern Ireland, this was done solely based on size of the wards in descending order. Systematic sampling using a sampling interval was used to select which wards would be included in the sample. The sampling interval was determined by the ratio of the number of wards in the populations for each country to the number of wards required in the samples for each country. Sample selection had a random initial point, and a fixed periodic interval (the sampling interval) at which subsequent wards were selected (Plewis, 2007b). The sample size of cohort members was not fixed in advance, and once wards had been selected, a list of all children turning 9 months within the MCS2001 eligible period was provided in the Analytical Services Directorate Information Centre's Child Benefit register, from what was then known as the Department of Social Security (now the Department for Work and Pensions). Eligible families were contacted by the Department of Social Security and were recruited on an opt-out basis (Centre for Longitudinal Studies, 2020; Plewis, 2007b). Families who had recently experienced child bereavement or whose children had been taken into care, families who were under benefit fraud investigation, or who had already taken part in the DWP Families and Children Survey were not contacted. Health visitors were enlisted to find families who may have been missed by Child Benefit records (Centre for Longitudinal Studies, 2020).

2.2.2. Sample

The MCS2001 sample includes all children born 1st September 2000-31st August 2001 (in England and Wales) and between 24th November 2000 and 11th January 2002 (for Scotland and Northern Ireland), alive and living in the UK when aged 9 months, and eligible to receive Child Benefits, who remain living in the UK at the time of sampling for each

subsequent follow up sweep (Plewis, 2007b). Government child benefit records were used to identify eligible children. This is a benefit with wide eligibility. However, families with temporary or uncertain residence status, such as asylum seekers, are ineligible (Connelly & Platt, 2014; Plewis, 2007b). Children who had died, emigrated from the UK or who were not established as a UK resident before they were aged 9 months were excluded from the sample population.

The first survey of the MCS2001, conducted in 2001-02, had a total sample of 18,552 families, with 18,818 cohort members (246 twins, 10 triplets and 6 families with two singletons who were eligible to participate), 75% of who were aged 9 months when interviewed — 2.9% were aged 8 months; 19% were 10 months and 2.6% were aged 11-12 months (Plewis, 2007b). This was 82% of all those eligible to participate. Response rates were higher among the advantaged stratum compared to the disadvantaged stratum (and ethnic stratum in England) for all of the countries (Plewis, 2007b).

In the second sweep, which was conducted in 2003-04, 1,389 New Families, who were eligible and living in sample wards but were not included in the first survey, were contacted in England; of these, 692 contributed. Prior to the second sweep fieldwork, interviewers contacted cohort families from the first sweep. Each parent was sent a letter to remind them about the survey and to inform them that an interviewer would be visiting in the following days. When interviewers visited the address, they asked if the cohort member still lived at the address. In cases where the cohort family were no longer residing at the last known address, interviewers attempted to find out the new address. In cases where these attempts were successful and the new address was local, the interviewer carried out the same steps as for the first address. If the address was not local, the cohort family was assigned a new interviewer. Full details of this procedure can be found in the Technical Report for the second sweep (Chaplin-Gray et al., 2010). The achieved sample at the second sweep, when cohort members were aged ~3 years old, was 15,590 families (78%), with 15,808 cohort members.

At each subsequent follow up sweep, families received a letter and leaflet providing information about the upcoming survey, prior to interviewers contacting them. Interviewers were required to make first contact with families either over the phone or by visiting the last known address. These details were provided on Sample Information Sheets by CLS (Burston et al., 2017; Centre for Longitudinal Studies & Ipsos, 2019; Chaplin Gray et al., 2009; Gallop et al., 2013). Families who were considered hard to reach, for example those who did not respond to the previous sweep, were prioritised for face-to-face contact. Interviewers were

required to make a minimum of 5 phone calls to each provided contact number and a minimum of 8 face to face visit attempts, before recording the family as being unable to contact. Where cohort families were no longer living at the last known address, interviewers attempted to trace the new address in a multitude of ways. These included: 1) speaking to the new people living in the last known address and neighbours to see if they had any information about family or friends who would know where the cohort family had moved to; 2) contacting nominated stable contacts (these were usually a close relative of one of the parents); and 3) contacting the school the cohort member was attending in the previous sweep. In cases where these attempts were successful, but the contact was unwilling to provide the contact details of the cohort family, they were given a letter containing information about the upcoming sweep and were asked to pass this on to the cohort family. If the new address was local to interviewers, they proceeded with making contact, otherwise the family was allocated a new interviewer (Burston et al., 2017; Centre for Longitudinal Studies & Ipsos, 2019; Chaplin Gray et al., 2009; Gallop et al., 2013).

15,246 families (79.2%) took part in the third sweep, which took place in 2005-06 when cohort members were roughly aged 5 years old, resulting in 15,459 cohort members. Sweep 4 took place in 2007-08, when cohort members were aged around 7 years old. 13,857 families (72%) took part, yielding 14,043 cohort members. When cohort children were aged around 11 years old, the fifth survey sweep took place, with 13,287 families (69.1%) and 13,469 cohort members. For sweeps 3-5, 19,244 families were eligible to take part. However, at each stage, some cases were not issued onto the field, due to ineligibility resulting from death, emigration, permanent refusal or sensitive family situations (n = 718 at sweep 3; n = 2,213 at sweep 4 and n = 2,581 at sweep 5) (Centre for Longitudinal Studies, 2020). When cohort members were aged 14, they took part in the sixth sweep of the survey. There were 19,243 families eligible to take part in this sweep, however 3828 families were not issued due to death, emigration, permanent refusal to take part or sensitive family situations. 15,415 cases were issued and of these, 11,726 families (60.9%) took part, resulting in 11,872 cohort members (Fitzsimons et al., 2017). The most recent sweep of data collection (sweep 7) took place when cohort members were aged 17 years old in 2018-19. Like the 6th sweep, 19,243 families were eligible to take part; of these, 4,747 families were not issued into the field, resulting in a total issued sample of 14,496 families. Of these, 10,625 (73.6%) families responded, resulting in 10,757 cohort members.

2.2.3. Survey Content

Throughout the MCS2001 sweeps, both parents were interviewed as either the main respondent or the partner respondent. These interviews were either "Main Interview", "Partner Interview", or "Partner Proxy Interview". Household members were allocated to either the main or partner respondent by the CAPI, based on both their relationship with the cohort member and their relationship with the remaining household members. Natural, step, foster and adoptive parents, and partners (including same sex partners) of parents were eligible. In cases where no parents were present in the household, the cohort member's main career and their partner were interviewed. In the first sweep of data collection, if the natural mother was present in the household then she was prioritised for interview; in a minority of cases, the natural mother completed the partner interview as a result of language barriers (Hansen, 2012). Where possible, the main respondent at sweep 1 was also the main respondent at sweep 2. However, in cases where the original main respondent was no longer in the household, but a biological parent was present, this parent would be the main respondent. If no biological parents were present, the main career was selected for the main respondent interview. In subsequent sweeps, the main respondent would be selected as the natural mother, the natural father if she was no longer present, the main informant at the preceding sweep, or the cohort member's main carer would be selected as the main respondent, in that order (Hansen, 2012). In the first sweep, 18,524 out of the total 18,815 families had the natural mother as the main respondent. 28 main respondents were the natural fathers (18 were single fathers). 99.6% of partner respondents in the first sweep were the natural fathers (Hansen, 2012).

Data collection methods throughout the MCS2001 included interviews, self-completion questionnaires, cognitive assessments, and interviewer observations. Computer assisted personal interviewing (CAPI) and computer assisted self-interviewing (CASI) instruments were used at each sweep. In the first sweep, when cohort members were around 9 months old, the main and partner respondents were interviewed and completed questionnaires. Information about the pregnancy, labour and delivery of the cohort member, the cohort member's health, development and behaviour, childcare, parental health and mental health, previous relationships and pregnancies, and SEC was collected. From the second sweep, and in each subsequent sweep, data about cohort member's height and weight were collected, and cohort members completed cognitive assessments. Further, information about any changes in SEC was collected at each sweep. When cohort members were aged 3

and 5, information about their temperament, health and behaviour was obtained in the parent interviews, in addition to any information about previous pregnancies and the home environment (Centre for Longitudinal Studies, 2020). Cohort member's vocabulary was also measured. A full overview of topics is provided in the User Guide (Centre for Longitudinal Studies, 2020).

In the age 11 sweep, parents were asked questions pertaining to their wealth (see 2.3.2.4), their relationship with the cohort member and parenting activities, and cohort members were asked about activities they did outside of school, the transition to secondary school and school experiences. A measure of vocabulary was also collected. In the sixth sweep of data collection, when cohort members were aged 14, parents completed a cognitive assessment, measuring their vocabulary, and data about cohort member's mental health, both in the form of self-report and parent-report were collected. Cohort members were also asked about their relationships, risky behaviours they engaged in, and other experiences such as bullying. Cohort member vocabulary was also measured (Fitzsimons et al., 2017). In the most recent sweep of data collection, when cohort members were aged 17, they provided self-reported qualifications data about qualifications they had achieved to date (for example, GCSEs, and qualifications they were currently studying for, such as A-levels). Cohort members were also asked about their relationships with their family, peer relationships, risky behaviours and any employment experiences (Fitzsimons, Haselden, et al., 2020). A full overview of topics of interest at each sweep of data collection can be found in the respective sweep's user guides.

2.2.4. Weighting in the MCS2001

Due to the disproportionate stratified clustered design of the MCS2001, whereby those from areas with high proportions of ethnic minorities, disadvantage and those from the smaller UK countries are overrepresented, the MCS2001 sample is not self-weighting, and individuals from the different sampled wards have unequal probabilities of being selected. For example, children born into families who are in areas with high rates of disadvantage are more likely to be selected into the sample than those living in advantaged areas (Plewis, 2007b). Furthermore, the clustered design within neighbourhoods means that observations are not independent. Sample design weights are needed to adjust for this stratified clustered design and to reduce sampling error, when obtaining estimates such as the mean, variance

and proportions, otherwise standard errors could be underestimated and subsequent significance tests invalid (Centre for Longitudinal Studies, 2020; Plewis, 2007b).

All longitudinal studies are susceptible to attrition between sweeps of data collection (detailed in section 2.4). Adjustment with non-response weights can account for any biases caused by unit non-response. In the MCS2001 cohort, the predictors of non-response between each sweep were used to produce these non-response weights, which can be combined with the sample weight to provide an overall weight for each sweep of data collection (Plewis, 2007b). At each respective sweep, the non-response weight is the inverse probability of response, predicted by a logistic regression model predicting response, using predictors from previous sweeps to predict response at the current sweep (Centre for Longitudinal Studies, 2020).

2.3. Variables of interest

This section provides details of the vocabulary and SEC measures used throughout this thesis. Vocabulary measures are considered as outcome variables in Chapter 4, and different measures of SEC are considered as predictors of vocabulary throughout childhood and into adolescence. Age 5 vocabulary is used in Chapters 3 and 5 as the main predictor of adolescent outcomes. Details about outcome variables (adolescent internalising symptoms and educational attainment at the end of secondary school) and potential confounding variables used in each analysis can be found in the respective empirical chapter. Because measures of SEC and vocabulary are used consistently throughout the thesis, they deserve detailed description, thus they are covered in detail in this section.

2.3.1. Vocabulary

This thesis uses a series of verbal cognitive assessments as an indication of vocabulary ability, that were administered to cohort members throughout childhood and into adolescence. In the BCS1970, tests at ages 5, 10 and 16 are utilised, and in the MCS2001, tests at ages 3, 5, 11 and 14 are utilised. Although not all tests directly measure vocabulary (for example, the Verbal Similarities test (see 2.3.1.2) is primarily a measure of verbal reasoning and knowledge), they are a measure of vocabulary knowledge to an extent, and so will be referred to generically as vocabulary skills throughout this thesis for simplicity. It must be noted that whilst all tests considered in this thesis measure vocabulary knowledge to

an extent, no test measures one skill in isolation, and each relies on wider skills for successful completion. Further, although language is usefully measured as a single construct, and vocabulary is thought to be a good proxy for wider language ability, unless the same test is used across cohorts and at each age, measurement is unlikely to be invariant, and it is possible that each test measures a primary skill, within the domain of vocabulary. This needs to be considered when interpreting the results of this thesis, as differences in the focus of each test may explain some of the results presented.

The following section gives an overview of the test used at each age in both cohorts, including the administration and scoring of the test, what each test specifically measures, and the skills required for each test. Details of the use of each vocabulary measure in specific analyses can be found in relevant chapters.

2.3.1.1. Vocabulary in the BCS1970

English Picture Vocabulary Test. Age 5. (EPVT; Brimer & Dunn, 1962)

The English Picture Vocabulary Test (EPVT) is a measure of receptive vocabulary. This test is a UK version of the Peabody Picture Vocabulary Test (Dunn et al., 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. The items became increasingly difficult as the test progressed, and the test stopped when the child made five errors in any sequence 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 et al., 1984). The BCS1970 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 2.1 (taken from the Age 5 Test Booklet, 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.

Although the EPVT measures language comprehension within the verbal domain (and specifically, receptive vocabulary), picture recognition and the ability to recall words from one's long-term memory are also important for completion of this test (Moulton et al., 2020).

Figure 2.1. Target words in the English Picture Vocabulary Test

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

Age 10: BAS word similarities (Elliott, Murray, & Pearson, 1979)

The BAS word similarities measures verbal knowledge and verbal reasoning. 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. For example, the first item was *orange*, *strawberry* and *banana*. Cohort members were asked "What would go with these" (for example, an apple) and "Why do they go together" (because they are all fruit). This first item was a teaching item, where cohort members were prompted with hints to

answer correctly. However, no input was given on subsequent items. Other items included *red, blue*, *brown,* and *daisy, rose, dandelion* (see https://cls.ucl.ac.uk/wp-content/uploads/2018/06/British-ability-scales-annotated.pdf for all items). In order to receive a point, cohort members had to correctly answer both parts of the question (Moulton et al., 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, the test was terminated. Items became progressively harder throughout the test, for example, the last item on the test was *democracy, justice, equality*. 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).

Although this test measures a primary skill within the verbal domain (specifically, verbal reasoning), expressive language skills, vocabulary and general knowledge, and an ability to differentiate between essential and superficial features of an object, are all crucial for completion of this test (Moulton et al., 2020). For example, since cohort members are required to state how the words in an item are related to one another, they must be able to recognise which features are essential to make something qualify as part of that category. Vocabulary and general knowledge are necessary for cohort members to be able to recognise and name other words consistent with the words in the item. Abstract and logical thinking, and the ability to recognise patterns and trends between the items is also important (Moulton et al., 2020).

Age 16: Vocabulary test (Closs, 1986)

The APU Vocabulary Test is a measure of vocabulary and measures one's word knowledge and vocabulary depth. 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 had the same meaning as the target word (Moulton et al., 2020; Parsons, 2014). Items became progressively harder as the test progressed (see Figure 2.2 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 2.2. Items in the Vocabulary Test (BCS1970 Age 16)

	(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). The full list of items can be found in the age 16 guide to BCS1970 data (Goodman & Butler, 1986)

2.3.1.2. Vocabulary in the MCS2001

British Ability Scales II (BAS II): Naming vocabulary. Ages 3 & 5 (Elliott, Smith, & McCulloch, 1996)

The BAS II Naming Vocabulary measures spoken expressive vocabulary. 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 resulted in a different, easier set of items being administered (further details for each age is provided below). The first two items were teaching items, whereby the interviewer confirmed if cohort members were correct, and gave the correct answer in cases where cohort members had answered incorrectly. On subsequent items, interviewers were only permitted to give encouragement to answer, rather than any specific guidance on the items.

Although this test is primarily a measure of vocabulary, children also need to be able to recognise pictures and retrieve the names of items from their long-term memory, rather than just knowing the words or understanding their meanings (Moulton et al., 2020). To progress through this test, children must say their answers out loud, and so a reluctance to speak may hinder progress through this test, despite strong vocabulary knowledge, which would not be reflected in their score on this test.

MCS2001 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, I 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 et al., 2020)

Age 5: The assessment started from picture 12. Progression through the test depended on the answers given by the cohort member, and the test ended when the child made five sequential errors. However, if early in the test, the child 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 (Gray, Gatenby, & Simmonds, 2009; Moulton, et al., 2020). Therefore, MCS2001 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)

The BAS II verbal similarities measures 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 (see Figure 2.3) (Moulton et al., 2020). Cohort members started the test at item 16, the start point for children aged 11, and completed up to item 28 (the decision point, based on their 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 et al., 2020). Again, progression through this test depends on performance, and poor

performance may result in a different, easier set of items being administered. I 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.

Although this test measures a primary skill within the verbal domain (specifically, verbal reasoning), expressive language skills, vocabulary and general knowledge, and an ability to differentiate between essential and superficial features of an object, are all crucial for completion of this test. Abstract and logical thinking, and the ability to recognise patterns and trends between the items is also important (Moulton et al., 2020).

Figure 2.3. Example items from 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	_____

Word Activity Task. Age 14 (Closs, 1986)

This test assesses the understanding of meanings of words and word knowledge (depth of vocabulary). Items were a subset of the items from the Applied Psychology Unit (APU) Vocabulary Test (Closs, 1986), which were used in the BCS1970 age 16 vocabulary measure. Cohort members were given a list of 20 target words, each presented alongside 5 other words. Cohort members had to choose the word that 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 et al., 2020). See Figure 2.4 for examples of items.

Figure 2.4. Example items from the 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	spurious	spasmodic	superfluous	pathetic	spurious
PROSAIC	commonplace	flowery	laudable	poetical	spacious
ASCETIC	artistic	dissolute	austere	antipathetic	charlatan
PUSILLANIMOUS	loud	living	timid	averse	correct

2.3.2. Family Socioeconomic Circumstances

I use multiple indicators of SEC in the analyses in this thesis, due to the multi-dimensional nature of the construct. These were parent educational attainment, occupational status, income, wealth, and relative neighbourhood deprivation. Measures of parent education and occupational status were taken from both cohorts, while income, wealth and relative neighbourhood deprivation were taken just from the MCS2001, as they were not available for the childhood sweeps of the BCS1970. The overall measures are covered in detail below, and details relevant to specific chapters can be found in the respective chapter.

2.3.2.1. Parent Education

Parent educational attainment is often reported in terms of the highest level of education a person has completed, or the highest qualification they have achieved. Educational attainment categories can be, to an extent, hierarchically categorised (Schneider, 2011). Measures of the highest level of education achieved do not necessarily indicate the level of performance in such qualifications. The highest level of education now covers both academic and vocational qualifications, given the increased societal value of and the number of people taking vocational qualifications. Educational attainment could also be classified in terms of years spent in education, assuming more years in education equates to higher levels of attainment, although this is not necessarily the case, given the fact that people may take a

break in education before returning to study, or may resit years of study as a result of failure to progress. Educational qualifications, on the other hand, are officially recognised certifications to demonstrate that a particular level of proficiency has been reached in a given educational domain, and are usually achieved through official examinations (Schneider, 2011). Highest educational qualification achieved is therefore seen as a more informative and valid indicator of educational attainment.

BCS1970

Information about parent educational qualifications was first collected when cohort members were aged 5; in the birth sweep, the age at which parents completed education was collected. Mothers were asked to report their own and the cohort member's father's highest educational qualification:

1. No qualifications
2. Vocational qualifications (e.g., shorthand, typing, trade apprenticeships, state enrolled nurse, and hairdressing diplomas)
3. O levels or equivalent (e.g., Certificate of Secondary Education, City and Guilds Intermediate Technical Certificate, and a City and Guilds Final Craft Certificate)
4. A-levels or equivalent (e.g., High School Certificates, Higher Grade of Scottish Leaving Certificates, City and Guilds Final Technical Certificate or Scottish Certificate of Education)
5. State registered nurse or Registered Nurse (Scotland)
6. Certificate of education or teaching qualification (Scotland)
7. Degree + (e.g., BSc, BA, PhD, Higher National Diploma/certificate, City and Guilds Full Technical Certificate, membership of a professional institution)

In Chapter 4, these categories were collapsed into a 4-category scale (see below) for harmonisation with the MCS2001 (see Chapter 4 for further details).

4 category scale:

1. No/low level qualifications
2. O-levels/ GCSE grades A*-C
3. A-Levels/ earning post-16 qualifications
4. University level qualifications

For this harmonised variable, the full scale was collapsed in the following way: no qualifications and vocational qualifications were collapsed into the 'no/low levels qualifications' category, O levels or equivalents were collapsed into the 'O-levels/GCSE grades A*-C' category, A levels or equivalent were collapsed into the 'A-levels/earning post-

16 qualifications' category, and State registered nurse, certificate of education and degree+ were collapsed into the "University level qualifications" category. Vocational qualifications were included in the no/low level qualifications, as they were not equivalent to O-level or A-level qualifications, and in the cross-cohort comparison analyses, we were interested in academic levels of qualifications and their effects on vocabulary; analysis of MCS2001 data in the cross-cohort comparisons considered only academic qualifications (see below for more information). In Chapter 3, these categories are considered on the full scale detailed above. The BCS1970 data was not analysed in Chapter 5.

MCS2001

Information about parent educational qualifications was first collected in the 9 months sweep of data collection. At each subsequent sweep, parents were asked if they had achieved any new qualifications since the time of the last interview, and data were updated accordingly. The MCS2001 asks about both academic and vocational qualifications achieved (see below). These are then collapsed into a combined academic and vocational scale, ranging from NVQ 1-5, with NVQ level 5 being the highest level of qualifications.

Academic qualifications:

1. Higher degree
2. First degree
3. Diplomas in higher education
4. A/ AS/ S levels
5. O level/ GCSE grades A-C
6. GCSE grades D-G
7. None of these qualifications
8. Other academic qualifications (including overseas)

Vocational qualifications:

1. Professional qualifications at degree level
2. Nursing/medical qualifications
3. National Vocational Qualification (NVQ), Scottish Vocational Qualification (SVQ) or General Scottish Vocational Qualification (GSVQ) level 3
4. Trade apprenticeships
5. NVQ/ SVQ/ GSVQ level 2
6. NVQ/ SVQ/ GSVQ level 1
7. None of these
8. Other vocational qualifications (including overseas)

National vocational qualifications, or Scottish vocational qualifications are taken by people who are employed full time, or who are in school or college with a work placement or part time job (UCAS, 2021), and over 1,000 subjects are offered. Since the availability and prevalence of vocational qualifications has increased, the MCS2001 vocational qualification data covers a much wider range of possibilities than the BCS1970 academic alternatives, and for this reason, when comparing the MCS2001 and the BCS1970, academic qualifications of MCS2001 parents were considered (Chapters 3 and 4).

In Chapter 4, when data were compared to BCS1970 data, the same 4 category scale was used as BCS1970 data, and in Chapter 3, the categories are considered on the full academic scale detailed above.

4 category academic qualifications:

1. No/low level qualifications
2. O-levels/ GCSE grades A*-C
3. A-Levels/ earning post-16 qualifications
4. University level qualifications

For this harmonised variable, the academic qualifications from the MCS2001 were collapsed in the following way: other academic qualifications, none of these qualifications and GCSE grades D-G were collapsed into the 'No/low level qualifications' category; O level/GCSE grades A-C were collapsed into the 'O-levels/GCSE grades A*-C' category; A/AS/S levels and diplomas in higher education were collapsed into the 'A-levels/ earning post-16 qualifications' category; and first degree and higher degrees were collapsed into the 'university level qualifications' category. The categories "other academic qualifications" and "none of these qualifications" were collapsed into the no/low levels qualifications category, because efforts were made in the MCS2001 to categorise alternative qualifications into the relevant category, so any remaining in this category were considered to be no or low level qualifications. Diplomas in higher education were collapsed into the post-16 qualifications category, since these are equivalent to 2 years of undergraduate study.

When the MCS2001 data was considered separately (Chapters 4 and 5), a combined measure of academic and vocational qualifications was used. This was derived in the following way:

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

In all analyses, the highest household level of parent education was the overall variable used.

2.3.2.2. Occupational Status

Occupational status, or social class based on occupation, has been used to classify members of British society throughout the 20th century, based on the idea that such classifications provide useful summaries regarding levels of social variation and inequalities in mortality and health, which are vital for informing and evaluating policies (Rose & Pevalin, 2005). Two measures of occupational status are utilised in this thesis: the Registrar General's Classification Scheme (RGCS; renamed Social Class based on Occupation in 1990), and the National Statistics Socio-economic classification (NS-SEC). The RGCS was used to collect information about occupational status in the BCS1970 cohort, and the NS-SEC was used to collect information about occupational status in the MCS2001 cohort. Although the NS-SEC replaced the RGCS as the national measure of occupational status, the RGCS was the standard classification system used by the British government from 1911-1980 (Rose & Pevalin, 2005), the period during which the BCS1970 began. The Registrar General's Office later became known as the Office of Population Censuses and Surveys, and eventually the Office for National Statistics (ONS).

The RGCS was based on the idea that British society is an ordered hierarchy of occupations, which are ranked based on skill and their position in society (Rose & Pevalin, 2001). This schema was used by the Office of Population Censuses and Surveys (what is now known as the Office for National Statistics; ONS), and while there were subtle changes to the classification of occupations and in the assignment of occupations to specific social classes with each census, the overall structure of the schema remained unchanged (Rose & Pevalin, 2001). The placement of individuals into social classes had three stages, with individuals first being assigned an occupational group based on the type of employment they were in. These occupational groups were subsequently allocated holistically to a specific social class,

regardless of differences between individuals within the same occupation groups. Individuals of a particular employment status, such as managers, were then placed into a different social class, regardless of the overall class their occupation was assigned (Rose, 1995). This process was completed by the Registrar General's office.

In 1994, a review of the government social classifications was undertaken by the ESRC, and subsequently, the NS-SEC replaced the RGCS (Rose & Pevalin, 2005). The NS-SEC was based on the Goldthorpe Class Schema, which aimed to classify individuals based on similarities in work and market circumstances, with a later emphasis on employment relations in occupations (Erikson & Goldthorpe, 1991; Goldthorpe et al., 1980). The decision to base the NS-SEC on the Goldthorpe Class Schema was made given it's conceptual clarity. The NS-SEC was developed to measure the employment relations and conditions of different occupations, in order to demonstrate the socioeconomic structure of society (Rose & Pevalin, 2005). Since this classification is based on social relations, it is not inherently an ordered hierarchy in the way that the RGCS was. However, there is some ordering between classes, with some occupations being more advantaged compared to others, for example, managers in large organisations, compared to those in intermediate occupations. Further, the 3-category measure (used in this thesis) can be conceptualised as hierarchical (Rose & Pevalin, 2005).The Goldthorpe Schema emphasised a distinction between employers, employees and the self-employed in terms of experienced employment relations and conditions. Membership of different classes reflects different sources and levels of income, as well as differing degrees of income and job security. The NS-SEC follows a similar distinction, but with an additional category of those who are excluded from the paid workforce. This includes those who are retired, those who have never worked and those who are in long term unemployment, ensuring an almost-universal coverage of the NS-SEC; the only group not covered by this schema are children (Rose & Pevalin, 2005).

The categories of the NS-SEC make distinctions between the positions resulting from social relationships in employment, regarding the regulation of employees with employment contracts. There are three types of social relationship: 1) service relationship, where the employee provides the employer with a service to receive compensation, in the form of a salary and job security; 2) a labour contract, where employees give distinct amounts of work in return for a wage, based on the hours worked; and 3) intermediate relationships, which are a combination of the service relationship and the labour contract (Rose & Pevalin, 2005). The concept of skill required for a specific occupation are not part of the NS-SEC; rather it aims

to establish qualitative distinctions in different employment relations, and members of different occupations may be advantaged, or disadvantaged, in different ways.

There may be some conceptual differences between the 3-class versions of these scales, which are employed in this thesis. As aforementioned, the RGCS ranked occupations based on the level of skill required, whereas the NS-SEC classifies individuals based on the employment relations and working conditions of their occupations, and the level of skill required plays no role. Although the 3-class measure of the NS-SEC can be considered as hierarchical, there may be differences in the types of occupations that make up these 3 classes, between the RGCS and the NS-SEC. For example, intermediate occupations of the NS-SEC consist of occupations that have a combination of a service relationship and a labour contract; for example, these occupations would not usually have any form of authority (Rose & Pevalin, 2005). The 'equivalent' class in the 3-class version of the RGCS used in this thesis consists of manual and non-manual skilled workers, and so these groups may not be directly comparable, as this class is informed directly by level of skill required. The possibility of such conceptual differences between these two measures needs to be kept in mind when interpreting results. However, the NS-SEC is now the nationally used, standardised classification of occupations, based on employment relations and conditions of different occupations. It is seen as theoretically clear and thorough, allowing explanations of relationships between occupational status and life outcomes to be made, and it is comparable to the older social class based on occupations (RGCS) (Rose & Pevalin, 2005).

Finally, the ONS states that the NS-SEC should be conceptualised at the household, rather than the individual level, since families are interdependent units who experience shared conditions, and the NS-SEC of one family member may have less of an effect than another on the life chances of other family members (Rose & Pevalin, 2005). It is therefore recommended to define a "household reference person" (HRP), who is often defined as the family member who owns, or is responsible for renting, the house. Where the house is under joint ownership, the person with the highest income is the HRP, and where incomes are equal, the older person is the HRP. Conceptualising the HRP in this way increases the chance that the HRP will be female (Rose & Pevalin, 2005). However, a "dominance" approach can also be taken, where people in occupational classes seen as "dominant" (i.e., the highest occupation – for example, full time work is dominant over part time work, and higher managerial occupations are seen as dominant to intermediate occupations, which are in turn seen as dominant to routine occupations) are given the role of HRP. In this thesis, a highest household level of occupational status approach was taken, since in the MCS2001, tenure

information was only available for the main respondent, yet NS-SEC categories were available for both the main and partner respondents, and in the BCS1970 childhood sweeps, tenure was not reported at the individual household member level. I therefore chose to select the highest household occupational status by way of defining the HRP. This also allowed harmonisation of the measure of occupational status in both cohorts for our cross-cohort comparisons.

BCS1970

The RGCS was used to collect information about occupational status in the BCS1970 in the birth sweep of data collection. When cohort members were aged 5 (data collected in 1975), the classification of occupations used in the 1971 census survey was used (Office of Population Censuses and Surveys, 1970). These categories are detailed below:

RGCS – Birth sweep (Institute of Child Health, 1970)

- Social Class I: Professional occupations. This included occupations such as doctors, lawyers, ministers of religion, university teachers
- Social Class II: Managerial and other professionals, such as nurses, school teachers, or company directors
- Social Class III (non-manual): non-manual skilled occupations, such as shop assistants and clerical workers
- Social Class III (manual): manual skilled occupations, such as mechanics and craftsmen
- Social Class IV: Semi-skilled workers, such as machine operators, postmen and caretakers
- Social Class V: Unskilled workers, such as cleaners and rubbish collectors

During the birth sweep, the mother reported her husband's current occupation, and this was then converted into the relevant social class group. For the mother, the occupation prior to her pregnancy with the cohort member was reported.

Classification of Occupations (Office of Population Censuses and Surveys, 1970) – formerly the Registrar General's Office; Age 5 Sweep

This had the same categories outlined above, with some changes to the classification of occupations and in the assignment of occupations to specific social classes, which was

typical with each census. Those who were unemployed or retired were given an occupational status based on their last job (Institute of Child Health, 1975).

For this thesis, a 3-category scale, with a 4th category for unemployment was used. This was derived in the following way:

1. Professional & Managerial (Social Classes I and II)
2. Skilled (Social Classes III NM and III M)
3. Semi-skilled and unskilled (Social Classes IV and V)
4. Unemployed (those who did not have an occupation reported and reported unemployment were coded as unemployed. This included students and volunteers, as they have no paid employment).

Where the age 5 measure of occupational status is used (see Chapter 4), if occupational status was missing, the occupation from the birth sweep was used, and if this was missing and unemployment reported, they were categorised as unemployed. This was calculated for both the mother and her husband, and then the highest household occupational status was selected as the final variable.

MCS2001

In the MCS2001, the NS-SEC was used as the measure of occupational status. The NS-SEC can be conceptualised as eight, five or three categories, and the full version can be found here: https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010#classes-and-collapses. The 5-classes version of the NS-SEC, which was collapsed into 3 classes for this thesis, can be found below.

NS-SEC 5 classes:

1. Higher managerial/administrative/professional occupations
2. Intermediate occupations
3. Small employers/self-employed
4. Lower supervisory and technical occupations
5. Semi-routine and routine occupations

In this thesis, the above 5 class version was collapsed into a 3- class version, with a 4th category for unemployment, in the following way:

1. Higher managerial
2. Intermediate (intermediate occupations, small employers, self-employed)
3. Routine (lower supervisory, technical, semi-routine and routine occupations)
4. Unemployed (those who did not have an occupation reported and reported unemployment were coded as unemployed. This included students and volunteers, as they have no paid employment)

I used a measure of occupational status at birth/9 months (Chapter 3); age 3 (Chapters 4 and 5); and age 5 (Chapter 4). In each case, where measure of occupational status was missing, the occupation from the previous sweep was used, and if this was missing and unemployment reported, they were categorised as unemployed. This was calculated for both the mother and her husband, and then the highest household occupational status was selected as the final variable.

2.3.2.3. Income

Income was considered as an indicator of SEC in the MCS2001 cohort only. The first measure of income in the BCS1970 was collected when cohort members were aged 10, and to ensure consistency with measures of early childhood SEC in the MCS2001 used in cross-cohort comparisons, the decision was made to omit a measure of income in analyses of BCS1970 data. Data about income was collected in a variety of ways in the MCS2001 — for example, gross and net earnings, income from benefits, earnings from second jobs, or jobs seekers allowance, and maternity allowance (see Centre for Longitudinal Studies, 2020, for an exhaustive list). Banded total income and estimated net weekly income were also included.

A measure of OECD equivalised income quintiles was used in all analyses in this thesis. OECD equivalised income accounts for household size and composition, including the age of members in the household, since the financial requirements of a household increase with each additional person present in the household: a large household with many members will need a greater income to experience the same standard of living as a smaller household (Office for National Statistics, 2015). Further, children will usually have lower living costs than adults. Generally speaking, single adult households with higher incomes will experience better standards of living, compared to a larger household with the same income. Once income has been equivalised, households with the same equivalised income have similar

standards of living. There are different methods available for equivalising income, however the MCS2001 applied the OECD-modified method to estimated net weekly income.

Equivalisation of income

Equivalisation scales are used to weight household income, considering household size and composition. The OECD-modified equivalence scale is as follows: single adult households are given an equivalence value of 1. Each additional adult present in the household is given a value of 0.5 (this smaller number is a result of "economies of scale", for example the sharing of water and electricity reduces the total living cost). Children who are aged 14 or over are also given a value of 0.5, as they are thought to have the same living costs as adults. Finally, any children aged 13 or younger are assigned a value of 0.3, due to their smaller living costs. Each household member is given an equivalence value. These are then totalled to give the household's equivalence value. Finally, the household's total net income is divided by the household's equivalence value to give the OECD adjusted income for that household (ONS, 2015).

In the MCS2001, this equivalisation method was applied to estimated net weekly income, and this was then divided into quintiles to give OECD weighted income quintiles. In Chapters 4 and 5 of this thesis, these quintiles were taken from the age 3 sweep, and in Chapter 3, they were taken from the 9 months sweep.

2.3.2.4. Wealth

Again, wealth was considered as an indicator of SEC in the MCS2001 cohort only. Total net wealth can be seen as the combination of "financial wealth" and "housing wealth". Financial wealth refers to the total value of any assets, such as current or savings accounts, Individual Savings Accounts (ISAs), endowments, stocks and shared, informal savings, net of outstanding debts, such as credit card debts, household bills, loans or student loans. Housing, or property, wealth, refers to an individual's self-valuation of their owned properties (this can include their main home and any other additional properties, such as holiday homes, owned), net of the outstanding mortgages owed on said properties (Office for National Statistics, 2019b).

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. These variables were used to derive a measure of total net wealth as follows:

Financial wealth = total assets and investments – total debts owed

Housing wealth = house value – outstanding mortgages

Total net wealth = financial wealth + housing wealth

The distribution of the wealth variable used in this thesis can be found in Appendix 2. Some caveats on this wealth measure must be noted. First, a recent study, which compared measures of wealth in the MCS2001 with wave 3 of the Wealth Asset Survey, demonstrated that whilst the MCS2001 captures housing wealth well, this was not the case for financial wealth, and underestimates of savings in the MCS2001 was evident. However, these authors concluded that the relative ranking of individuals was unlikely to be affected (Moulton et al., 2021). Second, this thesis did not account for those who are not homeowners, and would therefore automatically have a housing wealth value as missing instead of 0, as would be the case for financial wealth for those who had no savings or debts. This caveat is likely related to the high levels of missing data evident for this variable (see Chapters 4 and 5).

2.3.2.5. Relative Neighbourhood Deprivation

The final measure of SEC considered in this thesis is a measure of 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, 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, based on small area levels. These small area levels are called Lower Layer Super Output Areas, which are small areas of roughly equal size, which contain ~1500 people. An area can thus be considered as deprived *relative* to another area, if more people in that area are experiencing deprivation; the area itself is not deprived, rather areas get their deprivation statistics based on the deprivation experiences of the people living in that area (Mclennan et al., 2019). Not every individual living in a deprived area, defined by IMD ranks, will themselves be necessarily experiencing high levels of deprivation — rather the IMD is a measure of relative deprivation.

The indices are based on 7 indicators of deprivation, which are combined and weighted to form an overall IMD; the indicators used in each UK country are slightly different, meaning they are not directly comparable. This thesis used The English Indices of Deprivation (2004); Welsh Assembly IMD (2005); The Scottish Assembly IMD (2004); and

the NISRA Multiple deprivation measure (2005) (Centre for Longitudinal Studies, 2020). For England, the 7 indicators are income, employment, health deprivation and disability, education, skills and training, crime, barriers to housing and services and living environment (Noble et al., 2004). For Wales, the indicators are income, employment, health deprivation and disability, education, barriers to housing and services, and physical living environment (Stats Wales, 2005). In Scotland, the indicators used are income, employment, health, education, skills and training, geographic access to services, crime, and housing (Scottish Government, 2004). Finally, the indicators used in Northern Ireland are income deprivation, employment deprivation, health deprivation and disability, education, skills and training deprivation, proximity to services, living environment, and crime (Shaw et al., 2005). However, our aim in this thesis was not to compare deprivation across the UK countries. Rather, we were interested on the effects of how deprived a neighbourhood was on language ability, not whether this differed across UK countries.

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 et al., 2019). If an individual, family or area on which the IMD is based experiences more than one of the indicators of deprivation, this is seen as more deprived than experiencing only one of the 7 indicators.

IMD deciles were created from the total IMD ranks for each country. I used the IMD data collected when cohort members were 3.

2.4. Missing Data

2.4.1. Types of Missing Data

All longitudinal studies suffer from missing data. Non-response occurs when cohort members are completely missing from follow up sweeps. This can be in the form of attrition, which occurs when cohort members are permanently lost from follow up, or sweep non-response, which occurs when respondents are missing from a follow up sweep, but take part in subsequent sweeps. Unit non-response occurs when a whole interview or questionnaire is missing for a cohort member, despite that cohort member taking part in the overall follow up sweep: for example, parents may have completed the parent interview elements, but the cohort member may not have completed the cognitive assessment during a specific follow up.

Finally, item non-response occurs when cohort members do not complete every part of a questionnaire, by for example not answering all questions on a mental health scale, or skipping questions related to income.

In addition to reducing the sample size, missing data can result in the absence of cohort members with specific characteristics, and can therefore bias the remaining sample, making it no longer representative of the population. For example, in the BCS1970, males, ethnic minority groups, cohort members who were born to single mothers, or whose parents left school at an early age or had no qualifications, and those from lower social class backgrounds tend to be underrepresented in follow up sweeps (Mostafa & Wiggins, 2014). In the MCS2001, families living in rented accommodation, cohort members who were born to single mothers, cohort members who were born to young mothers, non-breastfeeding mothers and ethnic minority groups are likely to be underrepresented at follow up sweeps (Mostafa & Ploubidis, 2017; Plewis, 2007a).

There are three possible mechanisms of missing data: data can be missing completely at random (MCAR), missing at random (MAR) or missing not at random (MNAR) (Little & Rubin, 2019). MCAR assumes that the probability of missing data is not related to the dataset (both other measured variables and values of the missing variable itself). MAR assumes that the probability of missing data of a specific variable is related to other variables in the dataset, but not to values of the missing variable itself. Finally, MNAR assumes that the probability of missing data of a specific variable is related to values of that specific variable, even after adjusting for other variables that may be correlated with missingness (Enders, 2010).

There are several approaches to handling missing data in analyses. Listwise deletion (or complete case analysis) refers to restricting analyses to those with complete data on all variables of interest and can produce biased estimates if data are not MCAR. Single imputations, such as imputing the missing data with the mean values for that item can also produce biased estimates and smaller standard errors, as they treat the imputed data as if it were the observed data values, thus underestimating sampling error (Enders, 2010). Multiple imputation is a means of handling missing data in analyses, that does not result in attenuated standard errors.

2.4.2. Multiple Imputation using Chained Equations (MICE)

Multiple imputation generates multiple copies of the observed dataset, replacing the missing values in each with different plausible estimates. In doing so, multiple imputation accounts for the uncertainty of the missing data, as it provides a means of estimating additional source of sampling error (between imputation variance; see below), which single imputation methods do not do due to the fact that they treat the imputed data as complete observed data (Enders, 2010). Multiple imputation assumes the data is MAR, meaning missingness is related to other variables in the dataset, and once these variables have been accounted for, any remaining missingness is random (Graham, 2009). In order to increase the probability of satisfying this assumption, auxiliary variables must be included in the imputation model: these are variables that will not be included in subsequent analyses, but that are correlated with the missing variable or missingness (Graham, 2009). Including auxiliary variables as predictors during the imputation process aims to minimise bias in the imputations. In this thesis, the age of the mother at the birth of the cohort member, whether the cohort member was breastfed, housing tenure status, accommodation type and the number of carers present in the household were included as auxiliary variables, as these variables have been shown to be related to missingness (Mostafa & Wiggins, 2014; Mostafa & Ploubidis, 2017; Plewis, 2007a). Imputation models also included cohort member sex, ethnicity and the language spoken in the home, and measures of childhood SEC, as these variables were included in analysis models. Analysis models in Chapter 3 included measures of age 5 internalising and externalising symptoms, therefore these variables were used to inform imputations in Chapter 3, and as Chapter 4 included vocabulary measured at multiple ages, earlier vocabulary ability was included in imputation models in this chapter. In multiple imputation, the missing values are replaced with predicted values that take into account the observed values for a given cohort member, and relations that are present in the data for other cohort members (Azur et al., 2011).

There are three stages to multiple imputation: the imputation phase, the analysis phase, and the pooling phase. The multiple copies of the data are created in the imputation phase. Missing data in all analyses in this thesis was accounted for with multiple imputation using chained equations (MICE), using the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011). Compared to other multiple imputation approaches, MICE enables categorical and continuous data to be imputed in the same imputation model, as it imputes the data one variable at a time (Enders, 2010). MICE is an iterative process and imputes the data using a series of regression models (for example, linear regression for continuous variables, logistic regression for binary variables), where each variable is modelled against other

variables in the dataset (variables that will be included in subsequent analyses and auxiliary variables). In the first iteration, single imputations (such as a mean imputation) for every variable in the dataset are computed as "place holders", which are subsequently set back to missing for the first variable that is to be imputed. The observed data for this variable is the outcome variable in a regression model, which is predicted by all other variables in the imputation model. The missing values of the variable being imputed are replaced with predicted values from this regression model. Both the observed and predicted values for this variable are subsequently used as a predictor variable in regression models for the other variables. The regression steps are repeated for every variable that contains missing values in the dataset. Once this has been performed for all variables, the first iteration is complete, and the missingness in all variables has been replaced with predicted values which reflect relationships between the variables in the data. As MICE is an iterative process, these steps are repeated a number of times (in this thesis, 5 times), whereby the observed and predicted values imputed in the previous cycle are used to predict new values for each variable containing missing values. The imputed values are updated at the end of each iteration, and these new predicted values are used in the next iteration. In the final iteration, the imputed values are kept, resulting in the first imputed dataset. The imputation process repeats *m* number of times (the number of imputed datasets; in this thesis, I have 25 imputed datasets for each analysis) to produce multiple imputed datasets. In the resulting imputed datasets, the observed data will be identical and only the data that were originally missing will differ between each dataset (Azur et al., 2011; Enders, 2010).

In the analysis phase, the planned analysis (for example, a multiple regression) is ran *m* times (once for each imputed dataset). Importantly, the imputed datasets are analysed individually, and estimates are then combined, rather than the imputed datasets being combined and then analysed. In the final stage of the imputation, the pooling phase, parameter estimates and standard errors are pooled into a single set of estimates in the pooling phase, according to Rubin's rules (Rubin, 1987). Parameter estimates (such as regression coefficients) are simply averaged across all imputed datasets. The pooled standard error considers two sources of variation: the sampling error that would have occurred if the observed data been complete (within imputation variance), and the sampling error that occurs as a result of the missing data (between imputation variance). The between imputation variance refers to the variability of the parameter estimates between each imputed dataset. The pooled standard error is the square root of the combined within and between imputation

variance (Enders, 2010). The pooled results from analyses of multiply imputed data are the results that are then interpreted.

Chapter 3 : Does early child language predict internalizing symptoms in adolescence? An investigation in two birth cohorts born 30 years apart

This chapter is published in Child Development: Thornton, E., Patalay, P., Matthews, D., & Bannard, C. (2021). Does early child language predict internalizing symptoms in adolescence? An investigation in two birth cohorts born 30 years apart. *Child Development*, *92*(5), 2106-2127.

Fit within thesis

As discussed in Chapter 1, this thesis aims to investigate the relation between age 5 vocabulary and adolescent outcomes, having appropriately adjusted for potential confounding variables, such as childhood socioeconomic status. The published paper presented in this chapter focusses on the outcome of internalising mental health symptoms (depression and anxiety). I investigated the relation between age 5 vocabulary and adolescent self- and parent-reported internalising mental health, in two cohorts born 30 years apart. I adjusted for sociodemographic variables, and mother and child psychosocial variables in order to obtain the best possible estimates. I also conducted two exploratory analyses to investigate whether the pattern of results remained when considering vocabulary as a binary variable (to indicate clinical levels of difficulty) and internalising mental health as a binary variable (to indicate clinical levels of internalising symptoms). This paper is an important contribution to the overall aims of this thesis, as it identifies the nature of the relation between age 5 vocabulary and adolescent internalising mental health, and provides insight into the value of improving age 5 vocabulary to promote later mental health.

This paper was conceptualised and designed by myself, and my supervisors (Praveetha Patalay, Danielle Matthews, and Colin Bannard). I was responsible for all data management, wrangling and analysis tasks, and writing the paper. All authors were involved in the production of the final manuscript, following reviewer comments from Child Development. Two deviations from the published version of this paper are of note here: in the published paper, the term Socioeconomic Status (SES) is used, however for consistency with the rest of this thesis, here the term Socioeconomic Circumstance (SEC) is used. For further consistency, the cohorts are reported as BCS1970 (BCS in the published paper) and MCS2001 (MCS in the published paper).

3.1. Abstract

Language is vital for social interaction, leading some to suggest early linguistic ability paves the way for good adolescent mental health. The relation between age-5 vocabulary and adolescent internalizing symptoms was examined in two UK birth cohorts that are nationally representative in terms of sex, ethnicity and SEC: the 1970 British Cohort Study (BCS1970; N=11,640) and the Millennium Cohort Study (MCS2001 born~2001; N=14,754). In the BCS1970, no relation between receptive vocabulary and age-16 self-reported symptoms was observed (β=.00[-.03;.03]). In the MCS2001, better expressive vocabulary was associated with *more* age-14 self-reported symptoms (β=.05[.02;.07]). The direction of this effect was reversed for parent-reported symptoms. All effect sizes were small. The relation between childhood vocabulary and internalizing symptoms varies by generation and reporter.

3.2. Introduction

Early language skills are frequently claimed to be an important contributing factor to later mental health, with calls being made for early language interventions to prevent later mental health problems (Bercow, 2018; Miller et al., 2020; Oxford University Press, 2018). Theoretically, language ability could be important for mental health because it is the primary medium for social interaction and because it supports self-regulation (Redmond & Rice, 1998; Salmon et al., 2016). However, while there is some work to suggest poorer mental health outcomes for children with Developmental Language Disorder (DLD; Yew & O'Kearney, 2013), relatively little empirical work has tested this association in the general population. If an association holds across the continuum of language ability, this would support calls for widespread early language intervention to improve adolescent mental health.

This paper focuses on adolescent internalizing symptoms (emotional difficulties), which includes symptoms of the most common mental health problems such as anxiety and depression. These stand in contrast to externalizing symptoms (behavioral problems) such as poor impulse control and aggression (Willner et al., 2017), which we do not examine in this paper. Adolescence is a critical period for susceptibility to internalizing mental ill health (McLaughlin & King, 2016). If early language ability were to have any direct effect on mental health, there is good reason to expect there to be evidence of its impact on internalizing symptoms at this pivotal developmental stage. Since adolescent mental health difficulties 1) predict social exclusion, stigma, poor educational attainment, risky behaviours and poor physical health (Clayborne et al., 2019), 2) often persist into adulthood (Fergusson et al., 2005) and 3) are becoming more prevalent (Patalay & Gage, 2019), the prevention of mental ill health in adolescence is a priority (Thapar et al., 2012).

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

peers (see the Social Adaptation Model: Redmond & Rice, 1998). Since positive peer interaction is known to be important for adolescent mental health (Thapar et al, 2012), we might expect that a greater degree of language difficulty would put children at risk of later internalizing difficulties.

A number of studies have found that children with DLD when aged 4-7 years are at increased risk of later mental health difficulties when aged 8-19 years, compared with their typically-developing peers (Beitchman et al., 2001; Conti-Ramsden & Botting, 2008; Conti-Ramsden, Mok, Pickles, & Durkin, 2013; Wadman, Botting, Durkin, & Conti-Ramsden, 2011; although see Redmond & Rice, 2002; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006 for counterevidence). These findings are hard to interpret since, despite group differences, continuous measures of language ability do not always predict internalizing symptoms in the children studied (Conti-Ramsden & Botting, 2008; Wadman et al, 2011). A meta-analysis conducted by Yew and O'Kearney (2013) noted that children with early language difficulties experience emotional problems of an increased severity and frequency compared to their typically developing counterparts. However, very few studies to date have controlled for baseline emotional or behavioral difficulties and so we cannot be sure that language problems explain unique variance in later mental health difficulties. Nonetheless, recent research with the Millennium Cohort Study (which did account for such factors), found that those 'at risk' of DLD at age 5 (operationalized as having low vocabulary scores and/or parent reported language difficulties) were more likely to have parent-reported internalizing symptoms at age 11 (Forrest et al., 2020; St Clair et al., 2019). Furthermore, evidence from the 1970 British Cohort Study suggests 5-year-olds with language difficulties are more likely to self-report internalizing symptoms at age 34 (Schoon et al., 2010).

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

association between language ability and internalising symptoms (Hentges et al., 2021). However, there are a number of reasons why further research is warranted with large cohort studies run across generations in nationally representative samples.

First, few studies to date have adequately controlled for factors such as early child and parent mental health difficulties and socioeconomic circumstance (SEC). Vocabulary size and processing speed are positively associated with social advantage from 18 months in the United States and the United Kingdom (Fernald et al., 2013; McGillion et al., 2017; Pace et al., 2017), a relation that persists throughout the lifespan (Sullivan et al., 2021). We therefore included robust SEC confounders across our analyses.

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

Second, in studies to date, information about mental health has been obtained either by asking the individual concerned (for example, Conti-Ramsden & Botting, 2008; Wadman et al, 2011; Conti-Ramsden et al, 2013), by asking others such as parents (for example, St Clair et al, 2019; Forrest et al, 2020), or by some conflation of these measures (for example Bornstein et al, 2013; Miller et al, 2020). When considering internalizing symptoms (such as feelings of low mood or worrying), individuals themselves are uniquely well positioned to know how they are feeling. Many self-report measures have been validated for use with clinical and community samples (for example, Sharp, Goodyer, & Croudace, 2006; Thabrew, Stasiak, Bavin, Frampton, & Merry, 2018) and self-reports are the recommended measure to use according to The Good Childhood Report (The Children's Society, 2019). In contrast, parents and other adults may not know the full extent of internalizing symptoms, unless the adolescent discloses such feelings to them. Indeed, parent and self-report measures are not highly correlated (typical correlations ~0.2; Rescorla et al., 2013). In the current study, self-report was pre-registered as our primary outcome measure. In additional analyses, we then tested whether choice of self-report over parent-report affected our findings.

Finally, the relation between language and mental health could plausibly be changing over historical time (Yew & O'Kearney, 2013). It has been argued that the transition to a knowledge-based economy has increased the economic importance of cognitive resources,

including language (Beddington et al., 2008). At the same time, adolescent internalizing problems have become more prevalent (Bor et al., 2014; Patalay & Gage, 2019). The current research therefore explored the relation between vocabulary and adolescent internalizing mental health in two large, nationally representative cohort studies with cohort members born 30 years apart: the 1970 British Cohort Study (BCS1970, children born in 1970) and the Millennium Cohort Study (MCS2001, children born in 2000-02). This cross-cohort comparison allowed us to investigate the relation across a time period that has seen an increase in both reliance on cognitive ability, and in the prevalence of internalizing mental health difficulties.

We pre-registered two main analyses to assess whether early vocabulary is associated with self-reported adolescent internalizing symptoms in the general population. The first analysis assessed this with the BCS1970 and the second with the MCS2001. To better understand the findings and connect them with existing literature, we also report two exploratory analyses. The first repeated the main analyses, but with the vocabulary predictor dichotomised (whether or not the child had a language difficulty, operationalized as scoring 1SD below the mean for vocabulary). This permitted comparison with prior work that has sought to identify children with language difficulties in this way (for example, Schoon et al., 2010). The second exploratory analysis repeated the main analyses, but with self-reported adolescent internalizing symptoms considered as a binary outcome, according to clinical threshold cut-offs. This allowed us to check whether the relation between vocabulary and internalizing symptoms differed for those with clinical levels of internalizing symptoms. Finally, Appendix 1 reports three analyses that assessed the role of vocabulary when parent-reported adolescent internalizing symptoms were considered as the outcome in each cohort, and when an adult outcome point was considered for the BCS1970. The latter analysis allows comparison with Schoon et al.'s (2010) findings of an association between early vocabulary (dichotomous variable: difficulty or not) and adult mental health in the BCS1970 (see Appendix 1, Section 9).

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

3.3. Method

Data

Data from two national birth cohort studies were used: the 1970 British Cohort Study (BCS1970) and the Millennium Cohort Study (MCS2001). The BCS1970 follows 16,571 children born in England, Scotland and Wales during one week in 1970 (Elliott & Shepherd, 2006) and has 4 childhood sweeps (ages 0, 5, 10 and 16 years). More information about this cohort study can be found here: https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study/. The MCS2001 follows 19,244 young people born across England, Scotland, Wales and Northern Ireland in 2000-02 (Connelly & Platt, 2014) and there are currently seven sweeps (ages 9 months, 3, 5, 7, 11 and 14 years). The age 14 sweep of the MCS2001 took place in 2015, and therefore this cohort represents contemporary adolescence. More information about this cohort study can be found here: https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/.

Participants

For the BCS1970, information about all babies born between 5th April-11th April 1970 was requested (this was not restricted to babies born in the NHS; Institute of Child Health, 1970). The sample was supplemented with children who were born in the eligible week and had subsequently moved to the UK; there were an additional 79 new cohort members at age 5, 294 at age 10 and 65 at age 16 (CLS website: https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study/.).

For the MCS2001, a stratified clustered sample design was used, which specifically over-recruited subgroups of the population (ethnic minorities, disadvantaged areas and the smaller UK countries). Eligible children (living in the UK at age 9 months, born within the eligible time period — 1st September 2000-31st August 2001 for England and Wales, and 23rd November 2000-11th January 2002 for Scotland and Northern Ireland — and receiving child benefit at age 9 months) were identified by government child benefit records and sampled from electoral wards (Connelly & Platt, 2014). 72% of eligible families responded to the 9 months sweep of data collection. The original sample was supplemented in the age 3 sweep with families who were eligible to be included, but were not recruited due to recently moving to the eligible address; this resulted in an additional 692 families being interviewed (Connelly & Platt, 2014).

Families with multiple births in the cohorts were excluded due to possible differences in the language learning environments experienced by these children (BCS1970:189 pairs of

twins and 1 set of triplets — 2.30%; MCS2001: 251 pairs of twins, 11 sets of triplets and 6 families with two singleton cohort members — 2.84%; Thorpe, Rutter, & Greenwood, 2003).

For the BCS1970, we selected singleton cohort members with complete responses for the English Picture Vocabulary Test (EPVT; Brimer & Dunn, 1962; age 5). This resulted in a sample of 11,640 individuals. The majority of cohort members in our analytic sample were of White ethnicity (96%) and spoke only English (98%).

For the MCS2001, we considered singleton cohort members with complete responses on the British Ability Scale (BAS II) naming vocabulary scale (age 5; Elliott, Smith, & McCulloch, 1996), resulting in a final sample of 14,754 cohort members. 89% of our analytic sample were of a White ethnicity, and 90% spoke only English in the home.

Measures

Predictor variable: age 5 vocabulary

For the BCS1970, receptive vocabulary was measured at age 5 using the EPVT (Brimer & Dunn, 1962; see Chapter 2 for details). The EPVT has been reported to have a reliability coefficient of .96 (Osborn et al., 1984). For the MCS2001, expressive vocabulary was measured using the naming vocabulary sub-test of the BAS II (Elliott, Smith & McCulloch, 1996), which was administered to cohort members during the third sweep (aged around 5 years; see Chapter Two). The Naming Vocabulary subscale of the BAS has been reported to have a reliability coefficient of .65 for five-year-olds (Elliott, Smith, & McCulloch, 1997). Note that receptive and expressive vocabulary measures tend to be moderately to highly correlated (e.g. Conway et al., 2017).

Due to the nature of the naming vocabulary test, MCS2001 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. Therefore, in our analyses, we used ability scores, adjusted for item difficulty. The same set of items were administered to all BCS1970 cohort members and raw scores were therefore used. Because MCS2001 cohort members were born over a 2-year period (2000-02), they were different ages when they completed the naming vocabulary test (mean age of 62.51 months, range 52.87 to 75.52 months). Additionally, fieldwork in the BCS1970 age 5 follow up was conducted over 6 months in 1975, and cohort members were thus different ages when they completed the EPVT (mean of 60.92 months, ranging from 58.78 to 75.52 months). We therefore adjusted for age in months at the time of the test, in both cohorts. In both tests, higher scores indicated a higher ability. All scores and ages were converted to *z*-scores for analyses.

Outcome variable: adolescent internalizing symptoms

For the BCS1970, total scores on the 9-item Malaise Inventory (Rutter et al., 1970) were used as a measure of internalizing symptoms at age 16. Scores ranged from 0-9. For the MCS2001, cohort members were given the 13-item Short Mood and Feelings Questionnaire (SMFQ; Angold, Costello, & Messer, 1995) at age 14. Scores ranged from 0-26. For both scales higher scores indicate greater severity of internalizing symptoms. These two scales have similar items relating to the same domains, such as tiredness, restlessness and mood and both are reliable and valid indicators of internalizing symptoms (Daviss et al., 2006; Rutter et al., 1970). In the current samples, there was an alpha of 0.7 (BCS1970 Malaise Inventory) and 0.93 (MCS2001 SMFQ). All scores were converted to *z*-scores for analyses.

Potential confounding variables

Biological and SEC variables. Biological risk variables included in all models were the child's birthweight (in grams, converted to *z*-scores for analyses), gestational age in days (converted to *z*-scores for analyses), sex (male = 0, female=1). We also included ethnicity and the main language spoken in the home. SEC variables included in all models were the highest level of parental education achieved and highest occupational status in the household at birth (a 3-category measure with a 4th category for unemployment. BCS1970: 1. professional & managerial, 2. skilled, 3. semi-skilled & unskilled; MCS2001: 1. higher managerial, 2. intermediate, 3. routine & manual). The MCS2001 has a richer set of indicators of SEC, allowing us to include two further SEC variables for MCS2001 analyses: UK OECD weighted income quintiles (taken from the first sweep, an indication of household income 1=lowest, 5=highest) and net total wealth (converted to *z*-scores for analyses). We derived the latter measure by summing information collected at age 11 (MCS5) about net housing wealth (house value net of outstanding mortgage) and net financial wealth (total savings net of any owed debts).

Mother and child psychosocial variables. Maternal psychosocial variables included whether the cohort member's mother was a teenage mother at their birth (0=yes, 1=no), the marital status of the mother at birth (0=partnered, 1=not partnered) and maternal depression when children were aged 5. In the BCS1970, this was assessed using the Malaise Inventory (Full version; Rutter et al, 1970). In the MCS2001, this was assessed using the Kessler K6 scale (Kessler et al., 2003). Items on the two scales are similar; for example, both ask

questions regarding feelings of low mood and restlessness. These variables were converted to *z*-scores for analyses.

Child psychosocial variables were internalizing and externalizing difficulties at the age of 5. In the BCS1970, cohort member's parents completed the Rutter "A" scale (Rutter et al, 1970). For the current study, a neurotic score and an antisocial score were calculated, as detailed by Rutter et al. (1970) and used as indicators of internalizing and externalizing behavioral difficulties respectively. In the MCS2001, parents completed the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), which was developed as the successor to the Rutter scales and items on the SDQ emotional symptoms and conduct problems subscales are similar to those of the antisocial and neurotic subscales of the Rutter "A" scale. Total scores from the emotional symptoms and conduct problems subscales were calculated as indicators of internalizing and externalizing behavior problems, respectively. For both scales, a higher score indicates increased difficulties. These variables were converted to *z*-scores for analyses.

Data analysis

The main analyses consisted of 2 multiple linear regressions: 1) BCS1970 data, with age 16 internalizing symptoms as the outcome and 2) MCS2001 data, with age 14 internalizing symptoms as the outcome. These confirmatory analyses were pre-registered at the Open Science Framework website (OSF number: osf.io/a94bh).

Missing data strategy. Sampling weights were applied to the analyses of MCS2001 data, to account for the stratified clustered design of the data and the oversampling of subgroups. The BCS1970 does not have a complex survey design and therefore sample weights were not required for this cohort. 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). Each dataset was imputed 25 times, as this was greater than the percentage of missing cells in both cohorts (7.67% BCS1970, 10.26% MCS2001). Across our chosen samples for each analysis, no data was missing for the main predictor variable (vocabulary score) or sex. All analyses were conducted in RStudio (RStudio Team, 2020).

Analysis plan. Initially, the raw association between vocabulary and internalizing symptoms was estimated to assess whether or not there was an association before the addition of potential confounding variables. Subsequently, to determine whether there was a relation between age 5 vocabulary and adolescent internalizing symptoms after adjusting for SEC and childhood psychosocial variables, the following nested models were estimated for both BCS1970 and MCS2001 data. Biological and sociodemographic factors were added in a first

model. Mother and child psychosocial variables were then added in a second model. The vocabulary predictor was added to a third model, and quadratic and cubic terms were added to the vocabulary predictor in a fourth model to test for any non-linearities. Regression estimates were pooled based on Rubin's rules (Rubin, 1984). Mean centering was carried out for all continuous variables for all analyses.

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

Sensitivity and supplementary analyses

In order to assess whether the ethnic make-up of our selected samples could be driving any observed effects, we ran two sensitivity analyses: 1). BCS1970: age 5 vocabulary predicting age 16 internalizing symptoms: White, English-speaking sample; and 2). MCS2001: age 5 vocabulary predicting age 14 internalizing symptoms: White, English-speaking sample, with matching potential confounding variables to the BCS1970 analysis. In order to assess whether the different items tapping internalising symptoms for the two cohorts could be driving any differences, we ran a third sensitivity analysis using a harmonised matched subset of items from the self-reported internalising symptoms subscales from the BCS1970 and MCS2001. These can be found in Appendix 1.

In order to allow comparison with existing literature, we also carried out 3 supplementary analyses: 1) BCS1970: age 5 vocabulary predicting age 16 *parent-reported* internalizing symptoms; 2) MCS2001: age 5 vocabulary predicting age 14 *parent-reported* internalizing symptoms; 3) BCS1970: age 5 vocabulary predicting *age 34* internalizing symptoms. There are multiple potential reporters for adolescent mental health. Rates of agreement between parent and self-reported symptoms of adolescent internalizing symptoms are known to be low (Rescorla et al, 2013) and studies to date have varied in the measure used. Analyses 1 and 2 were therefore carried out in order to assess whether the size and

direction of any associations differed as a function of reporter. Analysis 3 was completed to complement our adolescent analyses - to see whether or not any relation persisted into adulthood in the BCS1970 sample, allowing for comparison with Schoon et al, (2010). Main findings for each can be found below, full details and results can be found in Appendix 1.

3.4. Results

Descriptive statistics

Descriptive statistics can be found in Table 3.1. These were estimated across 25 imputed datasets.

For the BCS1970, differences in the proportions between the full cohort sample and the analytical sample in this paper (i.e., everyone with a vocabulary score at the age of 5) are negligible (see Table 1, Appendix 1, Section 1). For the MCS2001, differences between the full cohort sample and the selected analytical sample are also negligible for most variables (see Table 2, Appendix 1, Section 1). However, there are more unemployed parents in the full sample compared to the analytical sample.

As expected, based on demographic trends in the UK, there were more White ethnicity participants in the BCS1970 compared to the MCS2001 and more parents in the MCS2001 had university level qualifications (higher degree, first degree, diploma in education; 38.96%) compared to BCS1970 (first degree, postgraduate degree, national diploma or certificate, membership of a professional institution, city and guilds full technical certificate, certificate of education, state registered nurse; 16.84%).

Table 3.1. Mean (SD), proportions (%) and 95% confidence intervals for analysis samples in the BCS1970 and MCS2001 cohorts

BCS1970: variable names	BCS1970 Mean (SD) or % [95%CI] *N=11,640*	MCS2001: variable names	MCS2001 Mean (SD) or % [95%CI] *N=14,574*
Mental health		***Mental Health***	
Age 16 self-reported internalizing symptoms	3.77(2.20) [3.73;3.81]	Age 14 self-reported internalizing symptoms	5.62(5.90) [5.53; 5.72]
Age 16 parent-reported internalizing symptoms	1.95(1.89) [1.92; 1.99]	Age 14 parent-reported internalizing symptoms	1.98(2.12) [1.95; 2.02]
Language		***Language***	
Age 5 language score (EPVT)	35.32(10.81) [35.12; 35.51]	Age 5 language (naming vocabulary)	109.21(15.61) [108.95;109.46]
Age of CM at the time of vocabulary test (months)	60.92(1.28) [60.89;60.94]	Age of CM at time of vocabulary test (months)	62.51(2.92) [62.46;62.56]
Biological Risk Variables		***Biological risk variables***	
Birthweight (g)	3331.09(517.01)	Birthweight (g)	3388.62(1959.92)

	[3321.70;3340.48]		[3357.00;3420.25]
Gestational age (days)	281.86(16.44) [281.56; 282.16]	Gestational age (days)	276.13(13.44) [275.91;276.35]
Sex (female)	48.21%	Sex (female)	48.93%
Ethnicity (White UK)	96.22%	Ethnicity (White)	88.61%
Ethnicity (Minority)	3.78%	Ethnicity (mixed)	2.91%
		Ethnicity (Indian)	1.78%
		Ethnicity (Pakistani/Bangladeshi)	3.47%
		Ethnicity (Black/Black British)	2.22%
		Ethnicity (Other, including Chinese)	1%
Sociodemographic variables		***Sociodemographic variables***	
Language used in home (English)	97.63%	Main language in home (English)	90.20%
Language used in home (other than English)	2.37%	Main language in home (English and another language)	7.86%
		Main language in home (only another language)	1.94%
Occupation (Professional & Managerial)	20.95%	Occupation (NS-SEC Higher managerial)	46.21%
Occupation (skilled manual & non-manual)	61.75%	Occupation (NS-SEC intermediate)	18.68%
Occupation (semi-skilled/unskilled)	16.33%	Occupation (NS-SEC Routine& Manual)	26.17%
Occupation(unemployed)	0.97%	Occupation(unemployed)	8.95%
Parental education (Degree +)	13.7	Parent education (higher degree)	7.64%
Parental education (Certificate of education)	1.66%	Parent education (first degree)	19.15%
Parent education (SRN (state registered nurse)	1.67%	Parent education (diploma in higher education)	12.69%
Parent education (A levels)	7.88%	Parent education (A levels)	10.05%
Parent education (O level)	21.18%	Parent education (O levels/GCSE grades A-C)	33.05%
Parent education (Vocational qualification)	13.2%	Parent education (GCSE grades D-G)	7.36%
Parent education: no qualifications	40.71%	Parent education (none of these/other including overseas)	10.07%
		Income (lowest quintile)	17.34%
		Income (second quintile)	18.87%
		Income (third quintile)	20.05%
		Income (fourth quintile)	21.60%
		Income (highest quintile)	22.14%
		Total net wealth	193416.77(518334.2) [185052.83;201780.72]
Childhood psychosocial variables		***Childhood psychosocial variables***	
Teen mum (yes)	8.85%	Teen mum (yes)	5.89%

Marital status (not partnered)	5.37%	Marital status (not partnered)	32.96%
Maternal depression (CM age 5)	4.32(3.63) [4.25;4.39]	Maternal depression (CM age 5)	3.14(3.77) [3.08; 3.20]
Age 5 CM externalizing difficulties	1.8(1.65) [1.77; 1.83]	Age 5 CM externalizing difficulties	1.23(1.56) [1.21;1.26]
Age 5 CM internalizing difficulties	1.5(1.5) [1.48; 1.53]	Age 5 CM internalizing difficulties	1.34(1.56) [1.31;1.36]

Means, SDs, proportions and 95% CIs are pooled across 25 imputed datasets (both cohorts) and are sample weighted (MCS2001 only)

Does age 5 vocabulary predict age 16 internalizing symptoms in the BCS1970 (born 1970)?

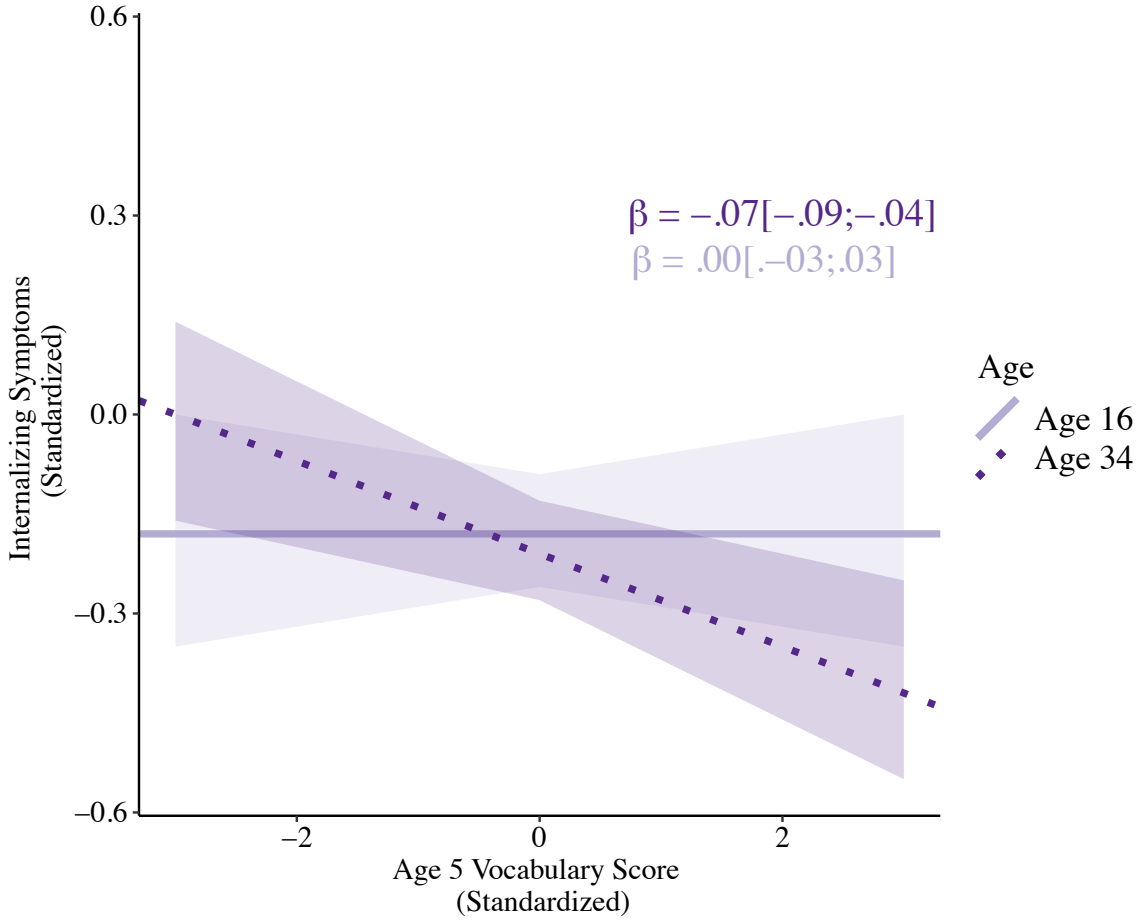
In an unadjusted model (i.e., not including any potential confounding variables), there was a significant negative relation between vocabulary and self-reported internalizing symptoms such that higher vocabulary scores were associated with fewer internalising symptoms (β= -.03 [-.06; -.01]; see Table 3.3). To test whether this relation held when potential confounding factors were included, we first tested whether two sets of potential confounding variables predicted internalising symptoms and then whether vocabulary explained variance over and above these variables. Compared to a model with no predictors, a model with biological and SEC control variables significantly improved the model fit, (Dm(14, 1336.91)=10.57, p<.001) (see Table 3 Appendix 1, Section 2). Compared to a model with only these variables, a model that also included mother and child psychosocial variables had a significantly improved fit, (Dm(5, 374.65)=5.24, p=.001). Compared to a fully adjusted model, adding receptive vocabulary scores (model 3) did not significantly improve the model fit, (Dm(2, 138.26)=0.04, p=.965; see Table 3.2). We examined a model with quadratic and cubic terms which did not improve the model fit, (Dm(2, 226.84)= 0.34 p=.710), suggesting the absence of non-linear relations between age 5 vocabulary and age 16 internalizing symptoms. These results suggest that age 5 vocabulary does not predict any unique variance in age 16 self-reported internalizing symptoms in this cohort after accounting for potential confounding variables. Given the unadjusted relation between age 5 vocabulary and age 16 internalizing symptoms, we ran a post-hoc analysis whereby we added the vocabulary predictor to a model containing biological and SEC variables. This was to check which potential confounder(s) removed the relation. Compared to a model with only biological and SEC variables, adding receptive vocabulary scores did not significantly improve the model fit, (Dm(2, 137.73)=0.28, p=.756). Sex was the only significant predictor in this model (β=.31[.25;.37], see Table 4, Appendix 1, Section 2), and we therefore conclude that the

appearance of a relation between vocabulary and internalizing symptoms in the unadjusted analysis is spurious and due to the co-linearity of sex and vocabulary size in this cohort (see Patalay & Fitzsimons, 2018; Stolarova et al., 2016 for further evidence that both mental health and vocabulary show sex differences).

Sensitivity analysis 1 (restricting the analysis to a White, English-speaking sub-sample) revealed a similar pattern of results (see Appendix 1). However, a different pattern of results was observed in supplementary analysis 1, which considered *parent-reported* adolescent internalizing symptoms as the outcome variable. In a fully adjusted model, this analysis revealed a significant *negative* relation between age 5 vocabulary and parent-reported adolescent internalizing symptoms (β= -.06[-.08;-.03]; see Table 3.3). This suggests that lower vocabulary scores in childhood were predictive of more parent-reported internalizing symptoms in adolescence (Table 10, Appendix 1, Section 5). A sensitivity analysis with parent-reported symptoms that considered only White, English-speaking cohort members yielded a similar pattern of results.

Measures of self-reported internalizing symptoms are also available in the adulthood sweeps of the BCS1970. In order to investigate the longer-term effects of age 5 vocabulary on internalizing symptoms, we ran an analysis with age 34 internalizing symptoms as the outcome variable, extending the findings of Schoon et al (2010), by considering vocabulary as a continuous predictor of age 34 internalizing symptoms. In a fully adjusted model, this analysis revealed a significant *negative* relation between age 5 vocabulary and age 34 internalizing symptoms (β=-.07[-.09;-.04]; see Appendix 1, Section 9), such that that those with lower vocabulary scores in childhood self-reported more internalizing symptoms in adulthood. This differs from the findings of the pre-registered analysis with age 16 self-reported symptoms as the outcome (see Figure 3.1).

Figure 3.1. Age 34 and Age 16 internalizing symptoms, predicted from age 5 vocabulary score.



Note: The scale of the standardised vocabulary measure ranges from −1.97 (5th percentile) to 1.45 (95th percentile). The age 16 internalising symptoms measure standardised scale ranges from −1.71 (5th percentile) to 1.92 (95th percentile). The age 34 internalising symptoms measure standardised scale ranges from −1.16 (5th percentile) 1.75 (95th percentile).

Does age 5 vocabulary predict age 14 internalizing symptoms in the MCS2001 (born ~2001)?

In an unadjusted model, a significant *positive* relation between vocabulary and self-reported internalizing symptoms was observed such that higher vocabulary scores were associated with *more* internalizing symptoms (β= .03 [.01; .06]; see Table 3.3). To test whether this relation held when control factors were included, we first tested whether two sets of control variables predicted internalising symptoms in the MCS2001 and then whether vocabulary explained variance over and above these variables. Compared to a model with no predictors, a model with biological and sociodemographic control variables significantly improved the model fit, (Dm(24, 4312.52)= 36.69, p<.001). Compared to a model with only

these variables, a model that also adjusted for mother and child psychosocial variables gave a significantly improved fit, (Dm(5, 429.41)=14.43, *p*<.001) (see Table 5 Appendix 1, Section 2). Compared to a fully adjusted model, adding expressive vocabulary scores in model 3 accounted for significantly more variance in the outcome, (Dm(2, 331.61)=9.93, *p*<.001; see Table 3.2), such that higher vocabulary scores were associated with more internalizing difficulties in adolescence. We examined a model with quadratic and cubic terms which did not improve the model fit, (Dm(2, 366.39)= 0.03, *p*=.975), suggesting the absence of non-linear relations between age 5 vocabulary and age 14 internalizing symptoms. In sum, for children born in 2000-02, age 5 vocabulary ability explains some unique variance in age 14 internalizing symptoms, such that better childhood vocabulary ability predicts *poorer* adolescent internalizing symptoms. The effect size of the vocabulary predictor (β = .05[.02;.07]) indicates that a 1 SD increase in vocabulary was associated with an increase of 5% of a standard deviation in internalizing symptoms. Despite being small in size, this effect is of comparable magnitude to other predictors in the model (maternal mental health (β =.07[.04;.09]), childhood externalizing symptoms (β=.05[.02;.07]) and is larger than that of childhood parent-rated internalizing symptoms (β=.02[-.00;.04]); see Table 5, Appendix 1, Section 2).

Table 3.2. β [95% CIs] for fully adjusted regression models (final model) for BCS1970 sample (N=11,640) and MCS2001 sample (N=14,574)

BCS1970: variable names	BCS1970 β [95% CI] ***Partial R^2*** *N=11,640*	MCS2001: variable names	MCS2001 β [95% CI] ***Partial R^2*** *N=14,574*
Biological Risk variables		***Biological risk variables***	
Birthweight (g)	-.02[-.05;.01], *p=.26* ***0.0002***	Birthweight (g)	.00[-.03;.04] *p=.81* ***0***
Gestational age (days)	.00[-.03;.03], *p=.99* ***0***	Gestational age (days)	.01[-.01;.03] *p=.49* ***0***
Sex (male)	REFERENCE	Sex (male)	REFERENCE
Sex (female)	.32[.26;.38]*, *p=.00* ***0.0247***	Sex (female)	.54[.50;.58]** *p<.001* ***.0728***
Ethnicity (White)	REFERENCE	Ethnicity (White)	REFERENCE
Ethnicity (Minority)	.05[-.11;.20], *p=.56* ***0***	Ethnicity (mixed)	.02[-.09;.12] *p=.76* ***.0008***
		Ethnicity (Indian)	-.11[-.27;.04] *p=.15*
		Ethnicity	-.21[-.34;-.08]**

	(Pakistani/Bangladeshi) *p*<.001
	Ethnicity (Black/Black British) -.12[-.25;.01] *p=.06*
	Ethnicity (Other, including Chinese) -.11[-.31;.09] *p=.28*
Sociodemographic variables	***Sociodemographic variables***
Language used in home (English) REFERENCE	Main language in home (English) REFERENCE
Language used in home (other than English) .01[-.17;.19], *p=.95* ***0***	Main language in home (English and another language) .01[-.09;.11] *p=.83* ***0***
	Main language in home (only another language) .01[-.16;.17] *p=.93*
Occupation (Professional & Managerial) REFERENCE	Occupation (Higher managerial) REFERENCE
Occupation (skilled manual & non-manual) .00[-.06;.07], *p=.96* ***0.0001***	Occupation (intermediate) .01[-.05;.07] *p=.76* ***0***
Occupation (semi-skilled/unskilled) -.01[-.10;.08], *p=.84*	Occupation (Routine& Manual) .02[-.04;.08] *p=.49*
Occupation(unemployed) .10[-.21;.40], *p=.53*	Occupation(unemployed) .03[-.06;.13] *p=.46*
Parental education (Degree +) REFERENCE	Parent education (higher degree) REFERENCE
Parental education (Certificate of education) .07[-.12;.25], *p=.48* ***0.0002***	Parent education (first degree) -.05[-.13;.02] *p=.18* ***.0001***
Parent education (SRN (state registered nurse) -.01[-.20;.18], *p=.91*	Parent education (diploma in higher education) -.05[-.13;.04] *p=.30*
Parent education (A levels) .01[-.09;.11], *p=.83*	Parent education (A levels) -.04[-.13;.05] *p=.37*
Parent education (O level) .02[-.07;.11], *p=.62*	Parent education (O levels/GCSE grades A-C) -.06[-.14;.03] *p=.17*
Parent education (Vocational qualification) -.03[-.14;.08], *p=.59*	Parent education (GCSE grades D-G) -.02[-.13;.10] *p=.79*
Parent education: no qualifications .02[-.06;.11], *p=61*	Parent education (none of these/other including overseas) -.05[-.16;.05] *p=.33*
	Income (lowest quintile) REFERENCE
	Income (second quintile) .03[-.04;.10] *p=*.41 ***.0006***
	Income (third quintile) -.03[-.10;.04] *p=.43*
	Income (fourth quintile) -.04[-.11;.04] *p=.34*
	Income (highest quintile) -.08[-.15;.00] *p=.05*
	Total net wealth -.02[-.04;-.00]* *p=.03* ***.0004***
Childhood psychosocial variables	***Childhood psychosocial variables***
Teen mum (no) REFERENCE	Teen mum (no) REFERENCE
Teen mum (yes) .10[.01;.19]*, *p=.03*	Teen mum (yes) .01[-.12;.14] *p=.88*

	0.0007		***0***
Marital status (partnered)	REFERENCE	Marital status (partnered)	REFERENCE
Marital status (not partnered)	.01[-.12;.14], *p=.92* ***0***	Marital status (not partnered)	.07[.02;.13]** *p<.001* ***.001***
Maternal depression (CM age 5)	.02[-.01;.05], *p=.29* ***0.0001***	Maternal depression (CM age 5)	.07[.04;.09]** *p<.001* ***.0038***
Age 5 CM externalizing difficulties	.04[.01;.07]*, *p=.01* ***0.0013***	Age 5 CM externalizing difficulties	.05[.02;.07]** *p<.001* ***.0016***
Age 5 CM internalizing difficulties	.05[.02;.08]**, *p=.00* ***0.0025***	Age 5 CM internalizing difficulties	.02[-.00;.04] *p=.07* ***.0003***
Age 5 language		***Age 5 language***	
Age of CM at time of taking vocabulary test (months)	.00[-.03;.03], *p=.82* ***0***	Age of CM at time of taking vocabulary test (months)	.01[-.01;.03] *p*=.33 ***0***
Age 5 vocabulary	-.00[-.03;.03], *p=.94* ***0***	Age 5 vocabulary	.05[.02;.07]** *p<.001* ***.0017***
R^2	0.0332	**R^2**	0.0893

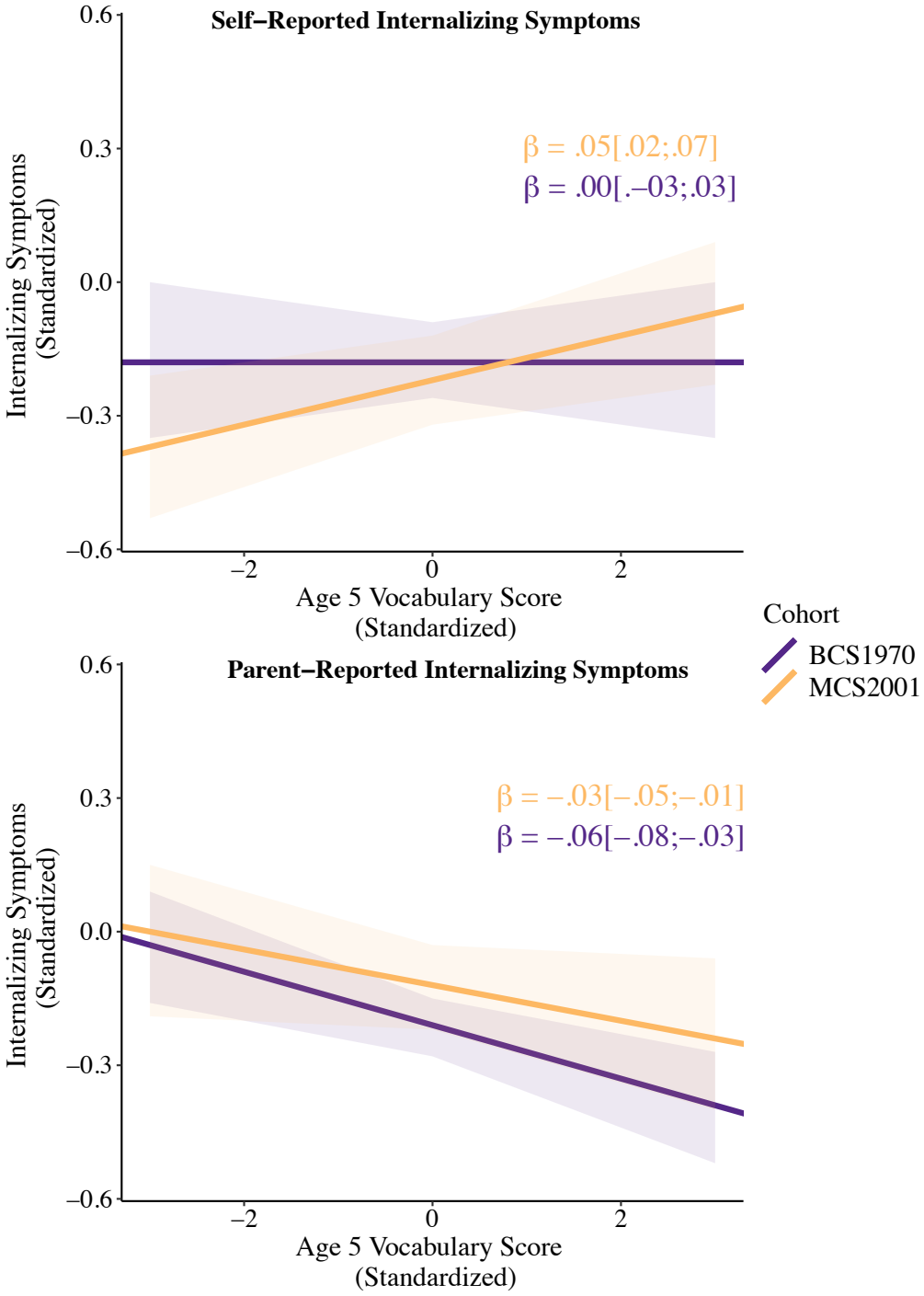
* *p* <.05, ** *p* <.01. These are the results for the final model (model 3) in the hierarchical regression. Please see Appendix 1, section 2 for the full regression model tables.

A sensitivity analysis restricted to a White, English-speaking sub-sample, with matched BCS1970 potential confounding variables revealed that a model with the vocabulary predictor was a significantly better fit than a model including only the biological, SEC, mother and childhood psychosocial variables: better vocabulary scores in childhood were associated with more self-reported internalizing symptoms in adolescence (see Appendix 1).

Supplementary analysis 2 considered *parent-reported* adolescent internalizing symptoms as the outcome variable and here there was a significant *negative* relation between age 5 vocabulary and adolescent internalizing symptoms, in a fully adjusted model such that better vocabulary scores in childhood were predictive of *fewer* parent-reported internalizing symptoms in adolescence (β= -.03 [-.05 to -.01]; see Table 3.3 and Table 12, Appendix 1, Section 7).

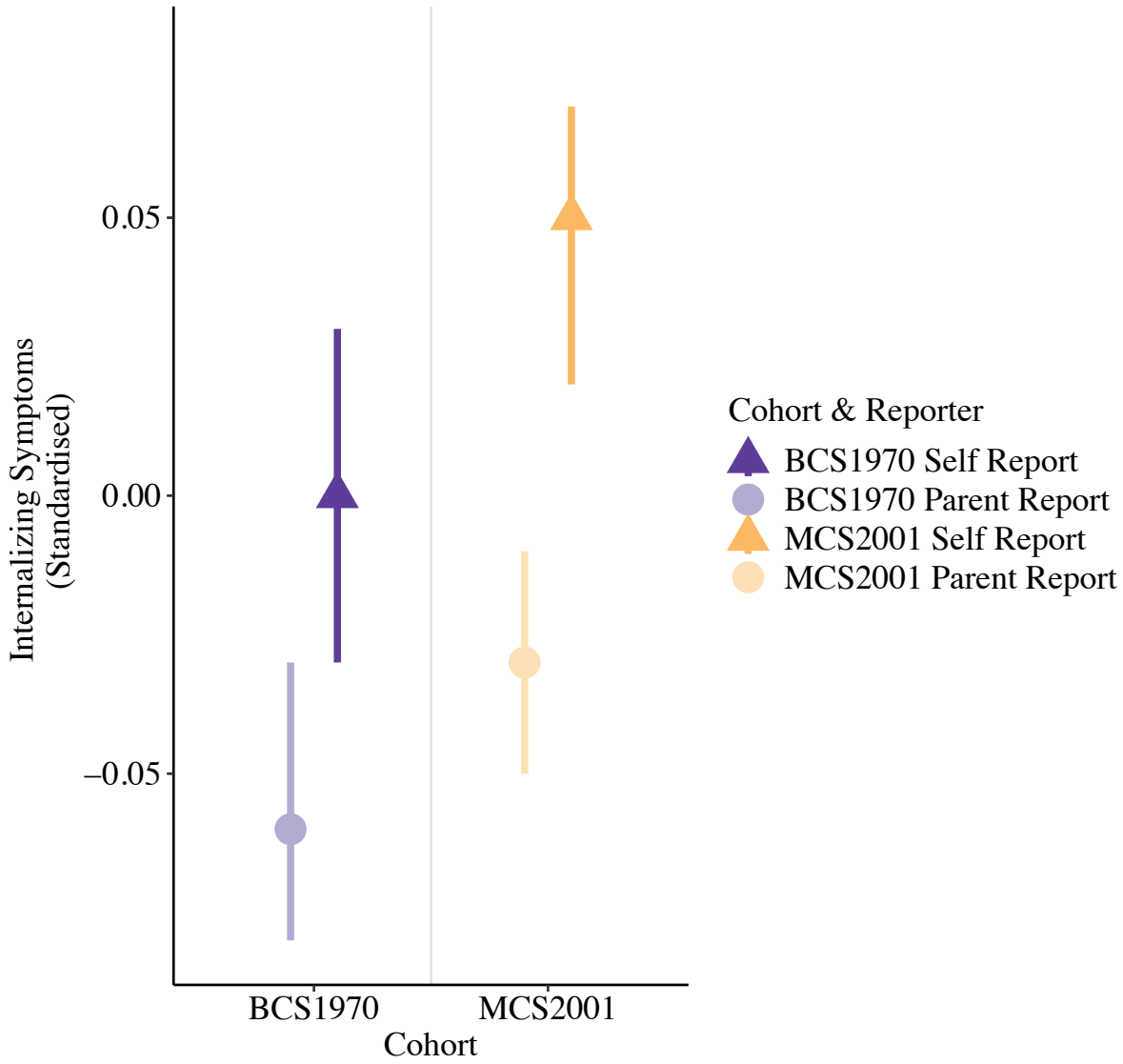
Overall, switching from self-report to parent-report of adolescent internalizing symptoms changes the direction of effect such that good early vocabulary predicts *fewer* internalizing symptoms in both cohorts (see Figure 3.2). As can be seen from Figure 3.3, the self- and parent-reported internalising symptoms measures are significantly different from one another (the confidence intervals for each do not overlap).

Figure 3.2 Adolescent self and parent reported internalizing symptoms, predicted from age 5 vocabulary score.



Note: For BCS1970 data, the scale of the standardised vocabulary measure ranges from −1.97 (5th percentile) to 1.45 (95th percentile). The self reported internalising measure standardised scale ranges from −1.71 (5th percentile) to 1.92 (95th percentile). The parent reported internalising symptoms measure standardised scale ranges from −1.03(5th percentile) to 2.13 (95th percentile). For MCS2001 data, the scale of the standardised vocabulary measure ranges from −1.79 (5th percentile) to 1.44 (95th percentile). The self reported internalising measure standardised scale ranges from −0.95 (5th percentile) to 2.12 (95th percentile). The parent reported internalising symptoms measure standardised scale ranges from −0.96 (5th percentile) to 1.83 (95th percentile).

Figure 3.3 β [95% CIs] for Internalizing Symptoms (Self and Parent Reported) in BCS1970 and MCS2001



Exploratory analyses

The following post-hoc, exploratory analyses were conducted to better understand the above results in the context of the broader literature. Model comparisons and tables for exploratory analyses can be found in Appendix 1, Section 12.

1. Age 5 vocabulary as a binary predictor

Schoon et al (2010) analysed the BCS1970 and found that vocabulary difficulties at age 5 (a dichotomous variable, where a difficulty was defined as vocabulary 1 standard deviation below the mean) were associated with poor mental health at age 34. Along with studies of DLD, Schoon et al's (2010) findings led us to predict that vocabulary ability across the full continuum would be negatively associated with internalizing symptoms in the general

population in adolescence. However, the main pre-registered results suggest that this is not the case when the full continuum of vocabulary ability is considered. To test whether the predicted association holds in adolescence when a dichotomised vocabulary predictor (vocabulary difficulty or not) is used, we ran two further models with data from the BCS1970 and the MCS2001. Absence of vocabulary difficulties was used as the reference category. Models were built in the same way as the main analyses. The vocabulary predictor was dichotomised at 1 SD below the mean, in line with the methodology of Schoon et al (2010). However, some research has classified language difficulty as 1.5 SD below the mean (Norbury et al., 2016), and we therefore also dichotomised the vocabulary predictor using this cut off as a sensitivity analysis (see Appendix 1, Section 13).

In the BCS1970 sample, 1872 cohort members (16% of the cohort) had vocabulary scores 1 SD below the mean. Results for this analysis can be found in Table 18 (Appendix 1, Section 12). There was no relation between vocabulary and internalizing symptoms when vocabulary was considered as a binary predictor in a fully adjusted model (β= -.02 [-.10;.05]; see Table 3.3). This remained the case when the more stringent cut off of 1.5SD below the mean was used (1114 cohort members (10%) had scores 1.5SD below the mean (See Appendix 1, Section 13).

In the MCS2001 sample, 2919 cohort members (20%) had vocabulary scores 1 SD below the mean. Results for this analysis can be found in Table 19 (Appendix 1, Section 12). There was no significant relation between age 5 vocabulary and adolescent internalizing symptoms in a fully adjusted model when this cut off was considered (β= -.05[-.01;.01]; see Table 3.3). However, when the more stringent cut off of 1.5SD below the mean was used, there was a significant *negative* relation between age 5 binary vocabulary and age 14 internalizing symptoms (β=-.14[-.24;.-.05]. See Appendix 1, Section 13). This suggests that poor vocabulary was predictive of *fewer* internalizing symptoms at age 14. While the effect size was again small, this unexpected direction of effect is consistent with the outcome of the main pre-registered analyses reported above. When using this cut off, 8% of MCS2001 cohort members were classed as having vocabulary difficulties (1204 cohort members). This maps on to national prevalence levels for DLD, which are estimated to be around 7.5% (Norbury et al, 2016).

Binary internalizing symptoms as the outcome variable

In our main pre-registered analyses, we found that for BCS1970 cohort members, there was no relation between age 5 vocabulary and internalizing symptoms. For MCS2001 cohort members, there was a positive relation between age 5 vocabulary and self-reported internalizing symptoms. Therefore, in a second set of exploratory analyses, we investigated whether these trends remained when we considered those with clinical levels of internalizing symptoms, (scores ≥4 on the Malaise inventory in the BCS1970 and scores ≥12 on the SMFQ in the MCS2001), with binary logistic regressions, whereby 0 = non-clinical levels and 1 = clinical levels of internalizing symptoms.

In the BCS1970 sample, 4188 cohort members had scores ≥4 on the Malaise inventory. This analysis revealed that the odds of having clinical levels of internalizing symptoms in adolescence did not differ as a function of vocabulary (see Table 3.3). This finding is in line with the main pre-registered analysis for the BCS1970, which also suggests no relation between early vocabulary and the continuous internalizing symptoms measure.

In the MCS2001 sample, 2013 cohort members had scores ≥12 on the SMFQ. This analysis revealed that for every SD unit increase in vocabulary, there was a 16% increase in the odds of having clinical levels of internalizing symptoms (OR=1.16[1.07;1.25], see Table 3 & Table 21, Appendix 1, Section 12). This is in line with the finding of the main pre-registered analysis, whereby MCS2001 cohort members with better vocabulary in childhood were found to have more internalizing symptoms in adolescence. However, it is worth noting that compared to a model with all potential confounding variables, adding expressive vocabulary scores did not significantly improve the model fit (see Appendix 1, Section 12).

Table 3.3. Main exposure (age 5 language) – outcome (internalizing symptoms) relationship in each analysis

	Analysis details	BCS1970 Age 5 language (β Coef [95% CI], *p* value)	MCS2001 Age 5 language (β Coef [95% CI], *p* value)
Self-report: unadjusted	Age 5 language predicting self-reported adolescent mental health, unadjusted model.	-.03 [-.06; -.01]**, *p<.001*	.03 [.01; .06]** *p<.001*
Self-report: fully adjusted	Main, pre-registered analysis. Age 5 language ability predicting adolescent mental health, after adjusting for demographic,	.00[-.03;.03], *p=.94*	.05[.02;.07]**, *p<.001*

	socioeconomic and childhood indicators of poor mental health		
White subset	Analysis conducted as a sensitivity check. Model built in the same way as the main analysis, this time considering only English-speaking cohort members of a White ethnicity.	.01[-.03;.04], *p=.72*	.03[.00;.05]*, *p=.02*
Parent-report	Supplementary analysis. Models built in the same way as the main analysis. Here, the outcome variable was parent-reported mental health, to assess whether the main analysis results changed when considering a different perspective.	-.06[-.08;-.03]**, *p<.001*	-.03[-.05;-.01]** *p<.001*
White subset (parent-report)	Analysis conducted as a sensitivity check. Model built in the same way as the main analysis, this time considering only English-speaking cohort members of a White ethnicity.	-.06[-.08;-.03]** *p<.001*	-.04[-.07;-.02]**, *p<.001*
Exploratory: binary language those scoring 1SD below the mean defined as language difficulties (0=normal language, 1 = language difficulties)	Exploratory analysis 1. Models built in the same way as the main analysis. This time language considered as a binary predictor, to consider those specifically with language difficulties.	-.02 [-.10;.05], *p=.54*	-.04[-.09;.02], *p=.22*
Sensitivity: binary language those scoring 1.5 SD below the mean defined as language difficulties (0=normal language, 1 = language difficulties)	Sensitivity check with more stringent cut off point. Models built in the same way as the main analysis.	-.03[-.15;.08], *p= .59*	-.11[-.21;-.02]**, *p<.001*
Exploratory 2: binary internalizing symptoms OR [95% Cis]	Exploratory analysis 2. Internalizing symptoms outcome considered as binary.	1.00[.93;1.08], *p=.95*	1.16[1.07;1.25]**, *p<.001*

Sensitivity: complete case for language and internalizing symptoms	The main analysis considered all of those cohort members with just complete scores for the language predictor. However, due to debate around whether or not the outcome variable should be imputed, we also conducted a sensitivity check whereby we considered complete cases for language and the mental health outcome.	.00[-.03;.03], *p=.91*	.05[.03;.08]**, *p<.001*
Age 34 mental health as the outcome variable	Supplementary analysis. Models built in the same way as the main pre-registered analysis	-.07[-.09;-.04]**, *p<.001*	

Binary language [a]: coefficient for poor language, normal language = reference group.
β Coefficients for the White subset, binary language and parent-reported outcome models are taken from the fully adjusted models (see Appendix 1)
* $p < .05$, ** $p < .01$.

3.5. Discussion

In pre-registered analyses, we assessed whether early vocabulary in the general population is associated with self-reported internalizing symptoms in adolescence. The overall finding was that in a cohort of children born in 1970, there was no significant relation between early vocabulary and self-reported adolescent internalizing symptoms once a comprehensive set of potential confounding variables was included. However, a supplementary analysis revealed that a relation emerged in adulthood, such that better early vocabulary predicted fewer self-reported adult internalizing symptoms. This finding was in line with Schoon et al. (2010). Conversely, in the more recently-born MCS2001 children (born ~2001), better early vocabulary predicted *worse* self-reported adolescent internalizing symptoms, an effect that remained in a fully adjusted model. In general, findings for both cohorts did not differ when vocabulary or internalizing symptoms measures were treated as dichotomous measures. Overall, our results suggest that the relation between early vocabulary and self-reported adolescent internalizing symptoms varies by generation in the UK.

Given the low rates of agreement between self report and parent report (correlations of typically ~0.2; Rescorla et al, 2013; Patalay & Fitzsimons, 2018), we investigated whether or not the outcomes of the pre-registered analyses (which focussed on self-reported symptoms) differed when parent-reported symptoms were considered as the outcome

variable. Across cohorts, parents tended to report fewer internalizing symptoms if their child had better language early on. This finding is in line with St Clair et al., (2019) and a recent meta-analysis (which did not differentiate studies on the basis of reporter; see Hentges et al, 2021). Thus, for the MCS2001, the direction of effect reversed when parent reports were considered instead of self reports. Similar trends have been noted in the literature. For instance, a socioeconomic gradient is observed in parent-reported child mental health, but not in child-reported mental health (Johnston et al., 2014). In contrast, no significant differences by ethnic group are observed at age 14 in MCS2001 cohort members based on parent report, but substantial ethnic differences are observed based on self report (Patalay & Fitzsimons, 2018). Some previous research looking at this relation with general population samples has used an outcome measure where self- and parent-reported internalizing symptoms have been combined into one overall measure (Miller et al, 2020; Bornstein et al, 2013). However, it is unclear what a combined measure of self- and parent-reported symptoms represents given their low correlation. The current findings, where the direction of effect differs as a function of reporter, suggest that studies with a combined outcome measure should be interpreted with caution.

Given the differences in direction of effect as a function of reporter, an important question is whether one reporter is more reliable for identifying internalizing symptoms in adolescence. There are strong arguments to be made for self-report measures. First, the socio-demographic patterns of self report better match the latest national prevalence estimates based on clinical diagnoses, which are arguably the gold standard. This suggests that they better reflect population patterns in diagnosed mental health difficulties, compared to symptoms reported by others (Sadler et al., 2018). Second, generally speaking, young people are competent reporters of their own mental health (Sharp et al., 2006) and it is argued that they should be considered the primary reporter when assessing internalizing mental health (The Children's Society, 2019). Third, from a longitudinal perspective, self-report measures allow direct comparison with adult outcomes in the BCS1970 (for example, Schoon et al, 2010), where only self-reported measures are available, which is the norm in research on adult mental health. For these reasons, we pre-registered the self-report measure as the primary outcome and, while the current direction of findings for the recent cohort is surprising, we consider it important to take seriously the possibility that good vocabulary is not straightforwardly predictive of good mental health.

The finding that better childhood abilities predict more internalizing symptoms in adolescence in the MCS2001 is counterintuitive, and there are a number of possible

explanations for this direction of effect. For example, academic pressure may have increased in recent years: schoolwork, examinations and feeling pressured are commonly reported stressors among adolescents (Gray et al., 2011). It is possible that language ability is positively associated with pressure to succeed academically, resulting in adolescents with more advanced language abilities having a higher risk of feeling stressed and experiencing poor mental health. Adding to possible increases in academic pressure is the widening of social and generational inequalities in Britain over this period (Corlett et al., 2019), which increases the importance of academic qualifications in achieving economic stability in adulthood (Green et al., 2020).

Limitations and strengths.

This research used vocabulary as the sole measure of language ability. As a result, we might not have captured the kinds of language problems that lead to mental health difficulties. Recent research suggests that different measures of formal language tend to load on to the same factor (Fricke et al., 2017), so vocabulary is likely a good proxy for broader language ability. However, there is some evidence that pragmatic language skills cluster separately (Wilson & Bishop, 2019) and that they might be more directly related to mental health (Brenne & Rimehaug, 2019; Ketelaars et al., 2010). Likewise, we focused only on one domain of mental ill-health - internalizing symptoms - and it might be that a different picture would emerge if externalizing symptoms were analysed (for example, Chow & Wehby, 2018).

While receptive and expressive vocabulary tend to be moderately to highly correlated (e.g. Conway et al, 2017), the difference in self-reported findings between cohorts could be attributed to the use of a receptive vocabulary measure in the BCS1970 and an expressive vocabulary measure in the MCS2001. However, one would have thought that if the difference could be attributed to the use of different vocabulary measures, a similar cohort difference would emerge for the parent-reported outcome, which was not the case.

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

As with any longitudinal data analysis, missing data had to be accounted for. Those with mental health difficulties in one sweep of cohort studies are less likely to take part in the next sweep. Furthermore, males, particularly of a lower SEC, tend to be underrepresented in

subsequent sweeps (Elliott & Shepherd, 2006; Mostafa & Wiggins, 2014). Therefore, missing data could introduce bias into the results. To combat this, missing data were accounted for using multiple imputations, which are considered a "best effort" approach (Little & Rubin, 2019). Although we have aimed to capture the full continuum of vocabulary abilities, higher rates of attrition may occur among those with language difficulties, and it is therefore possible that our results underestimate effects. However, we have imputed missing data to minimise bias due to attrition.

Finally, as with any study, it is likely that some measurement error is present (Van Smeden et al., 2020). However, we have no reason to expect any differential or multivariate error for our variables, or to expect large amounts of non-differential error for the standardised, reliable measures we make use of.

Despite these limitations, the strengths of this research lie in the large and nationally representative samples with researcher-collected vocabulary measures, that make it possible to test the association between early vocabulary and later internalizing symptoms, while taking into account important control variables. As such, findings are generalisable to the United Kingdom. However, our cross-cohort comparison revealed that this relation has changed between generations. It could therefore also differ as a function of cultural and socio-economic conditions across the globe, meaning our findings for contemporary adolescents may not be generalisable beyond the UK.

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

Implications.

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

well founded for educational reasons (e.g., Fricke et al, 2017), caution is needed when looking for means to improve adolescent internalizing mental health.

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

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

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

3.6. Conclusion

The use of two cohort studies enabled us to test whether there is an association between early vocabulary and adolescent internalizing mental health, and if so, how any relation may have changed over 30 years. In the BCS1970, no relation was observed for self-reported adolescent internalizing mental health once controls were accounted for. In the contemporary generation, MCS2001 data indicate that, if anything, better childhood vocabulary predicts *poorer* self-reported adolescent internalizing mental health, regardless of whether vocabulary was considered as a continuous or binary predictor. Thus, the relation between age 5 vocabulary ability and adolescent internalizing symptoms varies with generation. When parent-reported adolescent symptoms were considered, lower childhood vocabulary scores predicted *poorer* adolescent internalizing mental health in both cohorts. Therefore, the relation also varies as a function of reporter. In all analyses effect sizes were small. In sum, the relation between childhood vocabulary and adolescent internalizing

symptoms varies by generation and reporter – good early language skills may not be protective for contemporary adolescents' internalizing mental health.

Chapter 4 : Tracking the relation between different dimensions of socio-economic circumstance and vocabulary across developmental and historical time

This chapter forms the basis of a submitted manuscript, which is available as a pre-print Thornton, E., Matthews, D., Patalay, P., & Bannard, C. (2021, August 13). Tracking the relation between different dimensions of socio-economic circumstance and vocabulary across developmental and historical time. https://doi.org/10.31234/osf.io/bu3px

4.1. Abstract

Despite policy-makers seeking interventions to address them, social inequalities in child vocabulary persist. To make better targeted progress, fundamental questions must be addressed: which aspects of socioeconomic circumstances (SEC) best predict vocabulary? Is this stable over developmental and historical time? Data from two large, nationally-representative (for sex, ethnicity, and SEC) datasets were analysed: the 1970 British Cohort Study (N=14,206) and the Millennium Cohort Study (N=17,082). 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: 6.72%-8.55%), income and occupation. Inequalities were generally stable over developmental and historical time. However, findings suggest a need to focus on widening inequalities both towards the start and end of compulsory schooling.

4.2. 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, 1995b; Pace et al., 2017) and policy makers have sought educational interventions to correct 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?

Are all indicators of SEC equally predictive of vocabulary?

While parent 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). Although there are some claims that maternal education is the most relevant SEC indicator for language development (Hoff, 2013; Hoff et al., 2012), we know of no empirical work explicitly testing this claim in nationally representative samples. Insight into which SEC indicator is most relevant for language development has important implications for understanding the pathways between SEC and language and whether the relationship between each predictor is directly or indirectly related to language development. The MCS2001 affords us a unique opportunity to address this question, with information on parent education, income, wealth, occupational status and relative neighbourhood deprivation available for a large nationally representative sample of children born in the UK between 2000-02.

The association between parent education and child language ability is often explained in terms of the quality and quantity of child directed speech, possibly due to parental knowledge about child development (Hirsh-Pasek et al., 2015; Hoff, 2003;

Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Pan, Rowe, Singer, & Snow, 2005; Rowe, 2008, 2012; Schwab & Lew-Williams, 2016). Importantly, parent knowledge about child development is positively correlated with levels of parent education, and such parent knowledge mediates the relation between parent education and child language outcomes (Rowe, 2008; Rowe et al., 2016; Suskind et al., 2018). More educated parents are more likely to seek advice from professionals (for example, nurses) and their immediate family (such as their own parents), whereas less educated parents are more likely to seek advice from extended family members (such as grandparents); seeking advice from professionals may be positively related to knowledge about child development (Bornstein et al., 2010; Rowe et al., 2016). Parent education may also reflect language related genetic factors, and the role of parent's own vocabulary skill in the language development of their child: for example, parents with larger vocabularies provide a higher quality language learning environment, even after adjusting for years of parent education (Rowe et al., 2005). Additionally, educational attainment can be thought of in terms of human capital, which refers to knowledge and skills, and it is thought that a parent's human capital shapes that of their child (Conger & Donnellan, 2007; Duncan & Magnuson, 2012). Therefore, a parent's human capital may affect the goals and aspirations they hold for their child and as a result they may invest more time in interacting with their child and in their child's learning than those with lower educational attainment and thus less human capital (Conger & Donnellan, 2007; Duncan & Magnuson, 2012; Sohr-Preston et al., 2013).

Income may also affect language development through the availability of learning resources and linguistically stimulating toys available in the household, which differs as a function of SEC (Bassok et al., 2016; Kaushal et al., 2011; Washbrook & Waldfogel, 2011). According to the family investment model, parents with more income and higher levels of education will invest more resources in child development compared to lower SEC parents, who may instead be forced to spend their limited resources on more immediate needs, and these investments may have a more positive impact on child development, such as language outcomes (Conger & Donnellan, 2007). Furthermore, higher income children may have more access to opportunities for linguistic stimulation, such as visits to museums or libraries. Environments that offer cultural opportunities to children are potentially important for linguistic development, and disparities between low and middle income children in access to such opportunities and resources have been reported (Neuman & Celano, 2001). Income may also be indirectly related to language development through its impact on attendance at high quality day-care settings, which has been linked to positive outcomes, including language,

particularly for those from disadvantaged backgrounds (Geoffroy et al., 2007; Magnuson et al., 2004; Schmerse, 2020). Despite this, those from more deprived backgrounds are less likely to be able to afford or have access to high quality settings (Bassok & Galdo, 2016; Petitclerc et al., 2017; Roy et al., 2014). The family stress model posits that economic difficulty can influence parenting as a result of having a harmful effect on emotions, behaviours and relationships (Conger & Conger, 2002; Conger & Donnellan, 2007). This in turn can affect language development by affecting the interactions parents have with their children (Perkins et al., 2013).

Parental interactions and the level of quality input that children receive from parents is also influenced by levels of household disorganisation, chaos, noise and living in crowded houses, which are associated with poverty and may in turn predict language outcomes (Evans et al., 1999; Vernon-Feagans et al., 2012). Further, high and middle SEC families who were primed to think about recent experiences of financial paucity directed lower quality input to their three year old children in a later play session, and engaged in fewer conversational turns at the end of a month of financial difficulty (Ellwood-Lowe et al., 2022). Income may also be related to brain activity in areas for which there is evidence of activation being associated with enhanced cognitive abilities.. The Baby's First Years project in the US is a randomised controlled trial investigating the causal impact of a poverty reduction intervention, where low income mothers were randomised shortly after giving birth to receive an unconditional cash gift of either $333 (high cash gift) or $20 (low cash gift) a month for the first few years of their babies lives ("Baby's First Years", 2018). In a subgroup analysis of data from this project, babies of mothers in the high cash group exhibited patterns of brain activity more consistent with subsequent enhanced language and cognitive abilities, than those seen in the low cash group (Troller-Renfree et al., 2022), suggesting that interventions that aim to reduce poverty lead to causal changes in brain activity.

Another indicator of family SEC is family wealth. Whereas income refers to financial resources at any one time, for example as a result of pay from employment, wealth refers to the total accumulated resources over time, and is usually conceptualised as total assets net of outstanding total debt (Keister, 2018; OECD, 2015). Assets can be financial (such as savings accounts, stocks and shares) and non-financial (such as home or vehicle ownership; Pfeffer & Schoeni, 2016). Wealth is a profoundly right skewed variable and is more unequally distributed than income; high levels of wealth are concentrated in the top 1% of the distribution, with the wealthiest 10% estimated to own 44% of all UK wealth in 2016 (OECD, 2015; Office for National Statistics, 2019b). Although certainly related to family

income, wealth and income are two distinct constructs — those who are wealthy can have little income, and those with high incomes can have little in the way of assets (Dorling et al., 2007.; Keister, 2018). The correlation between the two constructs is positive, but relatively small. This could be due to some households having a high income from their employment, but high spending / low savings behaviours and therefore low wealth. Conversely, households may have high wealth with little or no current income, because for example they are retired but saved during their working years, or inherited a large amount of wealth (Duncan & Magnuson, 2012; Keister, 2018; OECD, 2015).

There are several possible mechanisms through which family wealth could affect child language development. Firstly, wealth may be a more stable, accurate indicator of one's long term economic resources than income (Jez, 2014). Financial assets may complement income in the purchase of resources for child development, such as books and toys. However it is thought that they are most important in being able to buy a home (Moulton et al., 2021; Orr, 2003; Yeung & Conley, 2008). Rather than providing resources that enhance language ability in the way that income does, family wealth may instead provide families with the opportunity to invest in homes in expensive areas close to high performing schools and local amenities that provide the cultural capital for language development mentioned above (Karagiannaki, 2017; Machin, 2011). Relatedly, there is an established link between house prices and school quality (Department for Education, 2017a; Gibbons, 2012; Machin, 2011), with house value being higher near the best schools than in nearby areas, and prices near poor performing schools being lower than in nearby areas. It is possible that housing wealth (i.e., the ability to afford more expensive homes) may affect language development through children from wealthier families having access to higher performing schools, rather than through the direct provision of resources for language development. Indeed, housing wealth is much more common than financial wealth in the UK (with the opposite pattern being observed in the US; Banks et al., 2004; Cowell et al., 2019), therefore, at least in the UK, any relationship observed between family wealth and language ability could be reflecting the ability to afford homes in areas surrounding better performing schools. Wealth could also be a protective mechanism against the family stress model (see above), acting as a safeguard against any negative effects of sudden income losses, such as unexpected unemployment (Grinstein-Weiss et al., 2014; Killewald et al., 2017; Lusardi et al., 2011; Rothwell & Han, 2010). By acting as a protective mechanism against family stress, wealth may benefit language development. Finally, greater family wealth may benefit language development by allowing mothers to take longer periods of maternity leave. A positive relationship has been

found between the amount of time mothers spend at home in the first few postnatal months and cognitive development (e.g. Brooks-Gunn et al., 2002). A link between family wealth and maternity leave has been empirically identified, with wealthier mothers being able to afford to take longer periods of maternity leave (Ulker & Guven, 2011), allowing them to spend more time with their infants and promote language development.

Despite being an important construct for family economic position, wealth is less often the focus of studies on child development (Duncan & Magnuson, 2012; Killewald et al., 2017). This is likely due to income being a more commonly understood construct and data related to income being more commonly collected than wealth data (Killewald et al., 2017; OECD, 2015; Shanks, 2007). However, some research in the US, focusing on racial disparities, has reported family wealth to be related to children's cognitive attainment, specifically maths ability (Orr, 2003; Shanks, 2007; Yeung & Conley, 2008) and other educational outcomes such as higher education attendance (Conley, 2001; Jez, 2014). It is plausible that family wealth will affect child language development, therefore we also include this as an indicator of SEC.

Occupational status, or social class, reflects one's social position in the labour market, as well as their power and status (Sullivan et al., 2013). It is thought that people's social networks generally consist of others who are similar to them in terms of occupational status (known as occupational homophily; Griffiths & Lambert, 2012; Griffiths, Lambert, & Tranmer, 2011; McPherson, Smith-Lovin, & Cook, 2001). Indeed, Horvat et al (2003) found that parents' social networks were strongly driven by social class, and that working class and poor families often had connections that consisted of their relatives. On the other hand, middle class parents had abundant links with other parents of pupils at the school, presenting them with social ties to help with problems experienced at school. The existence of such occupational homophily may present children of parents in higher status occupations with wider opportunities for communicative development, through hearing their parents talk to a wide network of similarly advantaged people, compared to families with lower occupational status, who may not have as much social capital. The social network of one's parents may be indirectly related to language development, as children will adopt styles of speech and vocabulary used by their parents, when talking to them and when talking to individuals in their social network (Sullivan, 2007).

Parent education, income, wealth and occupational status may be thought of as direct measures of family SEC. However, Bronfrenbenner's ecological systems theory of child development emphasises the importance of factors outside the immediate family

environments, and the interactions between them, on development (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 (Neuman et al., 2018).Therefore, we also included area level deprivation (Indices of Multiple Deprivation) as an additional indicator of SEC to account for deprivation factors outside the immediate family environment.

Each indicator implies different pathways for the effect of SEC on language development, yet there is likely to be overlap in these pathways, making them difficult to disentangle. Directly comparing the predictive value of different SEC indicators is a pre-requisite for understanding why vocabulary inequalities exist and which mechanisms to target to support development. One indicator could be substantially more important than the rest, or multiple measures may be equally important in predicting vocabulary. Identifying the most important indicators for vocabulary provides insight as to which indicator(s) should be the focus of future research, both in investigating inequalities in vocabulary and in unpacking the relationship further, to inform policies and interventions.

There are also practical implications of determining the most relevant indicator(s) for vocabulary development: this would give researchers insight into whether considering one measure of SEC in isolation is sufficient to robustly understand (or control for) inequalities in language ability. Relatedly, it might be that reducing the individual indicators of SEC to a single composite factor may better capture the overall extent of inequalities, compared to any one measure considered, providing a more informative way of understanding inequalities in vocabulary skill. There are many practical contexts, for example in policy development, where a single measure capturing overall inequalities may be desirable to represent and summarise inequalities in vocabulary, yielding easy-to-understand implications. A composite measure of SEC may be useful for this, if demonstrated to be of good predictive value. Our first goal was thus to test whether five key indicators of SEC (parent education, income, wealth, occupational status and neighbourhood deprivation) each predict unique variance in child vocabulary and how much relative variance is predicted by each. We also condensed these indicators into a derived latent composite variable to establish whether a composite measure is a viable way of representing overall inequalities in language ability.

SEC and vocabulary across developmental time

Compelling arguments have been made in favour of early intervention to prevent social disadvantage affecting language before children reach formal education (Doyle et al.,

2009) and the majority of work to date has focussed on very young children. However, there is also evidence that the SEC gap in vocabulary is pronounced among adolescents (Spencer et al., 2012; Sullivan & Brown, 2015). Other work has indicated stability in the gap between childhood and early adolescence (for example, Farkas & Beron, 2004). Furthermore, there is evidence to suggest that inequalities in language increase even between early childhood years. For example, Fernald et al (2013) found the gap widened between 18 and 24 months and work with the UK Millennium Cohort Study has indicated that social disparities in vocabulary either remain stable or actually increase between ages 3 and 5 (Becker, 2011). While there is evidence of SEC differences in vocabulary at secondary school level (e.g., Maguire et al., 2018; Spencer, Clegg, & Stackhouse, 2012), we know of no work testing whether the gap is widening as children grow up in contemporary society.

The period between late childhood and early adolescence sees important advances in language development, with an estimated 3000-5400 words a year being acquired during this period (Landauer & Dumais, 1997; Nagy & Herman, 1987; Smith, 1941) and the development of longer, formal conversational styles occurring (Berman, 2007; Sullivan et al., 2021). Vocabulary advances in the late childhood and early adolescent years, in terms of quality and quantity, alongside other linguistic developments during this time period (Berman, 2007). It is thought that early language experiences shape the increasing complexity of later language development (Schwab & Lew-williams, 2016), and if children from a high SEC background have a richer language experience than their lower SEC counterparts, then it is plausible that the SEC gap in vocabulary may persist as children age. It is also plausible that the gap could widen throughout childhood, as children with larger vocabularies are likely able to identify words they don't know and extrapolate their meanings by drawing on the meanings of words that they do know. The social experiences of ageing children, such as participation in cultural activities, is likely to be important in the developing complexity of language ability (Berman, 2007), and these may differ as a function of SEC. It is thus important to investigate the influence of SEC on language at later stages of development.

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 parent 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. The MCS2001 has vocabulary test scores at ages 3, 5, 11 and 14 in a contemporary cohort of children, enabling us to investigate this developmental question while also including rich and detailed information on socioeconomic position.

SEC and vocabulary across historical time

Trends in the relationship between SEC and vocabulary over historical time are less understood. Large societal changes in the UK have seen an increase in the proportion of women in the labour market 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; Amaranayake et al., 2000; Holmes & Mayhew, 2012; OECD, 2015). Many more jobs are now also knowledge-based, making language and cognitive skills of great importance for the UK economy (Deloitte, 2016), and putting pressure on parents to support their children's cognitive development to open doors to the labour market. The education system has also seen considerable change over recent decades, particularly in terms of compulsory school leaving age, the proportions of individuals remaining in post-compulsory education, and an increase of parents attending university (Amaranayake et al., 2000; Bolton, 2012). For example, the school leaving age increased from 15 to 16 in 1972 (Bolton, 2012), affecting parents of the MCS2001, but not parents of the BCS1970. These broad shifts in society have the potential to change the association between different measures of SEC and vocabulary ability. Further, the MCS2001 cohort consists of children born 2000~2002, during the "New Labour" government, which had a focus on reducing inequalities in educational access and attainment, with initiatives and policies such as Sure Start and Every Child Matters, so it is plausible that inequalities in vocabulary in this cohort may be narrower compared to that of the BCS1970 cohort. 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.

The current study

Our three aims were split into four research questions, with the first two research questions addressing our first aim:

- RQ1: Establish inequalities in language development: at ages 3, 5, 11 and 14: what is the variation captured by each indicator of SEC individually?
- RQ2. How does a composite measure of overall socioeconomic position perform relative to individual measures and combinations of measures?
- RQ3. How does this relationship change over developmental time? (Language ability at ages 3, 5, 11 and 14)
- RQ4: Cross-cohort comparison: how has this relationship changed with historical time? (Relationship between SEC and language ability in two nationally representative cohorts, born 30 years apart)

The MCS2001 has abundant information regarding childhood socioeconomic position, enabling us to thoroughly investigate the explanatory value of individual measures of SEC. There are also follow up sweeps throughout childhood, at ages 3, 5, 11 and 14, each with measures of vocabulary, allowing us to investigate inequalities in language ability across developmental time. We were able to address our final aim with a cross-cohort comparison of the MCS2001 and the BCS1970, as cohort members were born 30 years apart, with increases in income and wealth inequality in the UK evident during this period, as well as changes to the occupational structure and educational system. In a series of pre-registered, confirmatory analyses, we met the first two goals by analysing data from the Millennium Cohort Study (17,082 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 (14,206 children born in 1970; BCS1970, and 16,033 children in the MCS2001).

4.3. Method

Data

We used data from two large nationally representative UK birth cohort studies: The Millennium Cohort Study (MCS2001) and the 1970 British Cohort Study (BCS1970). Research questions 1-3 consisted of analyses of the MCS2001 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. Research question 4 used data from the MCS2001 and BCS1970 cohorts in a cross-cohort comparison. The use of these two datasets for a cross-cohort comparison allows us to thoroughly examine inequalities in language ability in two generations born 30 years apart, during a period which has seen a rise in income and wealth inequality and changes to occupational and educational structures.

MCS2001. The MCS2001 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 and 14 years. More information on the MCS2001 can be found here: https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/.

BCS1970. The BCS1970 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 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 cohort) 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 MCS2001cohort members completed a battery of cognitive tests throughout childhood and into early adolescence. It is worth noting that in this cohort, at age 7, cohort members did not complete a vocabulary test - they instead completed a reading test (BAS II Word Reading subscale*)* - and we have thus not included age 7. Full details about the completed vocabulary tests can be found in Chapter 2.

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. 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 et al., 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 et al., 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 MCS2001cohorts respectively, referred to as late childhood language ability) and ages 16/14 (BCS1970 and MCS2001cohorts 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 et al., 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 (Sullivan et al., 2021). 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 et al., 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 MCS2001cohort members were aged 11, they completed the BAS II verbal similarities subscale (detailed above). As already mentioned, test scores for the MCS2001cohort 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 et al., 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 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 (Moulton et al., 2020). Scores were adjusted for age and converted to *z* scores for analyses.

Measures of socioeconomic position.

Analysis of 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 discussed below.

Parent education. As a measure of parent's education when cohort members were aged 3, highest household NVQ level was used (both academic and vocational qualifications derived into NVQ levels 1-5, with level 5 equating to higher qualifications).

We first tested maternal vs highest parent education in predicting vocabulary at each age (see Appendix 2, Section 3) and based on findings that highest household education

consistently predicted the most variance in vocabulary at each age, we use a measure of highest household education in our analyses.

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

Wealth. A measure of total net wealth, taken from the age 11 sweep of the MCS2001cohort. 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. 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 et al., 2017); thus, total net wealth was split into quintiles for our analyses.

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 3 years. If data were missing, occupational status at 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, 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 et al., 2019). Therefore, we used IMD deciles at age 3 (with 1= most deprived and 10=least deprived) as a measure of relative neighbourhood deprivation.

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 4.1.

Parental 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.

Potential confounders.

We adjusted for sex at birth (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 4.1.

Table 4.1 Cross-cohort Harmonization of Variables.

Measure	BCS1970	MCS2001 cohort	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
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 *Note: students/volunteers were categorised as*	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 This 5-class version was collapsed into a 3-class version, as shown here:	Composite variable, with a 4th category for unemployment: BCS1970: 1. Professional & Managerial 2. Skilled 3. Semi-skilled and unskilled 4. Unemployed MCS2001 cohort: 1. Higher managerial 2. Intermediate 3. Routine 4. Unemployed *Note: The convention used in the MCS2001 cohort was used*

	unemployed as they have no paid employment.	https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010#classes-and-collapses	*for the occupational status variables from both cohorts, for ease.*
Parental education: highest educational qualification (highest household level)	BCS1970 (Age 5) 1. No qualifications 2. Vocational qualifications 3. O levels 4. A-levels 5. State registered nurse 6. Certificate of education 7. Degree +	MCS2001 cohort (Age 5) 1. None of these qualifications 2. GCSE grades D-G 3. O level/GCSE grades A-C 4. A/AS/ S Levels 5. Diplomas in higher education 6. First degree 7. Higher degree 8. 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
Ethnicity	European UK European Other West Indian Indian-Pakistani Other Asian African Other	White Mixed Indian Pakistani and Bangladeshi Black or Black British Other Ethnic group (incl. Chinese, Other)	Categorical measures collapsed into 0=White, 1=Minority
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

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 (14.6%)(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, EAL or income quintiles. 14.71% of vocabulary scores at age 3 were missing, 12.42% of age 5 vocabulary scores were missing, 23.93% of age 11 vocabulary scores were missing and 36.83% of age 14 vocabulary scores were missing. 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 (see Appendix 2, Sections 11-14 respectively). Missing data among the components of our wealth variable were also high (ranging from 46.62% (total savings) to 63.63% (outstanding mortgage). We therefore conducted sensitivity analyses where we considered all cohort members with a response to ≤ 1 wealth component variable and ≤2 wealth variables (see Appendix 2, Sections 15-16 respectively). Overall, these sensitivity checks revealed a similar pattern of results to the main analyses.

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 (7.3% MCS2001cohort, 21.3% BCS1970 cohort (White et al., 2011). For the MCS2001cohort, 6.69% of age 5 vocabulary scores were missing, 18.95% of age 11 vocabulary scores were missing and 32.7% of age 14 vocabulary scores were missing. For the BCS1970 cohort, 17.2% of age 5 vocabulary scores were missing, 20.04% of age 10 vocabulary scores were missing and 62.85% of age 16 vocabulary scores were missing (as a result of the teachers strike in 1986). No data were missing for sex in either cohort. Again, combined sampling and attrition weights were applied to MCS2001cohort data. 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 (see Appendix 2, supplementary methods for details).

Analyses.

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 (see Tables S1 and S2, Appendix 2, Section 2). Mean language scores for each SEC group are reported (see Tables 4.2 and 4.3).

Establish 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 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 supplementary methods, Appendix 2 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. 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 (Burnham & Anderson, 2002; Fabozzi et al., 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 (Burnham & Anderson, 2002; Fabozzi et al., 2014). AIC values are needed here as the models are not nested, and 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.

How does this relationship 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 are 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.

Cross-cohort comparison: how has the relationship changed with historical time? (Relationship between SEC and vocabulary size in 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 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 vocabulary were standardised within each cohort, we were able to directly compare coefficients between cohorts and establish the rate of inequality in vocabulary at each age in the two cohorts.

4.4. Results

Analytic samples.

To address the first two research questions in a contemporary cohort, we analysed the data of 17,082 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, 85.99% were of White ethnicity and 88.50% 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 MCS2001 cohort are negligible (see Table S1, Appendix 2, Section 2).

For the cross-generation comparison, we analysed the data of 14,206 children in the BCS1970 and 16,033 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 4.1 for details of harmonisation). 48.88% of BCS1970 cohort members were female, 94.60% were of White ethnicity and 95.77% did not speak English as an additional language. In the cross-cohort comparison, 48.66% of MCS2001 cohort members were female, 86.04%% were of White ethnicity and 88.65% 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 S2, Appendix 2, Section 2).

Descriptive Statistics.

As can be seen in Table 4.2, 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. Examination of density plots (see Figure 4.1) revealed that mode value in language scores (indicating a high concentration of scores) for the most deprived SEC group are overall shifted to the left of the distribution. For the least deprived group, the mode is overall shifted to the right of the distribution. This indicates a higher frequency of poor language scores among the most deprived groups. The distributional differences in inequalities in language are slightly less pronounced for some indicators at age 14, although they are much more pronounced for education and income at this age. Overall, there is a similar distributional pattern across all ages and SEC indicators, indicating that inequalities are evident regardless of the age and SEC indicator.

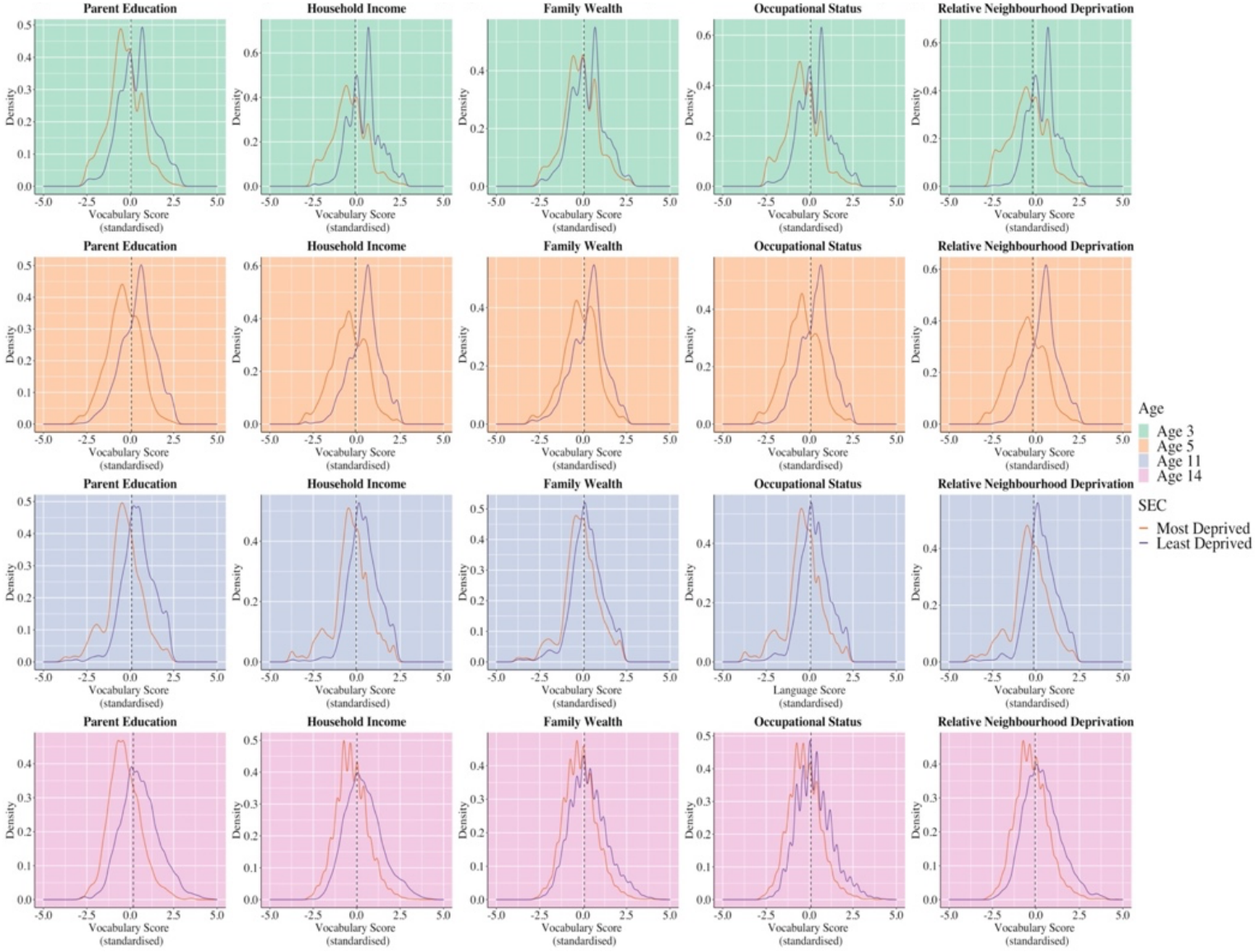
Table 4.2 Means (±SD) and 95% CIs for language scores in each SEC group at each age (MCS2001 cohort)

		Mean (SD) [95%CI]			
		Age 3 language *Range=20-80*	Age 5 language *Range=20-80*	Age 11 language *Range=20-80*	Age 14 language *Range = 0-20*
EDUCATION	Parent education (NVQ1)	45.27(±10.29) [44.64;45.90]	49.79(±10.49) [49.15;50.44]	55.00(±10.05) [54.38;55.62]	6.13(±2.39) [5.98;6.27]
	Parent education (NVQ2)	47.93(±10.64) [47.61; 48.25]	52.78(±10.29) [52.47;53.09]	56.75(±9.93) [56.45;57.05]	6.52(±2.35) [6.45;6.59]
	Parent education (NVQ3)	49.61(±10.67) [49.21;50.01]	54.25(±10.11) [53.87;54.62]	58.31(±9.34) [57.97;58.66]	6.80(±2.42) [6.71;6.89]
	Parent education (NVQ4)	52.39(±10.75) [52.11;52.67]	57.54(±10.18) [57.27;57.80]	60.72(±8.98) [60.49;60.96]	7.58(2.65) [7.51;7.64]
	Parent education (NVQ5)	53.46(±11.47) [52.82;54.10]	59.55(±10.51) [58.96;60.13]	63.29(±8.66) [62.81; 63.78]	8.53(±2.90) [8.37;8.70]
	Parent education (None of these/overseas qualifications)	41.36 (±11.50) [40.86;41.85]	46.45(±11.64) [45.95; 46.95]	54.07(±10.96) [53.59;54.54]	5.96(±2.26) [5.86; 6.06]
INCOME	Income quintile 1	44. 23(±11.36) [43.88;44.58]	49.20(±11.07) [48.85; 49.54]	55.33(±10.54) [55.00;55.66]	6.21(±2.36) [6.14;6.29]
	Income quintile 2	46.66(±11.04) [46.31;47.01]	51.56(±10.79) [51.22;51.90]	56.63(±10.18) [56.30;56.95]	6.56(±2.41) [6.49;6.64]
	Income quintile 3	50.30(±10.75) [49.94;50.67]	55.28(±10.03) [54.94;55.62]	58.69(±9.16) [58.37;59.00]	6.95(±2.47) [6.87;7.04]
	Income quintile 4	52.54(±10.54) [52.17;52.91]	57.19(±10.13) [56.83;57.55]	60.12(±9.19) [59.79;60.44]	7.39(±2.66) [7.30;7.49]
	Income quintile 5	53.73(±10.20) [53.36;54.10]	59.36(±9.82) [59.00;59.72]	62.26(±8.60) [61.94;62.57]	8.01(±2.75) [7.91;8.11]
WEALTH	Wealth quintile 1	47.43(±11.35) [47.05;47.81]	52.31(±10.98) [51.95;52.68]	56.89(±10.13) [56.55;57.23]	6.59(±2.46) [6.51;6.67]
	Wealth quintile 2	47.96(±11.34) [47.58; 48.34]	52.71(±10.82) [52.34;53.07]	57.24(±10.00) [56.90;57.57]	6.65(±2.43) [6.57;6.73]
	Wealth quintile 3	48.64(±11.28) [48.26;49.02]	53.54(±10.78) [53.18;53.90]	57.89(±9.98) [57.55;58.23]	6.85(±2.51) [6.76;6.93]
	Wealth quintile 4	49.91(±11.27) [49.53;50.29]	54.88(±10.97) [54.52; 55.25]	58.89(±9.72) [58.56;59.21]	7.04(±2.57) [6.96;7.13]
	Wealth quintile 5	51.99(±11.09) [51.62;52.36]	57.46(±10.80) [57.10;57.82]	60.90(±9.24) [60.59;61.21]	7.69(±2.81) [7.60;7.79]
OCCUPATIONAL STATUS	Occupational status (routine)	47.31(±11.13) [46.97;47.65]	52.20(±10.73) [51.87;52.52]	56.74(±10.00) [56.43;57.04]	6.56(±2.38) [6.49;6.63]
	Occupational status (intermediate)	50.12(±10.96) [49.74;50.50]	54.64(±10.62) [54.28;55.01]	58.69(±9.38) [58.36;59.02]	6.87(±2.46) 6.78;6.95]

Occupational status (higher managerial)	52.77(±10.67) [52.50;53.03]	58.27(±9.98) [58.02;58.51]	61.26(±8.89) [61.04;61.48]	7.74(±2.71) [7.68;7.81]
Occupational status (unemployed	44.26(±11.04) [43.90;44.62]	48.98(±10.86) [48.62;49.33]	54.98(±10.56) [54.64;55.33]	6.20(±2.40) [6.12;6.28]
Relative neighbourhood deprivation (most deprived decile)	43.71(±11.58) [43.28;44.13]	48.69(±11.17) [48.28;49.10]	54.83(±10.61) [54.44;55.22]	6.26(±2.40) [6.17;6.35]
Relative neighbourhood deprivation (10 - < 20%)	45.84(±11.79) [45.36;46.31]	50.53(±10.95) [50.09;50.97]	56.98(±10.15) [56.57;57.39]	6.56(±2.44) [6.47;6.66]
Relative neighbourhood deprivation (20 - < 30%)	48.03(±11.11) [47.55;48.52]	53.14(±10.60) [52.67;53.60]	57.58(±9.92) [57.14;58.01]	6.75(±2.54) [6.64;6.86]
Relative neighbourhood deprivation (30 - < 40%)	49.08(±11.21) [48.54;49.61]	53.75(±10.49) [53.25;54.25]	58.37(±10.04) [57.89;58.85]	6.86(±2.57) [6.74;6.98]
Relative neighbourhood deprivation (40 - < 50%)	49.53(±11.01) [48.97;50.09]	54.51(±10.89) [53.96;55.06]	58.34(±9.13) [57.87;58.80]	6.95(±2.53) [6.82;7.08]
Relative neighbourhood deprivation (50 - < 60%)	50.50(±10.93) [49.94;51.07]	55.58(±10.43) [55.04;56.12]	58.86(±9.91) [58.35;59.38]	7.03(±2.54) [6.90;7.16]
Relative neighbourhood deprivation (60 - < 70%)	51.56(±10.60) [50.96;52.17]	56.33(±10.40) [55.74;56.92]	60.12(±10.01) [59.55;60.69]	7.24(±2.70) [7.09;7.39]
Relative neighbourhood deprivation (70 - < 80%)	52.16(±10.51) [51.58;52.74]	57.48(±10.59) [56.89;58.07]	60.06(±9.11) [59.55;60.56]	7.50(±2.70) [7.35;7.65]
Relative neighbourhood deprivation (80 - < 90%)	52.24(±10.37) [51.69;52.79]	57.58(±10.21) [57.04;58.12]	60.18(±9.05) [59.70;60.66]	7.48(±2.56) [7.35;7.62]
Relative neighbourhood deprivation (least deprived decile)	53.62(±9.96) [53.10;54.14]	58.90(±9.55) [58.40;59.40]	61.47(±8.66) [61.02;61.92]	7.77(±2.76) [7.62;7.91]

Descriptive statistics combined across 25 imputed datasets. Descriptive statistics are sample and attrition weighted

Figure 4.1 Density plots showing the distribution of vocabulary (ages 3, 5, 11 and 14) in the highest and lowest SEC groups for each indicator



Which SEC measures predict child vocabulary?

Figure 4.2 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). Parent education explains the largest proportion of variance in vocabulary at each age (between 6.72% and 8.55% of variance), closely followed by income and occupational status. Wealth and relative neighbourhood deprivation consistently contribute the least variance in vocabulary scores, regardless of age.

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).

Age 3. Parent education (Dm(5, 4449.02)= 46.73, p<.001), income (Dm(4, 4311.06)= 17.36, p<.001), occupational status (Dm(3, 3601.65)= 17.68, p<.001) and relative neighbourhood deprivation (Dm(9, 8015.91)= 2.61, p=.005) all accounted for significant variance in language ability at age 3. Wealth did not account for significant variance (Dm(4, 357.98)= 1, p=.457).

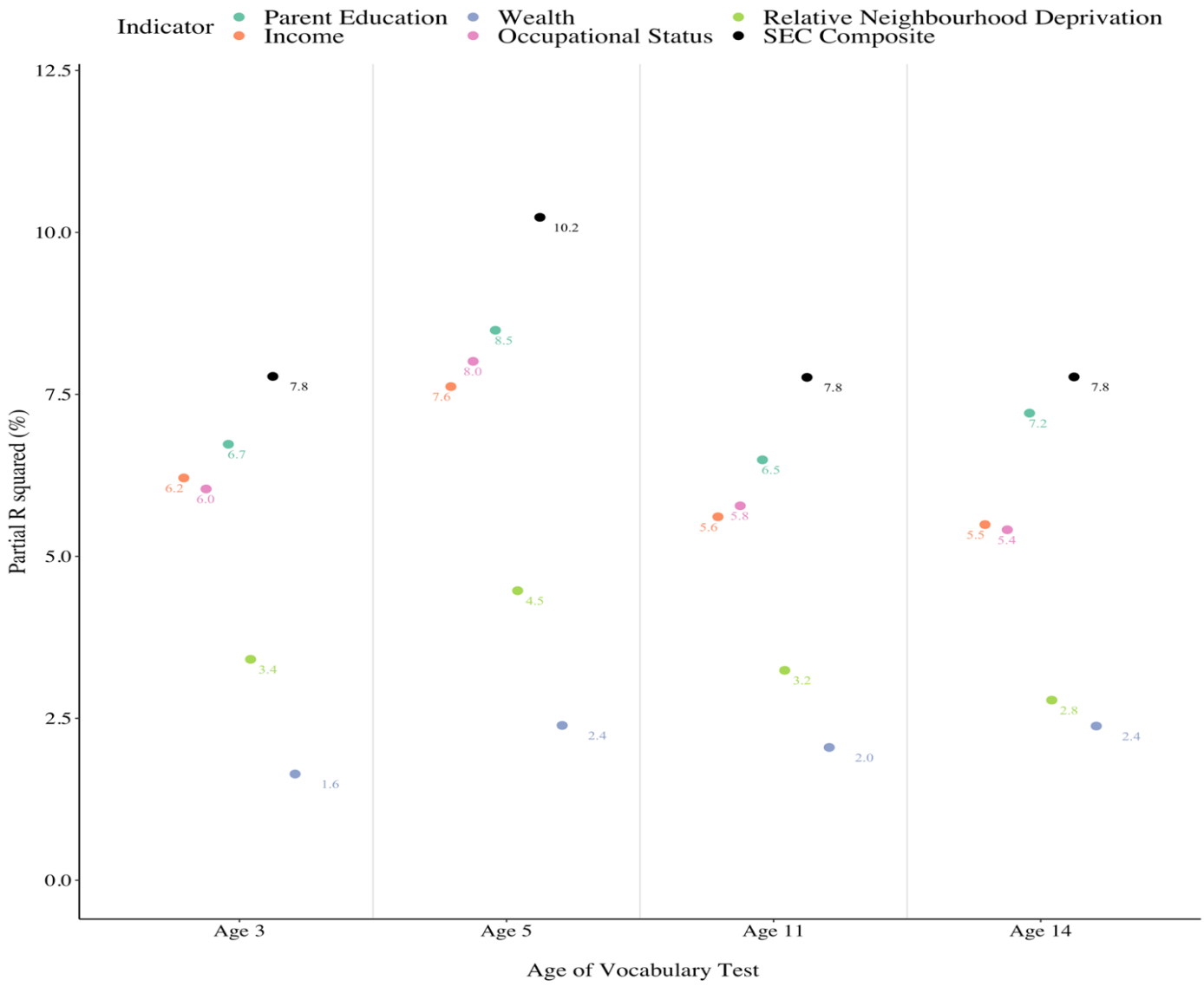
Age 5. Parent education (Dm(5, 3553.15)= 50.75, p<.001), income (Dm(4, 3378.3)= 13.2, p<.001), occupational status (Dm(3, 2861.34)= 30.24, p<.001) and relative neighbourhood deprivation (Dm(9, 6839)= 3.58, p<.001) all accounted for significant variance in language ability at age 5. Wealth did not account for significant variance (Dm(4, 327.59)= 1.88, p=.114).

Age 11. Parent education (Dm(5, 886.3)= 28.08, p<.001), income (Dm(4, 1105.92)= 7.59, p<.001), occupational status (Dm(3, 811.4)= 15.02, p<.001) and relative neighbourhood deprivation (Dm(9, 2298.22)= 3.14, p<.001) all accounted for significant variance in language ability at age 11. Wealth did not account for significant variance (Dm(4, 290.05)= 2.11, p=.079).

Age 14. Parent education Dm(5, 811.14)= 39.65, p<.001), income (Dm(4, 694.12)= 6.11, p<.001), occupational status (Dm(3, 339.7)= 7.93, p<.001) and wealth (Dm(4, 318.3)= 4.03, p=.003) all accounted for significant variance in language ability at age 14. Relative neighbourhood deprivation did not account for significant variance (Dm(9, 1706.49)= .81, p=.61).

In sum, this drop-one analysis revealed that parent education, income and occupational status accounted for significant unique variance in vocabulary at all ages. Neighbourhood statistics accounted for significant variance in vocabulary at ages 3, 5 and 11, while wealth only accounted for significant variance in vocabulary at age 14. Parent education is the "best" predictor, however, measures of parent education, income and occupational status all explain large amounts of unique variance in vocabulary skill at all time points considered. Measures of wealth and relative neighbourhood deprivation may be useful in predicting inequalities in vocabulary at some ages. However, measures of parent education, household income and occupational status are likely to give a better picture of inequalities than either of these two measures alone.

Figure 4.2. Variance explained by SEC indicators when predicting vocabulary in the 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).

Reducing individual indicators to a single composite factor may afford us an efficient way of communicating and understanding inequalities in vocabulary. Confirmatory factor analysis was therefore used to create a composite variable of SEC (see supplementary methods, Appendix 2), 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 at all ages (see Table S4 in Appendix 2, Section 5) and explained 7.4-10.1% of variance in language across ages (see Figure 4.2). 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

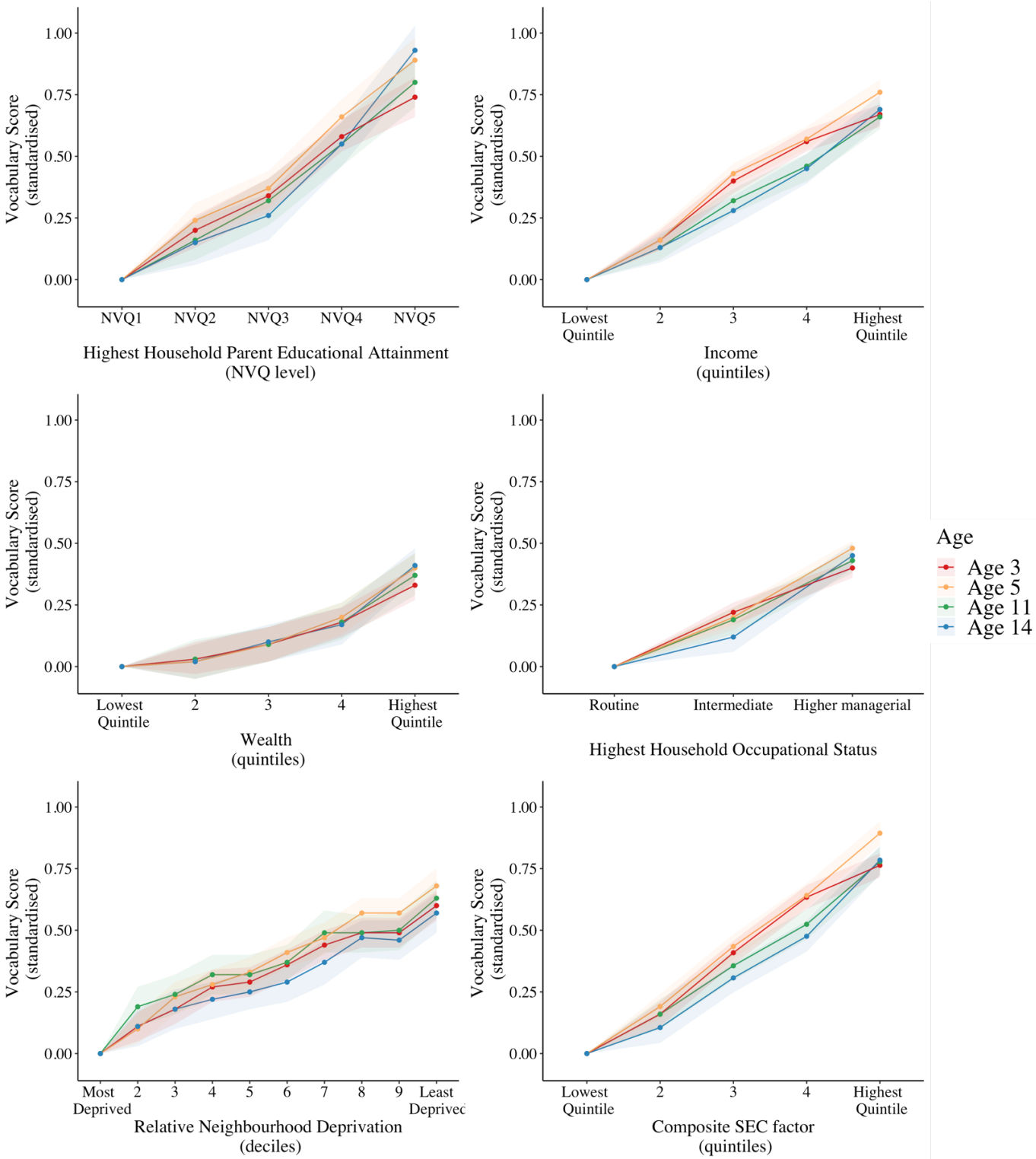
S5 in Appendix 2, Section 5). This indicates that if one needs to identify a single variable for use in analyses, then a composite variable would be a better choice than any of the original individual predictors. In the absence of such a constraint, including a set of multiple predictors would be preferable.

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

Figure 4.3 shows the relationships between each SEC indicator and vocabulary at each age (coefficients and 95% CIs plotted; see also Table S6, Appendix 2, Section 6). Because vocabulary scores were converted to *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 2, Section 7).

Figure 4.3. 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).

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 two measures of SEC available for both cohorts – highest household educational attainment and highest household occupational status. 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). 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 4.3)(Holmes & Mayhew, 2012). For both 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 2, Section 8).

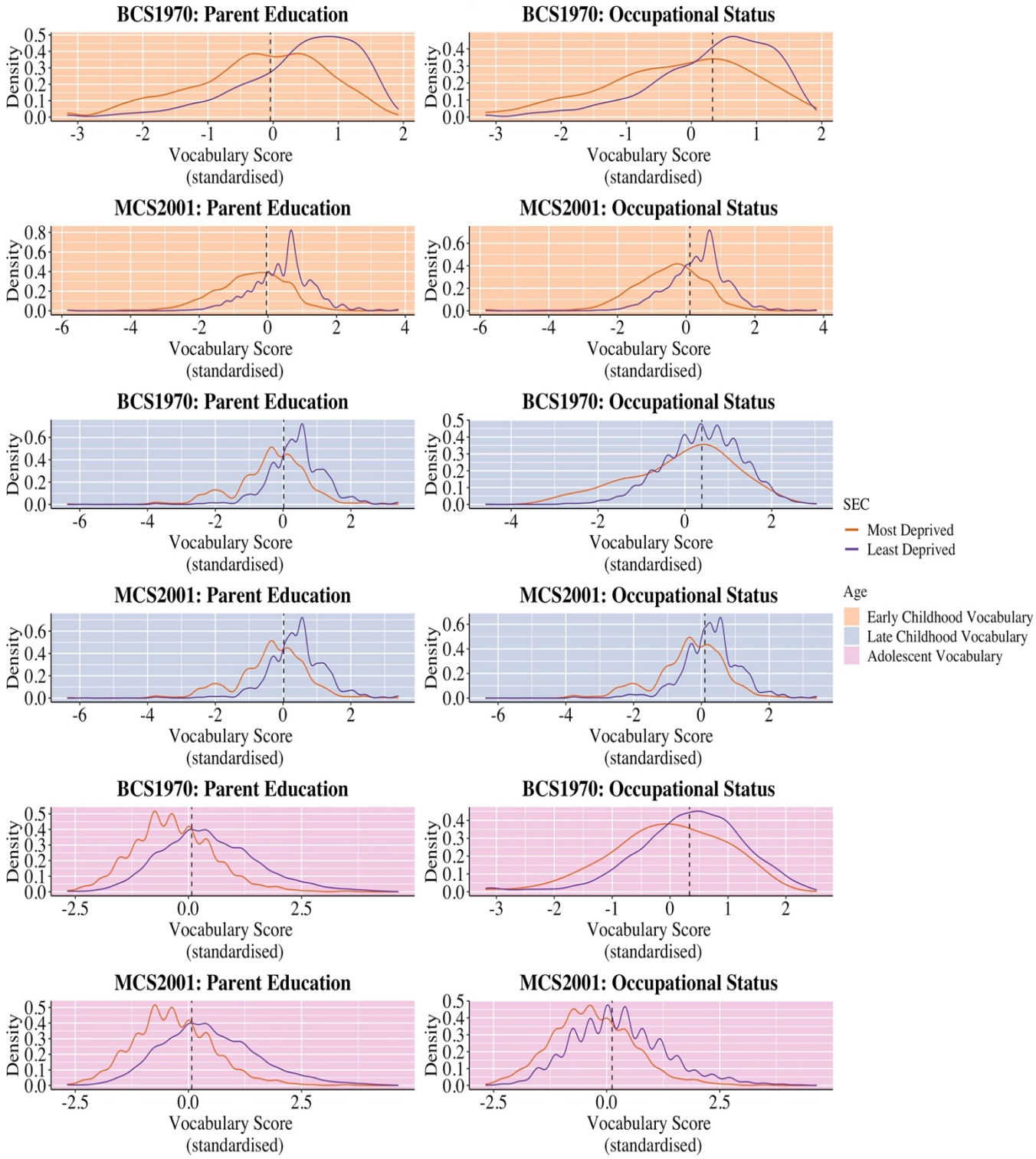
Examination of density plots for language at each age in the two cohorts (see Figure 4.4) revealed a similar overall distributional pattern in the two cohorts, with the distribution for the most deprived group being negatively skewed, indicating more scores below the mean in this group. Overall, the distribution for the least deprived groups was positively skewed, indicating more scores above the mean in this group. There are slight distributional differences based on the indicator of SEC used and the age of the language score in the two cohorts, however the overall pattern is similar. This indicates that there are inequalities in language ability based on parent education and occupational status, in both cohorts. It is also worth noting that the skew between SEC groups could be due to the suitability of the language tests used; if the vocabulary measures were developed using higher SEC children, then it is plausible that those in more disadvantaged groups will fall close to the floor of the distribution.

Table 4.3. Descriptive Statistics in MCS2001 and BCS1970 for the cross-cohort comparison

	MCS2001 Mean (SD) or % [95%CI] *N=16,033*	BCS1970 Mean (SD) or % [95%CI] *N=14,206*
Language		
Early childhood	107.98(±16.13) [107.73; 108.23]	34.59(±11.17) [34.40;34.77]
Late childhood	12017(±16.80) [119.91;120.43]	12.01(±2.64) [11.97;12.05]
Adolescence	7.01(±2.61) [6.97;7.05]	41.12(±13.04) [40.90;41.33]
Potential confounders		
Gender(male)	51.34%	51.12%
Gender(female)	48.66%	48.88%
Ethnicity (white)	86.04%	94.60%
Ethnicity (minority)	13.96%	5.40%
English as an additional language (no)	88.65%	95.77%
English as an additional language (yes)	11.35%	4.23%
SEC predictors		
Parent education (no/low level)	22.20%	55.84%
Parent education (O-levels/GCSEs grades A*-C)	31.93%	20.43%
Parent education (post-16 quals)	21.55%	7.52%
Parent education (university level quals)	24.31%	16.21%
Occupational status (routine)	24.83%	17.48%
Occupational status (intermediate)	20.80%	54.67%
Occupational status (higher managerial)	40.87%	27.17%
Occupational status (unemployed)	13.50%	0.68%

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

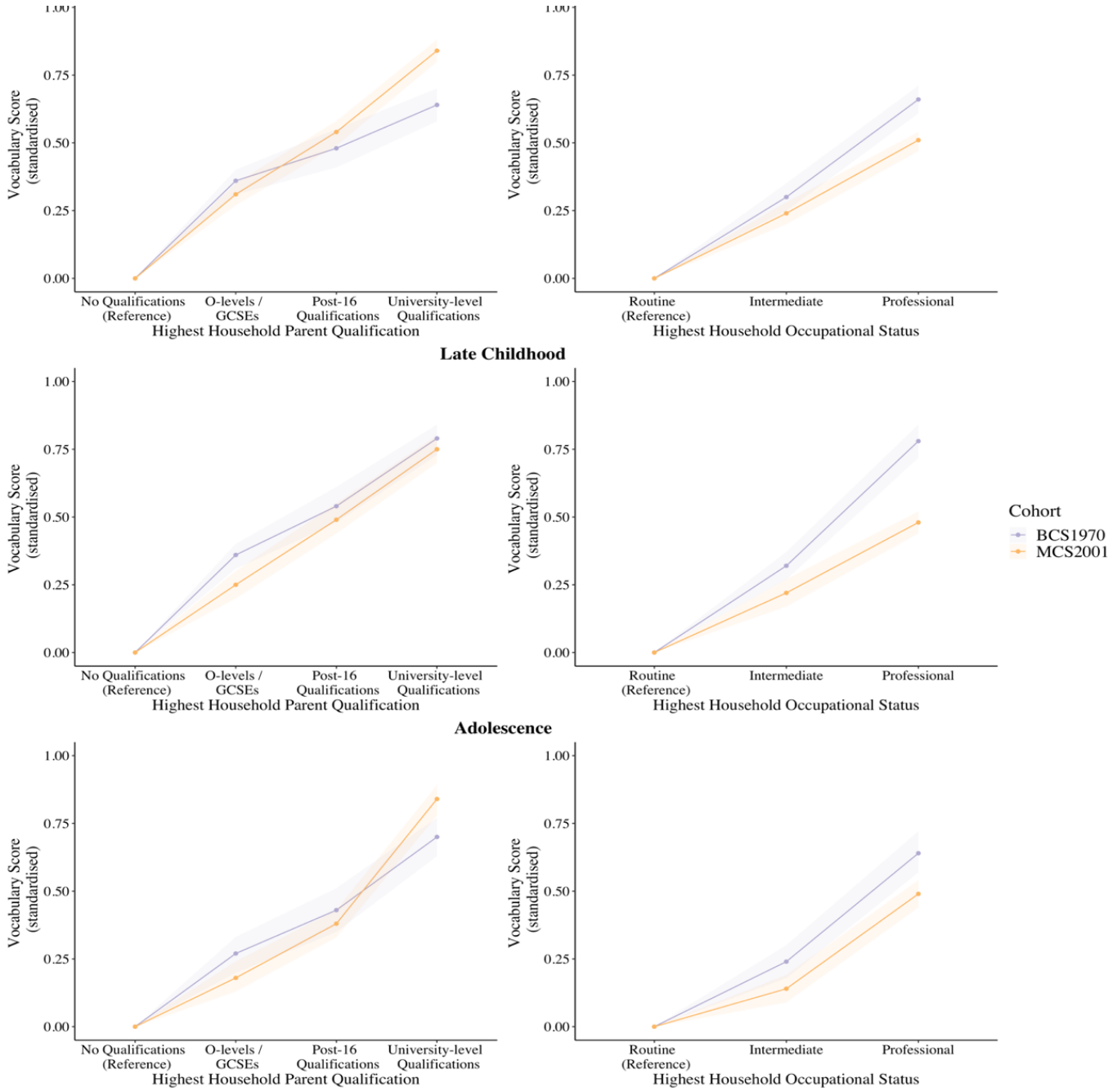
Figure 4.4. Density plots showing the cross-cohort distribution of vocabulary scores (early childhood, late childhood, and adolescence) in the most and least deprived SEC groups



As can be seen in Figure 4.5, vocabulary scores generally increased with SEC regardless of indicator and cohort (also see Table S10, Appendix 2, Section 8). 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 were wider for the BCS1970 cohort at all ages, as indicated by the steeper slopes for this cohort. Standardised coefficients are displayed in Figure 4.5, reflecting relative changes in inequalities in the two cohorts, based on two indicators of SEC. What is unclear from Figure 4.5 is whether or not the amount of variance explained by occupational status and parent education has changed between the two cohorts. It is important to investigate how much variation in language ability is still a function of SEC in the MCS2001 cohort, and whether or not this has decreased, or indeed increased since the BCS1970 cohort. The difference relative to the reference category of a given SEC indicator is indicative of effect size, and we would expect this to be correlated with R^2. We therefore display the R^2 values for parent occupational status and parent education in predicting language in early childhood, late childhood and adolescence in the two cohorts (see Figure 4.5).

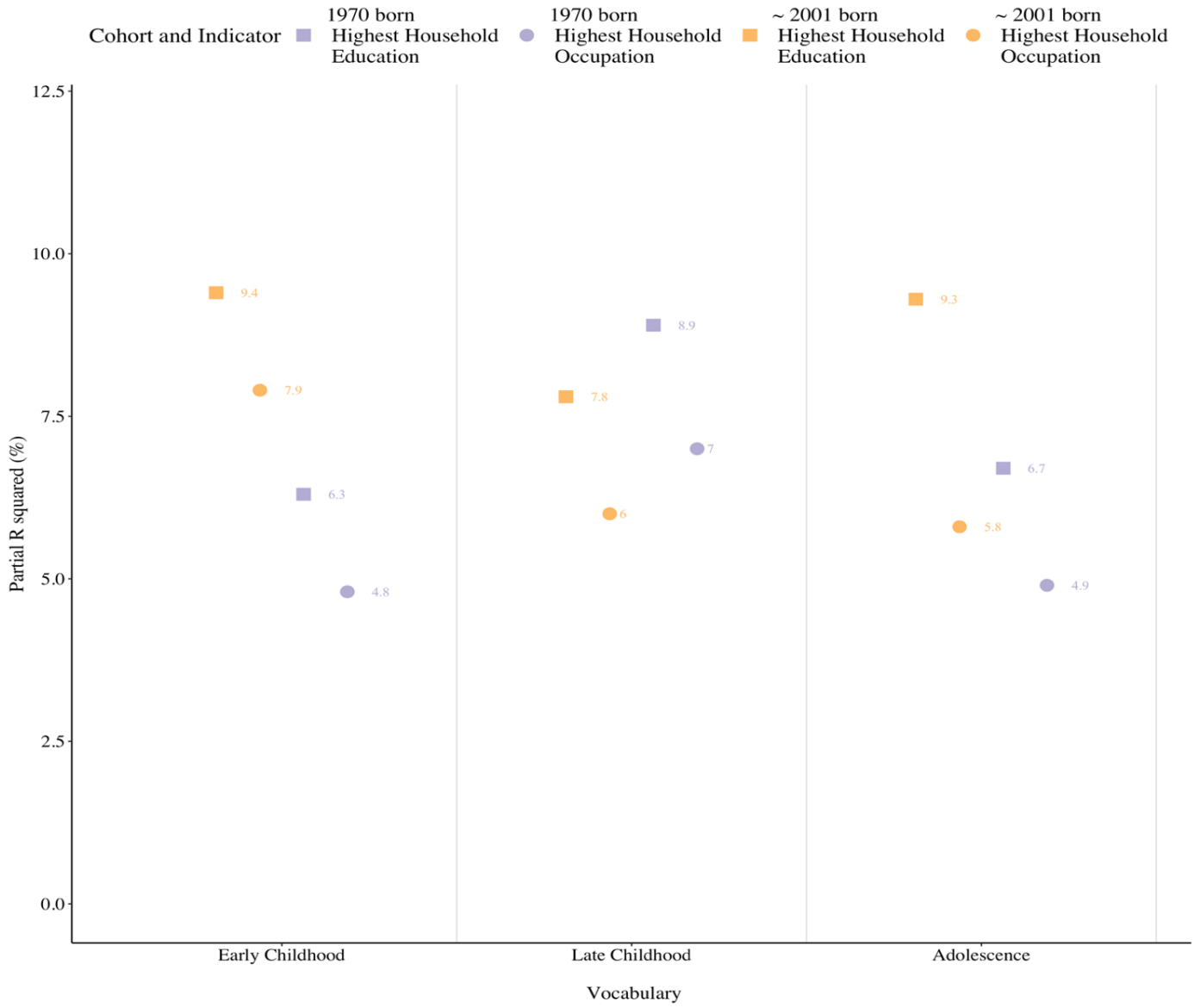
As can be seen from partial R^2 values (Figure 4.6), inequalities are substantial in both cohorts. There is no evidence of a decrease in SEC inequalities over the 30-year period and there is even some evidence that inequalities may have widened in early childhood, with SEC indicators explaining more 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. This pattern of results for late childhood vocabulary in the BCS1970 could be driven by the vocabulary measure used in mid-childhood in the BCS1970, as the amount of variance explained by both SEC indicators in mid-childhood language is much higher than early childhood and adolescence in the BCS1970, or compared to the MCS2001.

Figure 4.5. 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) and highest household occupational status (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 4.6. Variance in language explained by SEC indicators in the MCS2001 and BCS1970 cohorts



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.

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 2, Section 9)(Donaldson, 1998). Second, we restricted our analyses to those of a White ethnicity only

(see Appendix 2, Section 10). Neither analysis resulted in a change in the pattern of results observed.

4.5. 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, with parent education having the greatest predictive value (closely followed by income and occupation) and wealth the least, (ii) inequalities persist from ages 3 to 14 years, with SEC indicators explaining 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.

Are all indicators of SEC equally predictive of vocabulary?

Overall, the SEC predictor that explains the most variance in child vocabulary across development is parent education. However, income and occupational status also uniquely predicted large amounts of variance. For all three of these indicators, a step up from each level to the next was associated with a substantial step up in vocabulary. These indicators therefore each 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. Parent 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 et al., 2021), as well as mediating the relationship between the home learning environment and vocabulary. 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 et al., 2013). The role of genetics should also be considered here, as language ability is claimed 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): caregivers and infants with different genetic profiles shape

learning environments differently. Unravelling this will require rich data sets that include information regarding interaction dynamics.

One may expect wealth to supplement income, for example, in the provision of resources and learning materials (Karagiannaki, 2017), meaning their predictive value would be similar. However, while income explained about 6% of variance in children's vocabulary, family wealth explained far less (about 2%). Income is often assumed to affect vocabulary outcomes through the provision of learning resources (Duncan et al., 2017; Washbrook & Waldfogel, 2011). The ways in which wealth, on the other hand, influences language development are less clear. The unique variance explained by wealth for age 14 language captures variance not explained by the other SEC indicators, including income. It is possible that the variance in vocabulary that is explained by income better captures the ability of parents to provide resources for their children's language learning, than does the measure of wealth. Wealth is not a measure of disposable income, rather it reflects the long-term accumulation of and is usually conceptualised as total assets net of outstanding total debt (Keister, 2018; OECD, 2015). The finding that wealth only predicts unique variance in language ability at one particular age, and specifically age 14, is perhaps not surprising, as it has been suggested that wealth's role as an SEC predictor can have differential effects on development that vary with age. For example, family wealth may be an important predictor of success in late adolescence-early adulthood, through financial aid with access to higher education or the purchase of a home (Karagiannaki, 2017; Killewald et al., 2017; Pfeffer & Killewald, 2018). It has been proposed that wealthier families provide their children with more economic support than those from less affluent families and this is likely to continue well into adulthood (Pfeffer & Schoeni, 2016). It is also possible that family wealth is important for development very early in childhood, by enabling mothers to take longer periods of maternity leave, thus providing them with more time to interact with their child in the early months (Ulker & Guven, 2011). The earliest age at which we consider vocabulary ability here is age 3, and it is plausible that by this age the home learning environment plays an important role in language development, hence the increased predictive value of income compared to wealth at this age.

In line with our findings, Moulton et al., (2020) used the UK Millennium Cohort Study and found that wealth was not associated with language ability at age 11 after controlling for other indicators of SEC. 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 et al., 2019; Office for National Statistics, 2019b).

It is plausible then that in the UK, wealth is important for providing families with the opportunity to invest in homes in expensive areas close to high performing schools (Karagiannaki, 2017; Machin, 2011). The independent influence of wealth on age 14 vocabulary ability in our analyses may thus be a reflection of access to better secondary schools, as there is an established link between house prices and school quality (Department for Education, 2017a; Gibbons, 2012; Machin, 2011). International comparisons of the relative predictive value of different SEC indicators therefore have the potential to shed light on the mechanisms via which these SEC indicators are likely affecting language acquisition and inequalities.

When considering the role of income and wealth in language ability, it is worth noting that it is not clear exactly what our measure of wealth is capturing regarding mechanisms to language ability, and all proposed mechanisms for each indicator are speculation. Future research should seek to understand exactly how different SEC indicators influence vocabulary development. It is also worth noting that our income measure could be an underestimate of disposable income, as we only took income from the age 3 sweep of the data, in order to capture the influence of early childhood SEC on language development. Some suggest that income averaged over multiple sweeps gives a more accurate estimate of household income (e.g. Killewald et al., 2017). Our income measure is adjusted for household size, however our wealth measure is not; there is a general lack of agreement about whether family wealth should be adjusted for household size (Killewald et al., 2017).

There are a number of practical implications of these findings. Recent research has found the strength and significance of the SEC-language relationship differs based on indicator used, questioning the interchangeability of SEC indicators in this relationship (Gatt et al., 2020). However, they used maternal education, paternal education, maternal occupation and paternal occupation as their separate SEC indicators. Although parent education predicted the most variance in vocabulary, income and occupational status also uniquely predicted large amounts of variance. There are likely important pathways for each of these indicators that are important in understanding this relationship and targeting interventions effectively. Furthermore, it has been suggested that income is the component of SEC that is most susceptible to intervention (Duncan & Magnuson, 2012), and whilst targeting income as a means of reducing vocabulary inequalities may be somewhat effective, it is likely that a substantial amount of the relationship is driven by parent education and occupational status and thus targeting income may not fully reduce inequalities.

It has been suggested that parents may be more willing to provide information on their educational attainment and occupational status, rather than their income (Noble et al., 2007). Our finding that education, income and occupation explain similar amounts of variance in vocabulary ability at multiple ages suggests that an indication of inequalities can be determined excluding income as a measure of SEC. However, the fact that each explains unique variance also indicates that doing so would miss out on some relevant information. It is valuable, where possible, to have multiple predictors of SEC in order to gain the fullest possible picture of inequalities in language ability. The finding that relative neighbourhood deprivation, determined using IMD (a postcode-based measure of SEC) predicts unique variance up until adolescence is important; because it is based on an individual's postcode, it is easily and quickly calculated and does not rely on parent reports of information they may want to withhold (such as income). Therefore, using IMD as a measure of SEC will likely give some indication of the extent of inequalities in child language ability. However, the proportion of variance explained by this indicator is consistently small, and so if multiple measures can be obtained, they should be used.

Wherever practically feasible, utilising as many indicators of SEC as possible would be beneficial to gain a comprehensive picture of inequalities in vocabulary. Furthermore, where possible, it is advantageous to consider each indicator separately, rather than reduce these to a single composite measure. Although our analyses revealed a composite measure incorporating all indicators to be a better fit to the data compared to any single predictor alone, including all predictors as separate, individual indicators in a simultaneous model was the best fit to the data. It is not surprising that the composite measure was a better fit to the data over each single measure considered in isolation, since each contributed unique variance to the language outcome (with the exception of wealth at ages 3, 5 and 11 and IMD at age 14), so it is plausible that the combined effects of these will hold more explanatory value. However, it is useful to consider SEC indicators separately for some purposes, and their pathways to language ability should be considered carefully when assessing overall inequalities in language ability. A composite variable holds useful, practical value as it provides a single number estimate of overall inequalities in vocabulary, which can be attractive and easy to understand for lay readers, and is a simple, concise way of communicating such inequalities. However, a composite measure could obscure the different mechanisms of each single indicator driving inequalities in vocabulary, which may be of value for informing policies and interventions.

In sum, how one chooses to operationalise SEC should depend on the research question. For example, if one is interested in understanding the mechanisms by which SEC affects vocabulary, then multiple indicators are needed. However, if the goal is to understand overall inequalities in vocabulary, a composite measure may be the indicator of choice, as this provides a better estimation than any one predictor in isolation. If due to financial or practical constraints, data can only be collected on a single SEC variable, parent education was the best predictor, although was not dissimilar to income or occupation, and thus any of these could be used. Whilst of practical appeal, IMD predicted small amounts of variance in language, but explained unique variance and thus could be useful if it is the only option available.

SEC and vocabulary across developmental time

In the contemporary British cohort, inequalities in language ability widen between the ages of 3 and 5. This supports arguments for developing early interventions that seek to avoid inequalities becoming entrenched before children access formal schooling. Although not taken into account here, the widening inequalities observed between ages 3 and 5 could be partly explained by attendance at day-care settings. It has been suggested that attendance at a pre-school setting, such as day-care can benefit language development, particularly for those from disadvantaged backgrounds (Geoffroy et al., 2007; Magnuson et al., 2004; OECD, 2006; Schmerse, 2020). Becker (2011) found that pre-school attendance positively influenced the vocabulary of disadvantaged children, but did not affect the vocabulary of advantaged children. However, it was also observed that more advantaged children progressed more in their vocabulary development, regardless of pre-school attendance, thus attendance at a day-care setting did not provide a "catching up process" for lower SEC children; it appears as though the SEC gap in vocabulary is already engrained by the age of 3, and continues to widen despite improvements low SEC children may make at day-care settings (Becker, 2011). A possible implication of this finding is that interventions aimed at improving the quality of day-care for low SEC children may be beneficial as a means of targeting the SEC gap in vocabulary early on, as once it is established, the SEC gap in language appears to persist. Although RCTs of an intervention in the US to improve the responsiveness of early childcare educator have shown to be effective for the social and emotional development of toddlers, interventions appeared to have little impact on language and literacy outcomes, perhaps as a result of insufficient intensity (Landry et al., 2014, 2017, 2021). Our findings highlight that it is important to target inequalities in vocabulary early, before they become

entrenched by the age of 3 and worsen as children enter formal schooling. A two pronged approach whereby family support is provided at the same time as increasing the quality of provision in early years settings (Department for Education, 2017b; Gambaro et al., 2015) could be beneficial. Regarding the first prong, it is important 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)). In addition to parental support, quality pre-school provision has been found to benefit 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). It is therefore also likely important to test ways of improving the consistency and quality of pre-school education to help prevent inequalities becoming entrenched before entry to 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, than oral language (Elleman et al., 2019). It is plausible that these sources of input are influenced by SEC. For example, the availability of books and vocabulary-rich online content may be higher among higher SEC children. 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 et al., 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 et al., 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

school work (Jerrim, 2017). Additionally, children are likely to form peer groups with those that are similar to them (McPherson et al., 2001), a phenomenon which has been observed for social class, with children and adolescents being more likely to form friendships with peers in the same social class (Cherng et al., 2013; Kupersmidt et al., 1995; Lessard & Juvonen, 2019; Malacarne, 2017). Furthermore, it has been observed that children form friendships with peers who have similar levels of intelligence to them, specifically in terms of vocabulary (Boutwell et al., 2017). It is therefore possible that children with larger vocabularies and from more advantaged backgrounds will form peer groups with similar children, potentially providing a richer vocabulary environment for these children, and likewise for children with smaller vocabularies or from lower SEC backgrounds, emphasising inequalities in language ability throughout childhood.

Schooling may also be contributing to the SEC gap remaining wide throughout childhood and into adolescence. Children are often placed into 'sets' based on their ability in both primary and secondary schools (Hallam & Parsons, 2013; Muijs & Dunne, 2010), and because language lays the foundation for literacy and maths skills (Lervåg et al., 2018; Moll et al., 2015; Slusser et al., 2019; Snowling & Hulme, 2020), it is likely that those with poorer language skills could be placed into lower sets and thus targeted with lower-level ability education, that may not stimulate the same level of thinking as in higher sets (Anders & Henderson, 2019; Henry, 2015; Muijs & Dunne, 2010). Furthermore, there may be more disadvantaged pupils in lower sets compared to higher ability sets (Wiliam & Bartholomew, 2004). It is thought that parents will tailor the language they direct with their children to the perceived language abilities of the child (Schwab & Lew-williams, 2016) and teachers may also engage in this process, thus children of more disadvantaged backgrounds and in lower ability sets may not benefit from the same vocabulary exposures as their more advantaged peers. SEC can also influence the type of school children attend, with higher SEC schools having better resources and more experienced teachers (Willms & Somers, 2001; Palardy, 2008).

Children from disadvantaged backgrounds may require more input at school to support the development of an advanced vocabulary (Sullivan et al., 2021). However, because the type of school attended and level of teacher support available differs based on SEC, lower SEC children may not have as much of this support available Universal education aims to address inequalities in educational opportunity in the UK, but, when it comes to vocabulary, disparities persist throughout formal schooling. Further support across the

lifespan and particularly during adolescence is likely necessary to improve educational outcomes and open up employment opportunities (Deloitte, 2016).

SEC and vocabulary across historical time

Finally, the cross-cohort comparison performed suggest that inequalities in childhood language are generally similar across generations, despite decades of policy to reduce these inequalities.

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 child language in contemporary society. 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. Alternatively, it might be that the relative importance of the various proximal causal mechanisms themselves is changing with time.

Findings revealed that more variance is explained by both parent occupational status and parent educational attainment in the MCS2001 cohort compared to the BCS1970 cohort for early childhood vocabulary and adolescent vocabulary. However, for late childhood vocabulary ability, more variance is explained by both predictors in the BCS1970 compared to the MCS2001. There may be additional factors at the end of primary schooling that are important for vocabulary ability, including SEC indicators not included in the cross-cohort comparison, such as income, that are more important for the MCS2001 cohort than the BCS1970 cohort. Income inequality increased between 1970 and 2000 in the US, and although there were gains in household income across the distribution, these were largest for those at the top of the distribution (Pew Research Center, 2020). Over a similar period in the US, increases in parental financial investments on children, in terms of educational toys and activities and daycare, have also been reported (Kornrich & Furstenberg, 2013), with some evidence that these increases are concentrated at the top of the income distribution, meaning that as overall spending on children increased, so did inequalities in such investments (Kornrich, 2016).

Relatedly, across states in the US and over the period of 1975-2014, it was found that as income inequality increased, so did inequalities in parental investments in their children, which was explained in part by the increased income at the top of the distribution, with these parents having more disposable income for investing in their children. Importantly, general spending did not seem to increase, rather this was isolated to investments in children, perhaps reflecting increases in pressure to invest in children among areas of higher income inequality (Schneider, Hastings, & LaBriola, 2018). Similar increases in income inequality have been observed in the UK since the 1980s (Dorling et al., 2007.; Institute for Fiscal Studies, 2017; Rigby, 2017), and it is possible that corresponding increases in parental investments in children, similar to that in the US, have also occurred, perhaps increasing the importance of income as a predictor in the MCS2001 cohort. However, since a measure of household income in the early childhood years is not available for BCS1970 data, we cannot test the increasing relative importance of household income between these two cohorts.

We can only provide speculative explanations of this pattern of results, based on historical and contextual evidence of UK society. The UK has seen major changes to its occupational structure during the period that separates the two cohort studies used in our analyses. For example it is thought that the economy is becoming an hourglass structure, with more people in the highest and lowest classes, and fewer people in the middle classes (Holmes & Mayhew, 2012). Table 6 reveals a pattern in line with this for our two cohorts, with almost half of the BCS1970's parents falling into intermediate occupations, compared to 20% of the MCS2001's parents, and more MCS2001 parents having higher professional or routine occupations. There cannot be a simultaneous rise in the relative advantage that comes with higher class membership as the proportion of people within this higher class increases (Blackburn & Jarman, 1993) and so the advantage associated with having a high occupational status (higher managerial occupations) may be conferred as higher in the BCS1970 due to these occupations being less common.

In addition to this occupational shift to an hourglass economy, women became increasingly active in the workforce between 1971 and 1999 with many occupying flexible, paid part time jobs (Amaranayake et al., 2000). Inequalities in language ability appear to be wider among the BCS1970, based on parent's occupational status. The increased participation of women in the workforce may have increased social networks, which tend to consist of people who are similar to them in terms of occupational status (Griffiths & Lambert, 2012; Griffiths, Lambert, & Tranmer, 2011; McPherson, Smith-Lovin, & Cook,

2001). We therefore may see smaller inequalities in language based on occupational status in the MCS2001, as these families may have had a wider social network - children will adopt styles of speech and vocabulary used by their parents, when talking to them and when talking to individuals in their social network (Sullivan, 2007). In contrast, many mothers of the BCS1970 members were housewives (Institute of Child Health, 1970), and although the father was likely to be in employment, it may be more likely that they had smaller social networks. Indeed, by 1996-97, ~60% of couples in the UK with children had both parents in employment (Amaranayake et al., 2000). The national minimum wage was introduced in the UK in 1999, and although this is more directly related to income rather than occupational status, this policy may have benefitted the parents of the MCS2001 and subsequently be reflected in the effects of occupational status on language in these cohort members. Parents of a lower status may have benefitted from the introduction of the minimum wage, although we still see inequalities in language based on this measure in the MCS2001, these are just narrower than they were for the BCS1970.

The education system in the UK has also undergone significant change and expansion. Parents of the BCS1970 cohort were mostly in school when the compulsory leaving age was 15; this was raised to 16 in 1972 and therefore the parents of the MCS2001 had to remain in education until the age of 16. Inequalities in early childhood language and adolescent language were widest between those with no level qualifications and O-level qualifications in the BCS1970. This may be because the school leaving age was 15 for these parents, and perhaps there were characteristics of those that stayed in education after the compulsory age (for example they may have been more advantaged). This could be why we see larger inequalities between these two groups. Less advantaged students have consistently been under-represented in higher education (Mayhew et al., 2004; Woodward, 2019). Furthermore, for the BCS1970, around half of these parents had no qualifications and so the difference in inequalities of language ability would be present between those with no qualifications and O-level qualifications and inequalities may level off for higher groups in this cohort compared to the MCS2001. For parents of the MCS2001 cohort, these inequalities may manifest at different education levels, as these parents were more likely to stay in education to complete O-levels, but then not all of them will have continued to post-16 education. Further, because the school leaving age was 15 for the BCS1970 parents, many of these have no qualifications, but there were still a large proportion of this cohort in at least intermediate occupations. So, perhaps for this older generation, occupational status is a better indicator of inequalities in language ability, as the majority of fathers were in some form of

employment, despite smaller proportions of qualifications among parents in the older cohort; educational attainment may not be as relevant or as much of a reflection of standing in society as occupational status, as only a very small proportion stayed in education.

Based on the average age of the mother at the birth of cohort members, we can infer that the parents of the BCS1970 who attended university would likely have started in 1960-61, and this would have been between 1986-1988 for the MCS2001 parents. Higher education in the UK has seen massive expansion in recent decades, for example, between 1960 and 1970, where attendance increased from 5% to 14%, and reached 34% in 1997-98, with increases being dominated by those from higher SEC groups, and lower SEC groups remaining less likely to attend university (Mayhew et al., 2004; Woodward, 2019). It is thought that social inequalities in access to higher education increased in the 1980s and early 1990s, around the time that the parents of the MCS2001 would have attended university (Blanden & Machin, 2004; Chowdry et al., 2013; Galindo-Rueda et al., 2004; Machin & Vignoles, 2004) and this could partly explain why we see wider inequalities in language ability for children whose parents have university level qualifications in the MCS2001.

Access to higher education was more restricted for the BCS1970 parents, heavily contingent on gender and social class (Blackburn & Jarman, 1993). Although many of these mothers were likely at home with their child, perhaps they didn't have the skills from educational attainment that are important for language development, and this could be why we see wider inequalities at lower levels of education in the BCS1970. Perhaps we see wider inequalities based on university level qualifications in the MCS2001 due to the widening of participation in higher education. Furthermore, it is likely that there was an interaction between occupational changes and education expansion in the UK; as the number of graduates has increased, having a degree has become a more common entry requirement for many high-level occupations (Blackburn & Jarman, 1993), meaning there are less opportunities available for those with lower levels of educational attainment. The increasing demand for graduates among employers likely resulted in less opportunities for non-graduates, whereas the parents of the 1970 cohort likely didn't need a degree to be successful in at least intermediate occupations (Keep & Mayhew, 2004).

Some of our cross-cohort findings might be attributable to assortative mating, specifically educational homogamy. Assortative mating refers to the tendency for people to partner with others who possess similar characteristics to them, for example, in terms of educational attainment (educational homogamy; where people marry others with similar educational attainment levels). This is a well-established phenomenon and evidence points to

an increase in educational homogamy, particularly among those with degree level qualifications, across the period that separates the two birth cohorts studied here (Blackwell, 1998; Kalmijn, 1991a; Kalmijn, 1991b, Blossfeld & Timm, 2003). This has been attributed to the expansion of the education system in the UK , as such an expansion increases chances of meeting potential partners with similar educational attainment levels (Blossfeld & Buchholz, 2009). Historically, women have tended to display upward marriage tendencies because the average attainment level of women was inferior to that of men, and women usually stayed at home. However, as women have become increasingly active in education, this has decreased (Blossfeld & Buchholz, 2009). Furthermore, as language skills are the foundation for educational attainment and due to educational homogamy, it is conceivable that those with more advanced vocabularies are more likely to partner with someone who has a similarly advanced vocabulary (Sullivan et al., 2021). Thus, we could attribute our finding of more advantages for language ability among those with parents who have university level qualifications in the MCS2001 compared to the BCS1970 to increases in educational homogamy. Specifically, if educational homogamy has increased and this is particularly the case among those with university level qualifications, then it is plausible that in the MCS2001, where one parent has university level qualifications, both parents may have university level qualifications, meaning there are households where both parents have these high qualifications in this cohort. In contrast, in the BCS1970 where a parent has a university level qualification in this cohort it is likely to be the father and only one parent having such qualifications. There may be some benefit for language ability associated with two parents having degree level qualifications: because language skills are associated with educational attainment, it is likely that two parent households with university level qualifications will also have large vocabularies, and thus children of such households will be at an advantage due to being exposed to more advanced vocabularies , as well as from overhearing conversations parents have among themselves (Sullivan et al., 2021).

Although we can only propose these explanations based on contextual evidence, the implication of our cross-cohort research is clear. We found quite different patterns of results in the two cohorts based on whether we used occupational status or parent education as our indicator of SEC. There have been many changes to the occupational and educational structure in the UK between these two decades, with occupational status likely being a more important indicator of standing in society in the older, BCS1970, and educational attainment increasing in importance as an indicator as more people have gained higher qualifications in the UK. Therefore, it is important to use multiple indicators when researching the cross-

cohort relationship between SEC and language ability as a different pattern of results emerges based on the indicator used. However, in both cohorts it is evident that there are inequalities in language ability when using each indicator, suggesting that inequalities in language have not reduced or disappeared over historical time.

Limitations and strengths.

We used measures of single dimension of language ability — vocabulary — at multiple time points. This 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). This allowed us to investigate the extent of inequalities at each age considered, and by using a standardised score, we were able to make comparisons that reflect population distributions in these vocabulary outcomes. However, 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. For the cross-cohort comparison, we were limited to the two SEC indicators available in both cohorts – occupational status and educational attainment – and the operationalisation of these indicators might mean slightly different things in the two cohorts. However, when we conducted a sensitivity analysis using Ridit scores as a means of standardising SEC indicators, this revealed a similar pattern of results. 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). 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 and occupational status all explain substantial unique variance in child language. This suggests it is well worth testing the causal effects of supporting parent education (through

lifelong learning) and/or caregiver understanding 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, 22% of the UK population and 30% of children were living in relative poverty (after housing costs) in 2018-19 (Francis-Devine, 2020). Low 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, 2020a), and our cross-cohort comparison revealed inequalities in vocabulary are persistently wide across time, targeting these sustained inequalities may 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. Similarly, inequalities in vocabulary remain wide throughout childhood and the relative advantage of higher SEC 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 social activities, seeking out effective ways to support adolescent language development is important (Bercow, 2018; Spencer et al., 2012).

Finally, 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 (Greenwood et al., 2020; 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). Further, a 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 et al., 2019). However, the benefits for cognitive outcomes are currently less clear (Melhuish et al., 2010), perhaps because of a struggle to reach populations who would have derived the maximum benefit (Law, Parkin, & Lewis, 2012). The current analyses suggest that existing policies and interventions have not reduced social inequalities in language ability in recent decades, suggesting more needs to be done (for example, a

multi-pronged approach implemented at a meaningful scale), so as to reap sustained benefits and see the next generation of children reach their potential.

4.6. 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 parent education, income and occupational status. A composite measure of SEC can be useful if one's main objective is to understand overall inequalities in language ability. However, this may obscure possible mechanisms underlying the relationship. 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 reduced poverty, increased caregiver lifelong learning, improved early parenting support, improved quality of preschool education and sustained educational support through 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.

Chapter 5 : Socioeconomic circumstances moderate the predictive value of 5-year-old vocabulary for education outcomes at the end of secondary school: analysis of the UK Millennium Cohort Study

5.1. Abstract

Language ability is thought to be important for educational attainment as it lays the foundation for literacy and numeracy skills. However, this relation is confounded somewhat by the existence of socio-economic inequalities in both vocabulary and in educational attainment. It is not known whether vocabulary skill has any predictive value over and above socio-economic factors. Furthermore, the benefits of having a strong early vocabulary may be felt more strongly by some groups than others. Data from a large, nationally representative cohort across the UK (MCS2001; N= 15, 576) were analysed to investigate the role of age-5 vocabulary in predicting public examination performance at age 16. Vocabulary uniquely predicted educational attainment, above socio-economic circumstances (SEC) and caregiver vocabulary factors (OR = 1.61, 95% CIs = [1.50;1.72]; (β = .24, 95% CIs = [.21;.27]). SEC was found to moderate the relation between vocabulary and achievement, such that a high SEC background increased the predictive value of early vocabulary for achievement (OR = 1.97, 95% CIs = 1.85;2.1]). This moderating relation is not straight forward however, as the impact of increased vocabulary is greater in the middle SEC groups than in the highest SEC group, where educational attainment is high even in children with very low vocabulary scores. The impact of SEC on the relation between vocabulary and educational attainment is greater in both these groups relative to the lowest SEC group, where educational attainment is poor, regardless of vocabulary. Findings suggest that improving vocabulary early in childhood may improve education outcomes, but that to be most effective, policies and interventions need to be targeted at those who need them the most.

5.2. Introduction

Language skills, and in particular vocabulary, are thought to be an important contributing factor to educational success (Field, 2010; Oxford University Press, 2018, 2020; Tickell, 2011); Children need good language skills in order to access the curriculum, to exchange their thoughts and ideas, and to indicate that they understand what they are being taught (Alexander, 2020; Nagy & Townsend, 2012). In turn, educational success is important and poor educational attainment can have detrimental impacts on later life, such as poverty and worklessness, and the inability to support the learning of future offspring (Centre for Social Justice, 2013). Educational attainment is therefore seen as a key driver of social mobility. However, poor educational attainment is a barrier preventing disadvantaged people moving out of poverty (ESRC, 2012), as there are socioeconomic inequalities in educational attainment, with those from disadvantaged background performing worse than their more advantaged counterparts. Socioeconomic inequalities in vocabulary are also well established (Hart & Risley, 1995b; Pace et al., 2017). Our goals in this research were to a) investigate the relationship between early vocabulary and educational attainment at the end of secondary school, controlling for the effect of SEC; and b) to establish how much early vocabulary matters for educational attainment in different socioeconomic groups.

The fundamental purpose of education is threefold: ensuring that young people acquire the knowledge and skills required to succeed in a demanding economy and to contribute to society, the enabling of cultural participation and, practically, preparing people for adult life (Gibb, 2015). Further, education should facilitate one's understanding of and motivation to undertake further study, in addition to the ability to make informed decisions throughout life, including those about education and employment (Centre for Social Justice, 2013; Wellcome Trust, 2016). In the United Kingdom, students take academic qualifications in the form of public examinations, at the end of secondary school. In England, Wales and Northern Ireland, General Certificate of Education (GCSEs) are taken at the age of 16 and in Scotland, National Five (N5) qualifications are usually taken between the ages of 15 and 16. Success at this stage of education is seen as the gateway to post-16 success, whether this be further education, an apprenticeship, or entry to the workforce (The Children's Comissioner, 2019; The Sutton Trust, 2019). Further, success in specific subjects at GCSE can impact the subjects that students can take at A-level and which colleges they can attend. GCSE attainment is thought to be related to achieving predicted A-level grades (UCAS, 2016) and Crawford, Duckworth, Vignoles, & Wyness, (2011) found an association between poor GCSE achievement and NEET (not in education, employment or training) status. Those

leaving school without these basic qualifications, especially in English and Maths, therefore struggle in post-16 destinations and may also lack the confidence to make decisions surrounding their future, in addition to not meeting basic entry requirements for many apprenticeships and jobs (Lupton et al., 2021; The Children's Comissioner, 2019). It is clear, therefore, that success at the end of secondary school is an important stepping-stone for later outcomes.

Language ability and educational attainment

Early language ability has been shown to be important during schooling and for educational attainment. A relationship between early language and early educational attainment has been documented (Bleses et al., 2016; Durham et al., 2007; Lehrl et al., 2019; Morgan et al., 2015). For example, Roulstone, Law, Rush, Clegg, & Peters (2010) found that language skills at the age of 2 were predictive of achievement at school entry, and Pace, Alper, Burchinal, Golinkoff, & Hirsh-Pasek (2019) reported that language skills were predictive of later achievement across a range of subjects in the 3rd and 5th grades. Early language skill has also been shown to predict later educational attainment (Croll, 1995; Einarsdóttir et al., 2016) and adolescent vocabulary ability has also been linked to GCSE attainment (Spencer et al., 2017). Furthermore, research with populations identified as having Developmental Language Disorder has highlighted the importance of language ability for educational achievement at the end of secondary school and in early adulthood (for example, Conti-Ramsden, Durkin, Simkin, & Knox, 2009; Conti-Ramsden, Durkin, Toseeb, Botting, & Pickles, 2018; Dockrell, Lindsay, & Palikara, 2011; Snowling, Adams, Bishop, & Stothard, 2001).

Vocabulary forms the basis for literacy and mathematics development, which are key for later academic achievement (Lervåg et al., 2018; Moll et al., 2015; Slusser et al., 2019; Snowling & Hulme, 2020). For example, good language skill is needed to understand mathematical vocabulary and concepts (Slusser et al., 2019) and vocabulary is essential for reading comprehension, as a reader must understand the words to be able to understand the meaning of the text; difficulty with reading is likely to be a further barrier to accessing to the curriculum (Elleman et al., 2019; Ricketts et al., 2020). Furthermore, the majority of teaching is delivered through the medium of language: lessons tend to be delivered orally by teachers, and children use language (both written and oral) to interact with teachers and engage in lessons, and to communicate their understanding of, or difficulties with content (Alexander, 2020). Language is vital for the whole curriculum, as each subject has specific jargon

associated with it (known as Tier 3 vocabulary), and the understanding and learning of a subject is therefore contingent upon the language of that subject (such as the vocabulary of that subject) (Alexander, 2020; Oxford University Press, 2020).

Socioeconomic inequalities in educational attainment

There are socioeconomic inequalities in vocabulary that appear before schooling and persist throughout childhood and into adolescence (Pace et al., 2017; Sullivan et al., 2021; Thornton et al., 2021). There are also socioeconomic disparities in educational attainment, whereby those from less advantaged backgrounds perform worse than their more affluent peers (Centre for Social Justice, 2013; Chowdry et al., 2009; Crenna-Jennings, 2018; Hegedus, 2018; Ofsted, 2013; Pensiero & Schoon, 2019). This relationship has been estimated to be of a weak-moderate magnitude in a series of meta-analyses in the United States of America (Harwell et al., 2017; Sirin, 2005) and a recent meta-analysis suggested a weaker relationship between socioeconomic inequalities and educational attainment in developing countries (Kim et al., 2019).

Turning to the United Kingdom specifically, the situation is worse when compared to many other OECD countries: at all stages of schooling, children from lower SEC backgrounds fall behind their higher SEC peers, with the gap widening with progression through education (Crenna-Jennings, 2018; Jerrim et al., 2016). In the 2017 GCSE examinations, pupils from disadvantaged backgrounds had lower performance than their more advantaged peers (Department for Education, 2018d). Pupils eligible for free school meals are more likely to reach the end of free education at the age of 19 without the basic benchmark of qualifications (5 GCSEs grades A*-C/ 4-9) and it is thought this gap is widening further (The Children's Comissioner, 2019). SEC also appears to interact with ethnicity in affecting educational outcomes, with all ethnic minority groups of a low SEC background achieving higher than their disadvantaged White ethnicity counterparts at age 16 (Strand, 2011, 2014); for example, in the 2018-19 examinations, only 35.9% of disadvantaged White British pupils reached a key government benchmark of achieving a pass in their English and Mathematics GCSE, compared to 57.4% of disadvantaged Asian pupils, and 48.8% of disadvantaged Black pupils (Department for Education, 2019; House of Commons, 2021). There are also socioeconomic inequalities in the subjects chosen at GCSE level, with those from more advantaged backgrounds tending to study subjects that are sometimes considered to be more demanding, such as science, and less likely to take applied

subjects (Anders et al., 2018; Henderson et al., 2018).Subject choice at GCSE can affect the move to further education and employment —for example, success at GCSE influences whether that subject can then be taken at A-level (Henderson et al., 2018), which may therefore exacerbate socioeconomic inequalities in educational attainment.

SEC may lead to differentials in educational attainment in a multitude of ways. Children of parents with more educational qualifications may benefit more from parental input with school work, due to their parents having the ability to do so: children from disadvantaged backgrounds experience less parental input with homework (Jerrim, 2017). Adults with poor literacy skills will be poorly placed to help their children with school work, and in 2011, 14.9% of adults in England had literacy skills at the level expected of that of a 9-11 year old (Department for Business Innovation and Skills, 2012). It is estimated that 9 million adults aged 16-65 in England have low literacy levels (OECD, 2016a) and those with more educated parents tend to have higher literacy skills than those whose parents have no upper secondary qualifications (OECD, 2016b). It is therefore plausible that less educated parents want to help their children, but simply lack the necessary skills or feel inexperienced and inadequately placed to do so (House of Commons, 2021; Koshy et al., 2013; Treanor, 2017). Further, pupils from lower income households may lack the resources needed to succeed in education, such as access to a computer and the internet, which has been found to be associated with an increase in GCSE point scores, after accounting for prior attainment (Chowdry et al., 2009). Relatedly, more advantaged parents are more likely to provide homework support in terms of additional hours spent on school work and through private tuition (Jerrim, 2017). Private tutor provision is thought to be accelerating socioeconomic inequalities in educational attainment – it is estimated to cost between £24 - £32 per hour, and disadvantaged students are less likely to benefit from this, with affordability being the main reason (Kirby, 2016). Participation in extracurricular activities throughout school has also been linked to stronger educational attainment, such as being a member of a sports team, playing a musical instrument or being part of an arts or drama club. Low income is a likely barrier to participation in such activities (Donnelly et al., 2019).

Turning to school level factors, children from more disadvantaged backgrounds are likely to suffer from unconscious teacher bias, where they are judged less favourably by teachers than their more advantaged counterparts. For example, 7 year olds from low income families tended to be underrated by teachers in their perceptions of mathematics and reading ability (Campbell, 2015). Those of lower SEC are more likely to be placed in lower ability sets, which may coincide with poorer quality teaching and lower-level ability education that

may not stimulate the same level of thinking as in higher sets and may even negatively impact educational attainment (Henry, 2015; Muijs & Dunne, 2010). More advantaged parents may have access to higher quality schools, and high SEC children are more likely to attend private school than their lower SEC counterparts (Dearden et al., 2011; Moulton et al., 2021), with private and independent schools benefitting from smaller class sizes, more resources, increased support from teachers, and better outcomes for students compared to state funded schools, in terms of higher exam results, attendance at Russel Group universities and prestigious labour market positions (Henderson et al., 2020; Independent Schools Council, 2018). Pupil Premium (additional funding for schools to help disadvantaged students regardless of ability) has the aim of closing the attainment gap by investing in, for example, high quality teaching and the provision of additional support, such as one to one or small group tuition, and speech and language therapy. However, the amount of pupil premium a school receives is based on the proportion of pupils *registered* as eligible for free school meals in the previous 6 years, not the proportion of pupils who are eligible for free school meals, and due to a lack of awareness or stigma associated with free school meals, many eligible parents are not registered (around 1 in 10), meaning some schools are likely receiving less pupil premium than they are entitled to (Social Mobility Commission, 2019). Furthermore, schools in areas of disadvantage or with higher proportions of disadvantaged pupils tend to have less experienced teachers and fewer teachers qualified in specific subjects, so that pupils are more likely to experience poorer quality teaching, compared to more advantaged schools (House of Commons, 2021; Shaw, Baars, Menzies, Parameshwaran, & Allen, 2017), and for lower SEC pupils, the disparity between being taught by a good teacher and a bad teacher is equivalent to a year of learning (The Sutton Trust, 2011).

Language ability is also a likely mechanism by which SEC exerts effects on educational attainment (for example, Bukodi, Bourne and Betthäuser, 2017; Bourne et al., 2018; Slusser, Ribner and Shusterman, 2019; von Stumm et al., 2020). Individuals are more likely to form friendships with peers who display similar levels of intelligence, particularly regarding vocabulary (Boutwell et al., 2017). It is therefore possible that children with larger vocabularies and from more advantaged backgrounds will form peer groups with similar children, potentially providing a richer vocabulary environment for these children, and likewise for children with smaller vocabularies or from disadvantaged backgrounds. In adolescence, interactions with friends is the main exposure to vocabulary acquisition (Oxford University Press, 2020). It is thought that the need to 'code switch' between the language

used with friends and that required in formal education (which is more congruent with the language use of higher SEC children) can disadvantage the educational achievement of those whose language use is most different from that used in school, such as those from more deprived backgrounds (Bernstein, 1964; Oxford University Press, 2020; Spencer et al., 2013). Relatedly, non-academic routes may be more valued than academic qualifications among lower SEC groups, which may result in these pupils being disengaged in and undervaluing academic education (House of Commons, 2021). Those from a more disadvantaged background may hold fewer aspirations for their educational achievement than those from a higher SEC background , for example, in terms of pursuing post-16 education (Chowdry et al., 2009). Vocabulary difficulties may lead to poor self-esteem among adolescents (Oxford University Press, 2020), which could in turn lower aspirations of low SEC pupils further. Low SEC pupils may also have lower beliefs about their own ability, and about the control they hold over their educational outcomes, which has been linked to success at GCSE (Chowdry et al., 2010).

In an influential study for British policy, it was found that children of more affluent parents who had low scores in cognitive tests at the age of 22 months were more likely to perform better at the age of 10, than were a comparison group of children from disadvantaged parents (Feinstein, 2003). However it is worth noting that this study may have suffered from regression to the mean (Jerrim & Vignoles, 2013). Additionally, it has been suggested that socioeconomically disadvantaged individuals with high levels of cognitive ability tend to perform worse on cognitive and educational tests (of English and Mathematics, for example) and show less educational growth throughout school, than those from more advantaged backgrounds of a similar level of, or poorer, ability (Crawford et al., 2017; von Stumm, 2017).

Improvement in the language abilities of young children has been targeted with the aim of improving educational outcomes, via initiatives such as *Every Child a Talker* (Department for Children Schools and Families, 2009). The Education Endowment Foundation (EEF), which aims to break the relation between SEC and educational attainment, to allow every student to reach their full potential, has found oral language interventions to be effective in improving reading and mathematics abilities, equating to 6 months progress for those who received the interventions. However, the majority of these studies focused on reading outcomes in primary school (Education Endowment Foundation, 2021). More recently, the Nuffield Early Language Intervention (NELI) has been successfully delivered at scale to improve language in reception aged children, and is now being offered across

primary schools in England (West et al., 2021). The long-term effectiveness of such interventions is largely unknown, and where intervention studies have included long-term follow ups, effects on oral language have reduced (Law, Charlton, Dockrell, et al., 2017). Furthermore, the long-term effects of such oral language interventions on later educational attainment, especially in secondary school, remain largely unknown: might improving language ability in early childhood lead to higher educational attainment in secondary school? In order to establish whether early interventions can have knock-on benefits lasting to secondary school outcomes, it is important to first establish whether early language is related to educational attainment at the end of compulsory schooling (16 years).

The current study

We know that there are SEC inequalities in vocabulary and that there are SEC inequalities in educational attainment. However, we do not know whether vocabulary predicts educational attainment above and beyond SEC in early childhood. Further, it could be that vocabulary matters more for educational attainment in some SEC groups compared to others: it is possible that those of a higher SEC with lower vocabulary ability can draw on mechanisms, such as parental input and tutor provision (which are more likely to be available in higher SEC families), to protect them against the negative effects of poor language on educational attainment – something which those from a lower SEC are unable to do, and therefore may be more reliant on their vocabulary skill to succeed. Although vocabulary is already associated with SEC by the age of 5 (Thornton et al*.,* 2021), whether or not these initial individual differences in vocabulary have long lasting impacts, and whether these impacts are felt more by some than others, is less clear.

Using a large, nationally representative, contemporary British birth cohort, we investigated three research questions in a series of pre-registered analyses: 1) Does early childhood vocabulary predict a) whether or not cohort members achieve a functional level in core subject examinations at the end of secondary school and b) the level of achievement (regardless of pass/fail) of cohort member in these core subjects? 2) Does any such relation hold over and above SEC and caregiver vocabulary factors? and 3) Is the relation between age 5 vocabulary and attainment at the end of secondary school moderated by parent socioeconomic circumstances?

5.3. Method

Data

Data from the MCS2001 were used in these analyses, 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), which enabled us to look at the effects of vocabulary on achievement at the end of secondary school across the United Kingdom.

Sample selection. We selected all cohort members with either a response on the vocabulary measure at age 5 *or* with an educational attainment outcome. Where there are multiple cohort members from the same family, one of these was selected at random for inclusion in our sample. This resulted in a total sample of 15,576 cohort members.

Measures

Predictor variable: Age 5 vocabulary: Naming Vocabulary BAS II subscale (Elliott, Smith, & McCulloch, 1996).

At age 5, MCS2001 cohort members completed the Naming Vocabulary BAS test, as a measure of their expressive language (see Chapter 2). Cohort members were shown a series of images, one at a time, and 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 also born over a 1.5-year period (September 2000-January 2002) and assessed over a range of months, so their age at the time of testing may differ between cohort members. Therefore, we used *t*-scores adjusted for item difficulty and age. These were converted to *z* scores for analyses.

Outcome variables: Educational attainment at the end of secondary school.

Educational attainment in the United Kingdom

In England, Wales and Northern Ireland, General Certificate of Education (GCSEs) assessments are taken at the age of 16 and in Scotland, National Five (N5) qualifications are usually taken between the ages of 15 and 16. GCSE qualifications in chosen subjects are studied during the school years 10 and 11. English (language and/or literature), mathematics and science (and Welsh for those taking their GCSEs in Wales) are compulsory GCSE subjects. Levels of achievement in English, mathematics and science are often used as

indicators of success (for example at the level of the individual, the school, nationally and even internationally). The OECD's Programme for International Student Assessment (PISA) investigates the literacy, mathematics, and science abilities of 15-year-olds in OECD countries. Further, these subjects are defined as core subjects that are compulsory throughout every key stage of education, up to and including key stage 4 (i.e., GCSE) (Department for Education, 2014). Skills developed in these subjects underpin achievement in all other subjects and success later in life - for example, skills developed in science are important for understanding the world, particularly in a society that is becoming more technological (Department for Education, 2014). Scotland's education curriculum is known as the curriculum for excellence (CfE) for children aged 3-18, which has four fundamental goals to help pupils become successful learners, confident individuals, responsible citizens and effective contributors (Education Scotland, 2021). There are two phases to the CfE: broad general education (from early childhood education until the end of the third year in secondary school) and senior phase (fourth to sixth years in secondary school, i.e., aged 15-18). Among the 8 subject areas that make up the CfE, literacy and numeracy are seen as being of particular importance. The senior phase aims to offer the chance to achieve through studying qualifications and awards, and builds on the education obtained during the broad education, with the aim of allowing pupils to take subjects that suit their abilities and interests (Education Scotland, 2021). Progression from broad general education to the senior phase varies based on achievement at the end of the broad general education phase, and pupils can take a combination of qualifications in different subjects, based on their level of prior achievement (the National Parent Forum of Scotland, n.d.). Qualifications in the senior phase include National Four (internal added value assessments which are teacher graded as pass or fail), National Fives and Higher Grades.

A common GCSE benchmark across England, Wales and Northern Ireland is the achievement of 5 GCSEs grades A*-C (or 9-4 under the new grading system), usually including English and Maths. A pass of grade C or above in science subjects is also a requirement for entry to some university courses. In England, as of 2014, those failing to reach the benchmark in English and Maths are now required to resit these subjects as part of their post-16 education (Velthuis et al., 2018). Despite the emphasis on this benchmark, in one study, not all students failing to reach a least a grade 4 in English and Maths had poor attainment generally; 25% successfully gained 5 GCSEs grade 4 and above and 44% attained at least a grade 4 in one of the two subjects (Velthuis et al., 2018), yet they may still face barriers in their post-16 transitions. Among students of the same ability, those who fail to

reach the benchmark are more likely to drop out of education by the age of 18 and are less likely to enter into A-level or equivalent courses, or university, compared to those of the same ability who did reach the benchmark (Machin et al., 2020). We therefore focused on the attainment of this benchmark in core subjects, given its important for post-16 transitions.

In England, reforms to GCSEs were introduced in 2015, with the first reformed exams taken in 2017. These reforms were characterised by more challenging content, transitioning to linear qualifications (where all examinations are sat at the end of year 11), and the introduction of a new numerical grading system (9-1, with 9 being the highest grade) (Ofqual, 2017). In some subjects, marks are now allocated for correct spelling, punctuation and use of grammar (Ofqual, 2013). MCS2001 cohort members sat their GCSE examinations ~2017 and are among the first to have taken the new GCSEs in the English and Mathematics subjects, so using this cohort enabled us to investigate the effects of early vocabulary in a cohort that have taken the more challenging GCSEs in England. These reforms were introduced to GCSEs in England only: although pupils in England, Wales and Northern Ireland all sit GCSE qualifications, there are some differences between these qualifications. While GCSEs in England are now linear, Wales and Northern Ireland have retained the modular structure of GCSE exams, in addition to the old A*-G grading system. Across the three countries, the content and assessment objectives of specific subjects differs, however all involve students using their knowledge, understanding and skills of the subject material (AS and A Levels in England , Wales and Northern Ireland GCSEs in England , Wales and Northern Ireland, 2017). Although different qualifications are taken across these countries, harmonising these into one measure provides a valuable opportunity to look at the role of age 5 vocabulary in educational attainment across the UK: although the qualifications in each country are inherently different, they are each important for future success in their respective country, for example in terms of entry to post-16 qualifications and the labour market. It is important to establish the role of vocabulary on success in these qualifications as a possible route for improving educational attainment.

Previous research that has investigated the role of language in later educational outcomes has focused on language ability in middle childhood or adolescence (for example, Bukodi, Bourne, & Betthäuser, 2017; Spencer et al., 2017), or has conceptualised attainment in different ways - for example, vocabulary score, mathematics ability or reading ability (von Stumm et al., 2020). Those that have focused on success at GCSE have largely looked at the number of GCSEs grade A*-C, a binary variable of whether or not 5 grades A*-C were achieved, or a conversion of grades to a numeric score (when the old GCSE system was used)

and a continuous total score calculated, whereby a higher score indicated higher achievement (Bukodi et al., 2017; Conti-Ramsden et al., 2009, 2018; Spencer et al., 2017). However, it is important to note that these conceptualisations tend to disregard the fact that some subjects may be more difficult than others (Benton, 2015). There are also alternative vocational routes to achievement at the pre-16 level, such as Business and Technology Education Council qualifications (BTECs) and focussing only on GCSE attainment disregards this qualification. Those who are academically less able and/or from a disadvantaged background may be guided into less academic routes and qualifications (Henderson et al., 2018). It is plausible that those with lower language abilities and/ or of a lower socioeconomic status may take BTECs as an alternative to GCSE, and therefore estimates focussing solely on GCSE outcomes may be an underestimate of the effect of language on educational attainment (although note that some previous conceptualisations did take into account vocational equivalents; for example, Bukodi et al., 2017; Conti-Ramsden et al., 2009, 2018).

We focused on qualifications usually taken ~age 16 (i.e. GCSEs for England, Wales and Northern Ireland and N5 qualifications for cohort members in Scotland), and conceptualised educational attainment in the following two ways: 1) a binary variable of those who achieved ≥ grade 4 (or ≥ grade C for N5 qualifications) in the core subjects (0 = no; 1 = yes) and 2) a continuous variable of the average grade across core subjects. These are described in detail below.

Binary variable (grade 4 and above on core subjects (0 = no; 1 = yes).

"Core subjects" were identified as English, Mathematics and Science, as these are usually compulsory GCSE subjects in English schools – students take around 8-9 subjects at GCSE, and these must include English, Maths and Science. As can be seen from Table 5.1, due to this data being self-reported, there are multiple ways that these subjects have been reported. We decided that cohort members had taken the core subjects if they reported having at least one of each of the core category subjects listed in Table 5.1, i.e., cohort members needed: 1) at least one English subject, 2) at least one mathematics subject, and 3) at least one science subject). We created a binary variable of whether cohort members scored ≥ grade 4/C in these core subjects. As most post-16 education entry requirements ask for a minimum of this, we used this binary variable to see if vocabulary ability affects achievement of the minimum grade required in these core subjects. Subject choice also differs based on socioeconomic status, so entry to these subjects may be patterned by SEC (Henderson et al., 2018).

iGCSEs are international GCSEs, and these are exactly the same as GCSEs, except this qualification is recognised internationally and so if cohort members reported scoring above a grade 4 on core subject iGCSEs, they were included in the binary variable as meeting the criteria. This was calculated for GCSE/ iGCSEs and N5 qualifications separately and then combined into a single binary variable. Where cohort members reported having taken both GCSEs and N5 qualifications, they were given a score of 1 (i.e., scored 4/C and above in core subjects) if they scored this grade in either GCSEs or N5 qualifications. Those who reported having "no qualifications" were entered into this variable as 0 (not having passed the core subjects). There are no compulsory subjects for N5 qualifications, however for consistency we retained the "core subjects" classification. These subjects are also often mandatory in most Scottish schools, despite not being a national mandatory requirement. In Scotland, student may take some subjects at N5 and some subjects at higher grade at the same time. Therefore, we will consider students who have the core subjects at ≥grade C, in either N5 or Higher Grade qualifications. Since N4 qualifications are graded by teachers as pass or fail, and most pupils progress on to taking N5 qualifications, and the fact that most universities require at least a pass in English and Maths at N5 level, those who report having the core subjects at N4 level only will be classed as not achieving the core subjects.

Defining educational attainment in this way allowed those who took a mix of BTEC and GCSE qualifications, and importantly those with no qualifications, to be included in the outcome, rather than limiting the analysis to those who have only GCSEs or N5 qualifications. Some people may have taken the core subjects as a GCSE qualification, and other subjects as BTECs, for example. Some cohort members took English and Maths GCSEs, but science subjects as BTECs. A Level 2 BTEC is equivalent to a GCSE qualification grade A*-C. Cohort members who reported achieving above a grade 4 in English and Maths and passing a level 2 BTEC in a science subject were therefore classified as having achieved the core subjects at grade 9-4/ A*-C. By conceptualising a "core subjects" variable, we were able to include these individuals. The MCS2001 self-reported qualification data does not have the level of detail required to look specifically at BTEC achievement, but defining "core subjects" in a binary variable enabled us to include everyone with qualification data in the outcome variable, including those who have a mix of academic and vocational qualifications and even those who report no qualifications, who will be of particular interest here.

Different grading scales are used across GCSEs in England, Wales, Northern Ireland, and in N5 qualifications in Scotland. The old A* - G grading system and the new 9-1 grading

scale that is now used in England was harmonised onto a single scale, accounting for the fact that the new GCSEs have a wider range of grades: the following scale is used by the Department for Education for converting A*-G grades onto the 9-1 scale now used in England (see Ofqual, 2018 for an infographic demonstrating this scale):

A* = 8.5; A= 7; B = 5.5; C = 4; D=3; E = 2; F = 1.5; G = 1; Ungraded = 0

N5 qualifications in Scotland are graded A-D or No Award (those who achieve less than 40%). Rather than attempting to put this onto the same scale as the GCSEs, reaching a grade C or above was calculated separately for those with N5 qualifications and then combined into one variable with that of the GCSE benchmark binary variable.

Continuous variable: average grade across core subjects

The first outcome variable enabled us to investigate the role of age 5 vocabulary in achieving a pass in core subjects, an often-used government benchmark for GCSE achievement. However, we were also interested in whether age 5 vocabulary predicts variation in achievement in these core subjects, irrespective of whether the cohort member achieved a pass grade in these subjects. We therefore considered educational attainment in a second way. Anyone who reported taking at least one English subject, at least one mathematics subject and at least one science subject (see Table 5.1), and who reported a grade for these subjects was included in this variable. An English score was created (if the cohort member reported a grade for both English language and English literature, the mean of these two subjects represented the English score; if a cohort member reported only one grade, this was used as the English score). Similarly, mathematics and science scores were also calculated (again, if more than one subject from a core subject area was reported, the mean of these was calculated and this then represented that subject's score). The mean of the English, mathematics and science scores was taken as the outcome variable, to represent achievement in the core subjects.

We decided to operationalise educational attainment in this way, to account for the fact that each cohort member may have a different number of, or different combination, of subjects taken, and this way ensures that each contributes equal weighting to the outcome variable. This was calculated separately for GCSEs and N5s, given the different grading scales used. The harmonised GCSE grading scale defined above was used when calculating average GCSE score. Once the continuous score for GCSEs and N5s had been calculated,

these were converted to *z* scores within the relevant population, and then combined into a single variable for analysis.

Table 5.1. Core subjects reported in the MCS7 qualifications data

	GCSE	iGCSE	National Five	Highers	BTEC (level 2)
English subjects	English English language English literature	English (first language) English literature	Language: English	Language: English	NA
Maths subjects	Mathematics Mathematics (linear) Mathematics (numeracy) Further Mathematics Additional mathematics	Mathematics	Mathematics	Mathematics	NA
Science Subjects	Biology Chemistry Physics Applied Science Additional Applied Science Additional science Combined Science Further additional science Science Science (modular) Additional Science (modular) Human Biology Computer Science	Biology Chemistry Physics Science Computer Science	Biology Chemistry Physics Computer Science	Biology Human Biology Chemistry Physics Computer Science	Science Applied Science

Due to the self-report nature of the data, a wide variety of subjects are reported by cohort members, particularly for GCSE. We have no way of establishing exactly what cohort members mean when they report a specific subject, therefore we chose to include anyone who reported having at least one out of the English, Maths and Science categories in the above table. This

was a particular issue for Science, due to the number of different Science subjects reported. We cannot know for sure what a student was referring to, for example when they reported "Science" or "Additional applied science", therefore we chose to include all of these as contributing towards the "core subjects" variable. Computer Science is also included, as this counts towards the Science subjects (Department for Education, 2018d). In Wales, pupils can take *either* Mathematics or Mathematics (Numeracy) and so both of these are counted in the "core subjects" variable. English language is compulsory in all schools, and English literature is compulsory in most schools, therefore cohort members reporting at least one were classed as having "core subjects".

Modification variables: Socioeconomic circumstances

We considered 5 SEC indicators. These were first converted into a factor score to give a composite variable of parental SEC when cohort members were aged 3 (except wealth, which was measured when cohort members were aged 11), using confirmatory factor analysis (see Appendix 3, Section 1). This composite variable was used as the moderator variable in the main analysis, as in previous analyses, this composite variable was shown to explain the most variance in vocabulary throughout childhood and into adolescence (Thornton et al., 2021), when compared to each separate SEC indicator. The individual indicators that make up the composite factor score are detailed below.

Parental education: as a measure of parent's education, we used NVQ levels (both academic and vocational qualifications derived into NVQ levels 1-5, with level 1 equating to GCSE grades D-G or NVQ level 1 vocational equivalents, and level 5 equating to higher degree qualifications; see Rosenberg, 2012). We used the highest household NVQ level, as this was shown to predict the most variance in vocabulary (Thornton et al., 2021).

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

Occupational status: highest household occupational status (NS-SEC 4 categories: higher managerial; intermediate; routine; unemployed) at 3 years was used. If data was missing, occupational status at age 9 months was used instead.

Wealth: a measure of total net wealth, taken from the age 11 sweep of the MCS2001 was used. This measure was derived from 4 variables: amount outstanding on all mortgages, house value, number of investments and assets, and number of debts owed. Outstanding mortgages were subtracted from the house value, to give a measure of housing wealth. Debts owed was taken from the number 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.

Relative neighbourhood deprivation: Indices of multiple deprivation (IMD) deciles, ranging from the most deprived decile to the least deprived decile were taken from the age 3 sweep of the data. If data is missing, IMD deciles at age 9 months were used instead.

Potential confounding variables.

Demographic confounders. In all models, we controlled for cohort member's sex at birth (male, female); ethnicity (White, mixed, Indian, Pakistani and Bangladeshi, Black or Black British, other ethnic group (including Chinese); whether English was spoken as an additional language in the home (English only, English and another language, only another language); and the country that the cohort member lived in (England, Wales, Scotland, Northern Ireland).

Caregiver vocabulary. Vocabulary skill is thought to be partly heritable (Chow & Wong, 2021), as well as being shaped by the environment that infants are exposed to. The heritability of vocabulary may be influenced by SEC, in that those who are exposed to a higher SEC environment could be more likely to fulfil their genetic potential, which may be supressed in those from lower SEC families (Scarr & Mccartney, 1983). Since caregiver vocabulary likely reflects both the genetic component of vocabulary skill, and the language environment that cohort members are exposed to, we adjusted for this in our analyses. Caregiver vocabulary was measured in the age 14 sweep of the MCS2001, using the Word Activity Test (Closs, 1986): they were given a list of 20 target words, each presented alongside 5 other words, and had to choose the word which meant the same, or nearly the same as the target word, from the 5 options. In the MCS2001, there is a main respondent (usually the mother) and a partner respondent (usually the father), and both the main and partner respondents had the opportunity to complete the word activity test, with a different set of 20 words each. We used the mean score across respondents, under the assumption that the cohort member spends roughly an equivalent amount of time with both caregivers. Taking the mean enabled us to account for more of what cohort members are exposed to in terms of caregiver vocabulary and a proxy measure for the heritable component of vocabulary.

Data analysis

All analyses were pre-registered on the Open Science Framework website (OSF number osf.io/5bhx8), and all R code can be found on GitHub

(https://github.com/emmathornton/vocabulary-education). Analyses consisted of a series of multiple logistic and multiple linear regression models.

Missing data strategy. Multiple imputation using chained equations was employed to account for missing data, using the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011), to minimise the impact of biased attrition on the model estimates (Elliott & Shepherd, 2006). Imputations accounted for the interaction between SEC variables and age 5 vocabulary, since the interaction between these variables is of interest here (van Buuren & Groothuis-Oudshoorn, 2011). Each dataset was imputed 25 times, as this was greater than the overall percentage of missing data (14.8%)(White et al., 2011). No data were missing for sex at birth, relative neighbourhood deprivation or cohort member's country. 4% of age 5 vocabulary scores were missing. Figure S2 (Appendix 3, Section 2) displays the missing data for each variable in the dataset, which also includes auxiliary variables used in the multiple imputations. As can be seen from this figure, the percentage of missing data among the outcome variables is high (36.66% for the binary outcome variable and 50.79% for the continuous outcome variable). We therefore conducted sensitivity analyses on complete cases for each outcome variable, which did not change the overall pattern of results (see Appendix 3, Sections 3 and 4). The individual components of the wealth variable are also high (amount outstanding on mortgage 60.92% missing; house valuation 50.94% missing; total savings and investments 43.34% missing and outstanding debts owed 58.57% missing); however, sensitivity analyses in chapter 4, whereby all cohort members with a response to ≤ 1 wealth component variable and ≤2 wealth variables were considered, revealed similar patterns of results to the main analyses.

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 from the MCS2001 before the age of 5, when the vocabulary measure of interest was measured.

Analysis plan.

RQ1a & RQ2a. Does early childhood vocabulary predict whether cohort members achieve a functional level in core subject examinations at the end of secondary school? Does any such relation hold over and above SEC and caregiver vocabulary factors?

Initially, the unadjusted relationship between age 5 vocabulary and achieving ≥grade 4/C in the core subjects (binary variable) was estimated with logistic regression modelling, to

assess whether or not there was an association, before the addition of potential confounding variables. Subsequently, to test if any association remained after adding the potential confounding variables, the following multiple logistic regression models were estimated:

1. Sociodemographic confounding variables (sex, ethnicity, EAL, country, parent education, income, occupational status, wealth, and neighbourhood deprivation) were added to a model predicting whether cohort members achieved a pass in the core subjects.
2. Caregiver vocabulary was then added to this model
3. Finally, age five vocabulary was added to a model containing all potential confounding variables to assess whether vocabulary predicted the outcome after adjusting for potential confounders.

This allowed us to investigate whether age 5 vocabulary predicted reaching a functional level in (i.e., passing) core subjects at the end of secondary school.

RQ1b & RQ2b. Does early childhood vocabulary predict the level of achievement in the core subjects? Does any such relation hold over and above SEC and caregiver vocabulary factors?

Initially, the unadjusted relationship between age 5 vocabulary and the mean grade across core subjects was estimated with linear regression modelling, to assess whether or not there was an association before the addition of potential confounding variables. Subsequently, to test if any association remained after adding the potential confounding variables, the following multiple linear regression models were estimated:

1. Sociodemographic confounding variables (sex, ethnicity, EAL, country, parent education, income, occupational status, wealth, and neighbourhood deprivation) were added to a model predicting the mean grade across core subjects
2. Caregiver vocabulary was then added to this model
3. Finally, age five vocabulary was added to a model containing all potential confounding variables to assess whether vocabulary predicted the outcome after adjusting for potential confounders.

This allowed us to see if age 5 vocabulary predicted variation in achievement in the core subjects at the end of secondary school.

For RQ2a and RQ2b, a model containing demographic confounders (sex, ethnicity, EAL, parent education, income, occupational status, wealth, and neighbourhood deprivation) was compared to a model with no predictors. Each model was subsequently compared to the previous model to see if the different potential confounding variables predicted unique variance in educational attainment. An improvement in model fit was assessed using nested model comparisons for imputed data, using the method of Meng and Rubin (1992).

If an improvement in fit is seen when adding caregiver vocabulary to a model with demographic and SEC predictors, this would indicate that caregiver vocabulary predicts variance in educational achievement above and beyond SEC predictors. Similarly, if an improvement in model fit is seen after adding age 5 vocabulary, this would indicate that early vocabulary predicts unique variance in attainment, even after controlling for sociodemographic information, enabling us to answer the question of whether any relation holds above and beyond SEC and caregiver vocabulary factors.

RQ1 & RQ2 sensitivity analyses.

To establish whether our main findings were contingent on any of the analytic decisions outlined above, we ran a series of planned sensitivity analyses:

1. Welsh was also included as a core subject for cohort members in Wales; Welsh is a compulsory GCSE for those sitting the examinations in Wales; we therefore want to assess whether including this in our conceptualisation of core subjects changes the pattern of our results (see Appendix 3, Section 5)
2. We re-ran analyses on each country separately (England, Wales, Northern Ireland, and Scotland), due to the different education systems and examinations taken in each of these countries. This analysis allowed us to see if any one country is driving any particular finding (see Appendix 3, Section 6).

RQ1 & RQ2 additional analyses.

1. We investigated at the effects of age 5 vocabulary on attainment in English, Maths and Science separately at the end of secondary school, rather than as a combined measure, to see if language affects these subjects differently, or effects one subject more than others (see Appendix 3 Section 7).

RQ3. Is any relation between age 5 vocabulary and attainment moderated by SEC?

For this research question, the binary variable ascertaining whether cohort members achieved above a grade 4/grade C on the core subjects was the primary outcome variable. We made this analytic decision based on the fact that those who do not have GCSEs in the core subjects (who may be of a lower vocabulary ability and/or lower SEC) and those with no qualifications at all were included in this binary variable. The continuous measure does not include those without qualifications, and so we run the risk of conditioning on the outcome with this measure, and it may therefore be a biased measure in terms of establishing the role of vocabulary on educational attainment and whether this is moderated by any measures of SEC. To determine the moderating role of SEC on the relationship between vocabulary and educational attainment, the following multiple logistic regression model was built:

Binary educational attainment as outcome, sex, ethnicity, EAL, country, caregiver vocabulary, parent SEC composite* age 5 vocabulary interaction term

To test the significance of the interaction term, a model with the interaction term was compared to a model without the interaction term using nested model comparisons for imputed data (Meng & Rubin, 1992*)*, to see whether the interaction term predicts unique variance in the outcome.

RQ3 Planned additional analyses

We conducted a series of additional analyses to explore this question further, since each indicator of SEC could theoretically influence both vocabulary and education in different ways. To unpack any moderating effect of the composite variable and establish if it is any one SEC indicator in particular driving any interaction, the following models were estimated:

1. **Parent education as moderator.** Educational attainment as outcome, sex, ethnicity, EAL, country occupational status, income, wealth, neighbourhood deprivation and caregiver vocabulary as potential confounding variables, parent education* age 5 vocabulary interaction term
2. **Occupational status as moderator.** Educational attainment as outcome, sex, ethnicity, EAL, country, parent education, income, wealth, neighbourhood

deprivation and caregiver vocabulary as potential confounding variables, parent occupational status* age 5 vocabulary interaction term

3. **Income as moderator.** Educational attainment as outcome, sex, ethnicity, EAL, country, parent education, occupational status, wealth, neighbourhood deprivation and caregiver vocabulary as potential confounding variables, income* age 5 vocabulary interaction term
4. **Wealth as moderator.** Educational attainment as outcome, sex, ethnicity, EAL, country, parent education, occupational status, income, neighbourhood deprivation and caregiver vocabulary as potential confounding variables, wealth* age 5 vocabulary interaction term
5. **Neighbourhood deprivation as moderator.** Educational attainment as outcome, sex, ethnicity, EAL, country, parent education, occupational status, income, wealth and caregiver vocabulary as potential confounding variables, neighbourhood deprivation* age 5 vocabulary interaction term

Again, to test the significance of each interaction term, a model with the interaction term was compared to a model without the interaction term using nested model comparisons for imputed data (Meng & Rubin, 1992), to see whether the interaction term predicted unique variance in the outcome.

RQ3 sensitivity analysis

There is some concern in the literature that regression models with potential confounding variables simply being added as control variables do not properly adjust for the confounding effect of these variables on the interaction term – only on the potential confounding influence of the predictor variable (here, vocabulary) on the outcome (GCSE attainment) (Keller, 2014). To control for potential confounding effects on the *interaction term*, all potential confounders and interaction terms between the potential confounders and the predictor, and potential confounders and the moderator, must be entered into the model (Keller, 2014).

We therefore ran sensitivity analyses for our SEC* vocabulary moderations, whereby we included interaction terms between potential confounders (remaining SEC variables) and between the predictor (vocabulary) and potential confounders and the moderator (each SEC variable in turn), to ensure the confounding effect of SEC on the interaction term was accounted for (see Appendix 3, Section 9).

5.4. Results

Descriptive statistics.

Descriptive statistics were calculated across the 25 imputed datasets and can be found in Table 5.2. Analytical samples were compared to the whole cohort sample to see if there were any differences in characteristics of those included in the analyses. Proportions were similar between the whole cohort and our analytical sample. Means (±SD) for the average GCSE and average N5 grades were slightly higher in the whole cohort compared to our analytical sample. Average GCSE grades ranged from 0 to 8.9, with a higher grade indicating higher achievement and we had data on this variable for 7,136 out of the 14,474 cohort members in England, Wales, and Northern Ireland in the whole cohort sample. Average N5 grades ranged from 0-5, with 5 being the highest possible grade, and we had data for 539 of the 1,919 cohort members in Scotland in the whole cohort sample.

Table 5.2. Descriptive Statistics for Analytical Sample (N = 15,576) and Whole MCS2001 Cohort (N=19,243)

	Proportion (%) or Mean(±SD) [95% CIs]	
Variable	*Whole Cohort (N= 19243)*	*Analytical Sample (N=15,576)*
Vocabulary		
Cohort Member Vocabulary (Naming Vocabulary Score)	54.67(±10.97) [54.5;54.85]	54.54(±11.06) [54.37;54.72]
Caregiver Vocabulary (Word Activity Test Score)	11.46(±4) [11.39;11.53]	11.15(±4.04) [11.08;11.21]
Cohort Member Education		
Core Subjects Grade ≥ 4: No	34.28	38.21
Core Subjects Grade ≥ 4: Yes	65.72	61.79
Average GCSE grade (England, Wales & Northern Ireland) [1]	5.37(±1.59) [5.33;5.41]	4.84(±1.69) [4.81;4.86]
Average N5 grade (Scotland) [2]	4.33(±0.69) [4.27;4.38]	3.69(±0.94) [3.65;3.74]
Demographics		
Sex (Male)	51.31	50.99
Sex (Female)	48.69	49.01

Ethnicity (White)	85.9	86.2
Ethnicity (mixed)	3.37	3.3
Ethnicity (Indian)	1.87	1.9
Ethnicity (Pakistani & Bangladeshi)	4.42	4.4
Ethnicity (Black/ Black British)	3.06	2.91
Ethnicity (other incl. Chinese)	1.39	1.29
EAL (English only)	88.42	88.73
EAL (English and another language)	9.02	8.86
EAL (only another language)	2.56	2.4
Country (England)	82.11	82.21
Country (Wales)	5.04	4.95
Country (Scotland)	9.02	8.98
Country (Northern Ireland)	3.83	3.86
Socioeconomic Circumstances		
Parent Education (NVQ1)	5.82	5.43
Parent Education (NVQ2)	25.45	25
Parent Education (NVQ3)	15.99	15.98
Parent Education (NVQ4)	34.9	36.34
Parent Education (NVQ5)	7.23	7.58

Parent Education (None of these/overseas qualifications)	10.61	9.66
Income Quintile 1	21.62	21.03
Income Quintile 2	20.58	20.12
Income Quintile 3	19.7	19.86
Income Quintile 4	19.11	19.43
Income Quintile 5	18.98	19.55
Wealth Quintile 1[3]		17.95
Wealth Quintile 2		17.86
Wealth Quintile 3		19.36
Wealth Quintile 4		21.21
Wealth Quintile 5		23.62
Occupational Status (routine)	22.18	21.99
Occupational Status (intermediate)	19	19.29
Occupational Status (higher managerial)	38.45	39.88
Occupational Status (unemployed)	20.37	18.84
Relative Neighbourhood Deprivation (most deprived decile)	12.85	12.22
Relative Neighbourhood Deprivation (10 - <20%)	10.99	10.6
Relative Neighbourhood Deprivation (20 - <30%)	10.33	10.18
Relative Neighbourhood Deprivation (30 - <40%)	9.18	9.14
Relative Neighbourhood Deprivation (40 - <50%)	9.76	9.9
Relative Neighbourhood Deprivation (50 - <60%)	9.71	9.79
Relative Neighbourhood Deprivation (60 - <70%)	8.84	8.95
Relative Neighbourhood Deprivation (70 - <80%)	9.02	9.3
Relative Neighbourhood Deprivation (80 - <90%)	9.4	9.73

Relative Neighbourhood Deprivation (least deprived decile)	9.9	10.2

[1]Average GCSE grade refers to the average grade across English, Mathematics and Science subjects and was calculated post imputation on data from England, Wales and Northern Ireland.

[2]Average N5 (National Five) grade refers to the average grade across English, Mathematics and Science subjects and was calculated post imputation on data from Scotland.

[3]Wealth was derived post imputation and therefore we do not present the proportions for Wealth quintiles in the whole cohort, due to the high level of missing data among the wealth components.

In analyses, the average GCSE grade and average N5 grade were converted into z scores and combined into one variable pre-imputation.

Note that descriptives for the whole cohort (N=19,243) do not indicate the level of missing data for these variables in the whole cohort. Missing data proportions can be found in Figure S2.

Analytical sample descriptive statistics were calculated and pooled across 25 imputed datasets and are weighted to account for the sample design of the MCS2001 and attrition.

RQ1a & RQ2a. Does early childhood vocabulary predict whether cohort members achieve a functional level in core subject examinations at the end of secondary school? Does any such relation hold over and above SEC and caregiver vocabulary factors?

In an unadjusted model (i.e., not including any potential confounding variables), there was a significant relation between age 5 vocabulary and achieving ≥ grade 4 on the core subjects at the end of secondary school, such that with every SD unit increase in age 5 vocabulary, the odds of passing the benchmark of ≥ grade 4 on the core subjects increased by 85% (OR = 1.85, 95% CIs = [1.76; 1.95]). To test whether this relation held when potential confounding factors were included, we first tested whether two sets of potential confounding variables predicted whether the benchmark of ≥ grade 4 on the core subjects was reached (see Table 5.3). We subsequently assessed whether vocabulary explained variance over and above these variables. Compared to a model with no predictors, a model with sociodemographic confounding variables improved the model fit (Dm(36, 3351.11)=35.87, p<.001). Further, compared to a model containing only sociodemographic predictors, a model that also included caregiver vocabulary improved the model fit (Dm(1, 79.09) = 116.58, p<.001), indicating that caregiver vocabulary predicted variance in achieving the benchmark above sociodemographic variables. Finally, adding age 5 vocabulary scores to a model containing sociodemographic and caregiver vocabulary factors improved the model fit (Dm(1, 66.2) = 205.53, p<.001), such that higher vocabulary scores were associated with increased odds of passing the benchmark grade threshold: after controlling for sociodemographic and caregiver vocabulary factors, with every SD unit increase in age 5 vocabulary, the odds of passing the

benchmark of ≥ grade 4 on the core subjects increased by 61% (OR = 1.61, 95% CIs = [1.50;1.72]; see Table 5.3).

RQ1b & RQ2b. Does early childhood vocabulary predict the level of achievement in the core subjects? Does any such relation hold over and above SEC and caregiver vocabulary factors?

In an unadjusted model (i.e., not including any potential confounding variables), a positive relation was observed, such that higher vocabulary scores were associated with higher levels of overall achievement (β = .38, 95% CIs = [.36;.40]). To test whether this relation held when potential confounding factors were included, we first tested whether two sets of potential confounding variables predicted level of achievement on the core subjects (see Table 5.3). We subsequently assessed whether vocabulary explained variance over and above these variables. Compared to a model with no predictors, a model with sociodemographic confounding variables improved the model fit (Dm(36, 3203.3)=88.95, p<.001). Further, compared to a model containing only sociodemographic predictors, a model that also included caregiver vocabulary improved the model fit (Dm(1, 70.91) = 226.71, p<.001), indicating that caregiver vocabulary predicted variance in achieving the benchmark above sociodemographic variables. Finally, adding age 5 vocabulary scores to a model containing sociodemographic and caregiver vocabulary factors improved the model fit (Dm(1, 65.14) = 331.76, p<.001), with higher vocabulary scores predicting higher levels of overall achievement, above and beyond sociodemographic and caregiver vocabulary factors (β = .24, 95% CIs = [.21;.27]; see Table 5.3).

Complete case analyses for both outcome variables, including Welsh as a core subject for those who were living in Wales, and analyses conducted on each country separately did not change the overall pattern of results for the first two research questions: vocabulary predicted unique variance in educational attainment in all analyses (see Appendix 3, Sections 3-6).

Exploratory analyses on attainment in English, Mathematics and Science as separate subjects revealed that vocabulary similarly predicted achievement in each of these subjects, and predicted unique variance above and beyond SEC and caregiver vocabulary factors (See Appendix 3, Section 7).

Table 5.3 Predicting Educational Attainment (≥grade 4 on core subjects and average grade) (N=15,576)

	Binary Outcome (OR [95% CIs]			Continuous Outcome (β [95% CIs]		
Variable	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Sociodemographic confounders						
Sex (male)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Sex (female)	1.33[1.21;1.45] * * * p<.001	1.33[1.21;1.46] * * * p<.001	1.30[1.18;1.43] * * * p<.001	.20[.16;.25] * * * p<.001	.20[.15;.24] * * * p<.001	.18[.14;.23] * * * p<.001
Ethnicity (White)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Ethnicity (mixed)	1.17[.91;1.51] p= .223	1.18[.91;1.53] p= .215	1.18[.91;1.54] p= .218	.09[-.01;.20] p= .087	.10[-.01;.21] p= .067	.10[-.01;.20] p= .070
Ethnicity (Indian)	2.17[1.37;3.43] * * p= .001	2.53[1.59;4.01] * * * p<.001	2.61[1.63;4.18] * * * p<.001	.23[.07;.39] * * p= .005	.33[.17;.48] * * * p<.001	.32[.17;.47] * * * p<.001
Ethnicity (Pakistani & Bangladeshi)	1.41[1.02;1.97] * p= .041	1.69[1.20;2.37] * * p= .003	2.04[1.44;2.88] * * * p<.001	.15[.01;.28] * p= .030	.26[.13;.40] * * * p<.001	.35[.22;.49] * * * p<.001
Ethnicity (Black/ Black British)	1.48[1.11;1.99] * * p= .009	1.72[1.28;2.30] * * * p<.001	1.97[1.47;2.63] * * * p<.001	.11[-.02;.24] p= .092	.20[.07;.33] * * p= .002	.27[.14;.39] * * * p<.001
Ethnicity (other incl. Chinese)	2.09[1.30;3.34] * * p= .002	2.45[1.51;3.97] * * * p<.001	2.97[1.80;4.90] * * * p<.001	.49[.31;.67] * * * p<.001	.59[.41;.77] * * * p<.001	.66[.48;.84] * * * p<.001
EAL (English only)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE

EAL (English and another language)	1.45[1.12;1.89] * * p= .005	1.69[1.29;2.23] * * * p<.001	2.06[1.55;2.72] * * * p<.001	.18[.08;.28] * * * p<.001	.26[.16;.35] * * * p<.001	.34[.24;.43] * * * p<.001
EAL (only another language)	1.55[1.08;2.22] * p= .017	2.00[1.40;2.87] * * * p<.001	2.77[1.92;3.99] * * * p<.001	.21[.05;.37] * p= .010	.35[.19;.50] * * * p<.001	.49[.34;.64] * * * p<.001
Country (England)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Country (Wales)	.95[.79;1.15] p= .595	.97[.80;1.18] p= .769	1.00[.82;1.21] p= .985	-.08[-.16;.01] p= .092	-.06[-.15;.03] p= .183	-.04[-.13;.04] p= .331
Country (Scotland)	.46[.37;.56] * * * p<.001	.45[.36;.55] * * * p<.001	.43[.35;.53] * * * p<.001	-.73[-.84;-.62] * * * p<.001	-.74[-.85;-.63] * * * p<.001	-.74[-.85;-.64] * * * p<.001
Country (Northern Ireland)	1.38[1.11;1.72] * * p= .004	1.43[1.15;1.79] * * p= .002	1.39[1.11;1.74] * * p= .005	.22[.12;.31] * * * p<.001	.23[.14;.33] * * * p<.001	.22[.12;.31] * * * p<.001
Parent Education (NVQ1)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Parent Education (None of these/overseas qualifications)	1.16[.87;1.55] p= .308	1.21[.90;1.62] p= .196	1.23[.91;1.66] p= .167	.07[-.04;.19] p= .217	.10[-.02;.21] p= .097	.11[-.01;.22] p= .063
Parent Education (NVQ2)	1.49[1.20;1.85] * * * p<.001	1.40[1.12;1.73] * * p= .003	1.36[1.09;1.69] * * p= .007	.19[.09;.29] * * * p<.001	.15[.05;.25] * * p= .005	.13[.03;.23] * p= .013
Parent Education (NVQ3)	1.82[1.44;2.31] * * * p<.001	1.63[1.28;2.07] * * * p<.001	1.56[1.23;2.00] * * * p<.001	.28[.18;.39] * * * p<.001	.21[.10;.31] * * * p<.001	.18[.07;.29] * * p= .001
Parent Education (NVQ4)	2.53[1.98;3.23] * * * p<.001	2.05[1.60;2.61] * * * p<.001	1.89[1.48;2.42] * * * p<.001	.54[.43;.64] * * * p<.001	.40[.29;.51] * * * p<.001	.35[.24;.45] * * * p<.001

Parent Education (NVQ5)	3.99[2.86;5.57] * * * p<.001	2.76[1.97;3.85] * * * p<.001	2.50[1.78;3.51] * * * p<.001	.89[.77;1.02] * * * p<.001	.66[.52;.79] * * * p<.001	.60[.46;.73] * * * p<.001
Income Quintile 1	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Income Quintile 2	1.13[.95;1.34] p= .154	1.10[.93;1.31] p= .275	1.11[.93;1.32] p= .232	.03[-.05;.11] p= .473	.01[-.07;.09] p= .778	.01[-.06;.09] p= .702
Income Quintile 3	1.39[1.14;1.68] * * p= .001	1.34[1.10;1.62] * * p= .004	1.28[1.05;1.57] * p= .017	.14[.06;.23] * * p= .001	.12[.04;.20] * * p= .006	.10[.01;.18] * p= .025
Income Quintile 4	1.65[1.36;2.01] * * * p<.001	1.56[1.29;1.90] * * * p<.001	1.51[1.23;1.84] * * * p<.001	.23[.14;.32] * * * p<.001	.19[.11;.28] * * * p<.001	.17[.08;.25] * * * p<.001
Income Quintile 5	1.91[1.47;2.48] * * * p<.001	1.73[1.34;2.24] * * * p<.001	1.64[1.26;2.15] * * * p<.001	.38[.29;.48] * * * p<.001	.32[.22;.41] * * * p<.001	.29[.19;.38] * * * p<.001
Occupational Status (routine)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Occupational Status (unemployed)	.85[.72;1.01] p= .057	.84[.70;1.00] * p= .045	.88[.74;1.05] p= .157	-.03[-.11;.04] p= .366	-.04[-.11;.04] p= .305	-.01[-.08;.06] p= .802
Occupational Status (intermediate)	1.30[1.13;1.49] * * * p<.001	1.22[1.06;1.40] * * p= .005	1.20[1.04;1.39] * p= .012	.12[.05;.20] * * p= .001	.09[.02;.16] * p= .017	.08[.01;.14] * p= .033
Occupational Status (higher managerial)	1.87[1.58;2.21] * * * p<.001	1.64[1.38;1.93] * * * p<.001	1.60[1.34;1.90] * * * p<.001	.35[.26;.43] * * * p<.001	.26[.18;.34] * * * p<.001	.24[.16;.32] * * * p<.001
Wealth Quintile 1	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE

Wealth Quintile 2	1.09[.90;1.32] p= .352	1.10[.91;1.33] p= .313	1.10[.92;1.33] p= .286	.03[-.05;.11] p= .507	.03[-.05;.11] p= .441	.03[-.04;.11] p= .402
Wealth Quintile 3	1.11[.90;1.37] p= .308	1.11[.90;1.37] p= .316	1.11[.90;1.36] p= .327	.04[-.05;.14] p= .390	.04[-.06;.13] p= .415	.04[-.05;.13] p= .409
Wealth Quintile 4	1.18[.96;1.46] p= .109	1.17[.95;1.45] p= .132	1.17[.95;1.44] p= .136	.09[.00;.18] * p= .040	.08[-.00;.17] p= .056	.08[-.00;.16] p= .053
Wealth Quintile 5	1.37[1.08;1.75] * p= .012	1.31[1.02;1.68] * p= .033	1.31[1.02;1.68] * p= .037	.25[.16;.34] * * * p<.001	.22[.13;.31] * * * p<.001	.21[.13;.30] * * * p<.001
Relative Neighbourhood Deprivation (most deprived decile)	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
Relative Neighbourhood Deprivation (10 - <20%)	1.10[.92;1.31] p= .295	1.09[.92;1.30] p= .326	1.08[.90;1.30] p= .391	.06[-.02;.14] p= .168	.05[-.03;.13] p= .208	.05[-.03;.13] p= .240
Relative Neighbourhood Deprivation (20 - <30%)	1.11[.91;1.34] p= .300	1.09[.89;1.32] p= .395	1.06[.87;1.30] p= .555	.06[-.03;.15] p= .189	.05[-.04;.13] p= .291	.03[-.05;.12] p= .442
Relative Neighbourhood Deprivation (30 - <40%)	1.18[.96;1.44] p= .110	1.14[.93;1.40] p= .196	1.12[.91;1.38] p= .285	.06[-.03;.15] p= .188	.04[-.05;.13] p= .397	.03[-.06;.11] p= .537
Relative Neighbourhood Deprivation (40 - <50%)	1.22[1.03;1.46] * p= .024	1.17[.98;1.40] p= .083	1.15[.96;1.38] p= .135	.08[-.00;.17] p= .052	.06[-.03;.14] p= .186	.05[-.04;.13] p= .290

Relative Neighbourhood Deprivation (50 - <60%)	1.28[1.06;1.56] * p= .013	1.23[1.01;1.50] * p= .039	1.19[.97;1.46] p= .089	.06[-.02;.14] p= .124	.04[-.05;.12] p= .385	.02[-.06;.10] p= .640
Relative Neighbourhood Deprivation (60 - <70%)	1.32[1.08;1.61] * * p= .007	1.23[1.01;1.51] * p= .043	1.22[.99;1.50] p= .064	.12[.02;.21] * p= .013	.07[-.01;.16] p= .102	.06[-.03;.15] p= .172
Relative Neighbourhood Deprivation (70 - <80%)	1.43[1.17;1.76] * * * p<.001	1.33[1.08;1.64] * * p= .007	1.29[1.04;1.60] * p= .023	.16[.07;.25] * * * p<.001	.11[.02;.20] * p= .018	.09[-.00;.18] p= .058
Relative Neighbourhood Deprivation (80 - <90%)	1.79[1.43;2.23] * * * p<.001	1.68[1.34;2.11] * * * p<.001	1.66[1.32;2.09] * * * p<.001	.23[.14;.32] * * * p<.001	.19[.10;.28] * * * p<.001	.18[.09;.27] * * * p<.001
Relative Neighbourhood Deprivation (least deprived decile)	1.77[1.41;2.23] * * * p<.001	1.64[1.29;2.07] * * * p<.001	1.56[1.22;1.98] * * * p<.001	.23[.14;.32] * * * p<.001	.18[.08;.27] * * * p<.001	.15[.06;.24] * * p= .002
Caregiver Vocabulary						
Caregiver Vocabulary (Word Activity Test Score)		1.46[1.36;1.57] * * * p<.001	1.34[1.24;1.44] * * * p<.001		.23[.20;.26] * * * p<.001	.18[.15;.21] * * * p<.001
Cohort Member Vocabulary						
Cohort Member Vocabulary (Naming Vocabulary Score)			1.61[1.50;1.72] * * * p<.001			.24[.21;.27] * * * p<.001
R2 (%)				27.78[26.33;29.23]	29.93[28.39;31.47]	33.1[31.53;34.67]

RQ3. Is any relation between age 5 vocabulary and attainment moderated by SEC?

When controlling for sex, ethnicity, EAL status, country and caregiver vocabulary skill, a positive relation between a composite measure of SEC and the likelihood of achieving ≥grade 4 on the core subjects at the end of secondary school was observed, such that with every SD unit increase in SEC, the odds of passing this benchmark increases (OR = 1.97, 95% CIs = 1.85;2.1]; see Table 5.4). Similarly, a positive relation between vocabulary skill and achieving ≥grade 4 on the core subjects was observed, such that for every SD unit increase in vocabulary, the odds of passing this benchmark increases (OR = 1.63, 95% CIs = [1.53; 1.74]; see Table 5.4).

Further, the relationship between age 5 vocabulary and achieving ≥ grade 4 on the core subjects at the end of secondary school is moderated by one's SEC, with an additional increase in the odds of passing the benchmark threshold with each SD unit increase of vocabulary, for each additional SD unit increase in SEC (OR = 1.09, 95% CIs = [1.03; 1.14]; see Figure 1 and Table 4). To determine the significance of the interaction term, a model with the SEC composite*age 5 vocabulary interaction term was compared to a model without this interaction term. Compared to a model controlling for sex, ethnicity, EAL status, country, and caregiver vocabulary, a model which also included an SEC composite*age 5 vocabulary score interaction term increased the model fit (Dm(1, 242.21) = 9.41, p=.002), indicating that the relationship between age 5 vocabulary and the likelihood of achieving ≥ grade 4 on the core subjects is moderated by early childhood SEC.

As can be seen from Figure 5.1, which displays the composite measure of SEC broken down into quintiles, as vocabulary skill increases, the probability of successfully achieving ≥grade 4 on the core subjects increases. The strength of this relation incrementally increases with each of the first two steps up in SEC, being strongest in the middle SEC group. As can be seen from Figure 5.1, there are also stark absolute differences between the highest and lowest quintiles of SEC: for those with a vocabulary skill at the top end of the distribution at the age of 5 in the highest SEC quintile, the probability of passing the threshold is 92%, compared to 52% for those with the same level of vocabulary skill in the lowest SEC quintile.

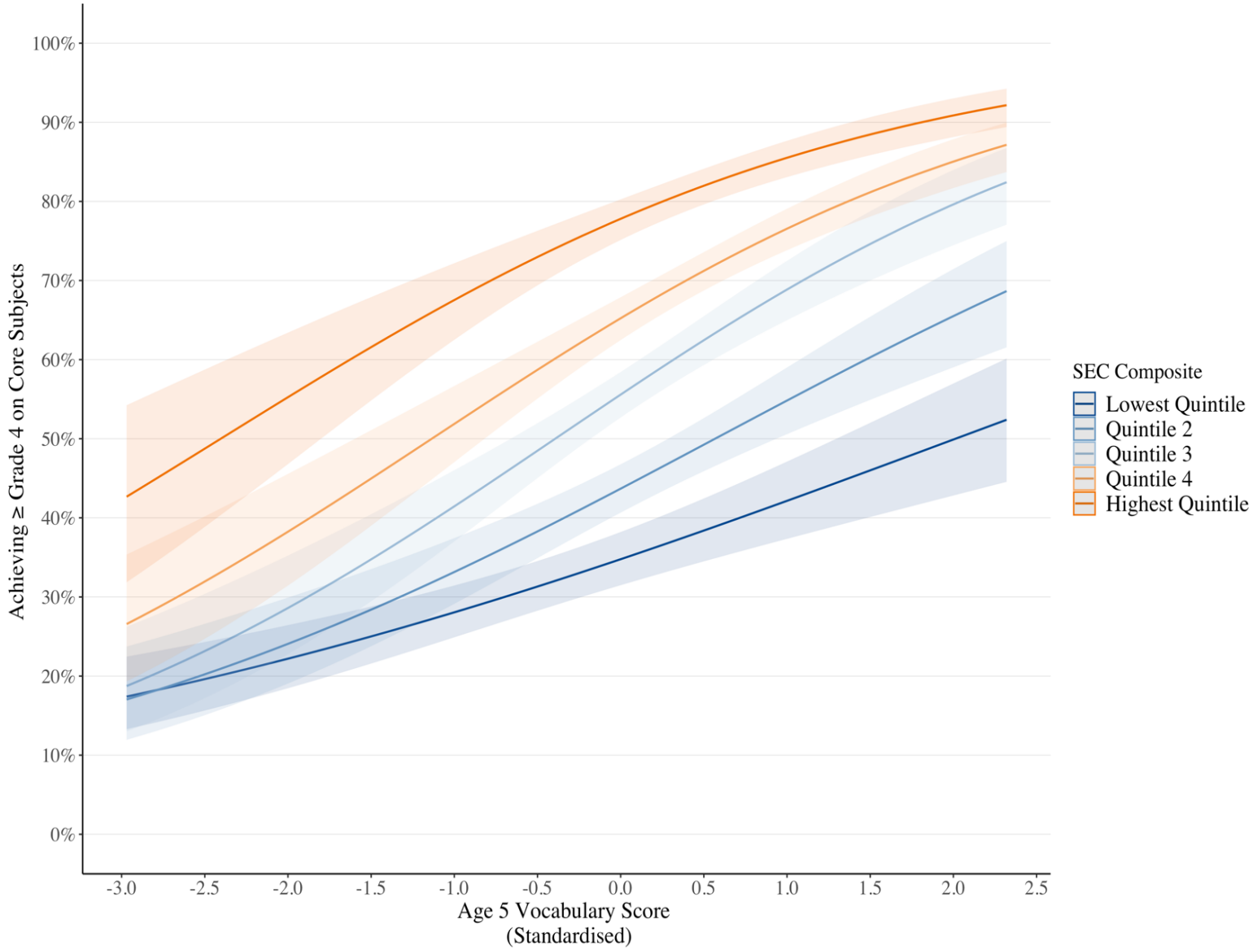
For those with vocabulary at the very bottom of the distribution at the age of 5, and therefore potentially with clinical levels of vocabulary difficulties, there are reduced chances achieving ≥grade 4 on the core subjects at the end of secondary school, although there are

slight increases in this probability with each step up in SEC: in the lowest three quintiles of SEC, those with poor language skills at the age of 5 are only 17-19% likely to pass the threshold. Even those in the fourth quintile of SEC who have very poor language skills at age 5 are only 27% likely to pass the threshold. Being in the most affluent SEC group (quintile 5) affords some advantage compared to the lower SEC groups for those with poor language skills, with these children being 43% likely to reach the threshold of achievement, which is already a much higher probability of achieving these grades than those in lower SEC quintiles of the same ability. This probability is similar to that of those in the lowest SEC group, with vocabulary at the top end of the distribution (52%). Thus, for those with very poor vocabulary skill at the age of 5, SEC offers little protection against the impact of such vocabulary skill on later educational achievement, particularly for those who are in the lowest SEC groups. Of these children who are in the highest SEC quintile, SEC appears to boost educational achievement relative to those in lower SEC groups of similar vocabulary skill. However, SEC clearly amplifies the effects of age 5 vocabulary on later educational achievement at all levels of vocabulary skill, with the probability of achieving ≥grade 4 on the core subjects increasing with every step up in SEC.

Interestingly, a sensitivity analysis whereby Welsh was included as a core subject for those who were living in Wales found that SEC did not moderate the relation between age 5 vocabulary and educational attainment (see Appendix 3, Section 5). This may reflect SEC and other related differences (such as confidence and attitudes towards the language) in the uptake of studying Welsh across different areas of Wales, which is something to be further explored in a Welsh context (National Survey for Wales 2017-18, 2018). Sensitivity analyses on England, Wales, Scotland, and Northern Ireland separately revealed a significant moderation effect of SEC in England only. There was no moderation of SEC on the relation between age 5 vocabulary and educational attainment in Wales, Scotland, or Northern Ireland (see Appendix 3, Section 6). As can be seen from Figure S3, effect sizes are similar in the four countries, therefore this pattern of findings may be a result of analyses in the smaller UK countries being underpowered. In summary, very poor levels of vocabulary at the age of 5 impede the likelihood of achieving ≥grade 4 on the core subjects at the end of secondary school at all levels of SEC, although this effect is attenuated in the highest SEC quintile, with these children being somewhat protected against the negative impact of poor vocabulary on attainment. As vocabulary skill increases, so does the probability of passing this benchmark of attainment in all SEC groups, and the probability of this achievement is particularly bolstered in the middle SEC groups. However, inequalities in the effects of vocabulary on

educational attainment are clear at all levels of vocabulary skill, with those in higher SEC quintiles being more likely to pass the benchmark threshold relative to those of a lower SEC with the same level of vocabulary skill (see Figure 5.1 and Table S9).

Figure 5.1. Predicted probabilities of achieving the benchmark threshold for vocabulary in each SEC quintile



Predicted probabilities when sex, ethnicity, EAL status and country set to reference levels and caregiver vocabulary score set to the mean. See table S9 for predicted values.

RQ3 Additional analyses

To unpack the moderating effect of the SEC composite further, we investigated the moderating effect of each individual SEC indicator considered separately (see Figure 5.2 and Table 5.4). As can be seen from Figure 5.2, when SEC is broken down by indicator, parent education and occupational status particularly clearly reflect the pattern observed when

considering SEC as a composite variable (Figure 5.1), with the probability of achieving ≥grade 4 on the core subjects being higher for the highest group at the bottom of the vocabulary distribution, relative to the lower SEC groups. The steepest increase in outcomes is again observed for the middle groups of the educational attainment and occupational status variables.

Parent educational Attainment

The relationship between age 5 vocabulary and achieving a grade 4 or above on the core subjects at the end of secondary school was moderated by level of parent education (see Table 5.4). Compared to a model controlling for sex, ethnicity, EAL status, country, caregiver vocabulary, income, occupational status, wealth and relative neighbourhood deprivation, and parent education and age 5 vocabulary as main effects, a model which also included a parent education*age 5 vocabulary score interaction term increased the model fit (Dm(5, 701.87) = 2.91, p=.013), indicating that the relationship between age 5 vocabulary and the likelihood of achieving ≥ grade 4 on the core subjects is moderated by level of parent education. As seen with the composite variable, parent education boosts the effects of vocabulary for attainment, particularly in the middle education groups, and ameliorates the impact of a negative vocabulary for those whose parents have the highest qualification levels (NVQ 5), with a higher probability for achievement observed for those in this group compared to lower levels of parental qualifications, even at the bottom of the vocabulary distribution.

Household Income

The relationship between age 5 vocabulary and achieving a grade 4 or above on the core subjects at the end of secondary school is moderated by household income (see Table 5.4). Compared to a model controlling for sex, ethnicity, EAL status, country, caregiver vocabulary, parent education, occupational status, wealth and relative neighbourhood deprivation, and household income and age 5 vocabulary as main effects, a model which also included an income*age 5 vocabulary score interaction term increased the model fit (Dm(4, 787.5) = 3.11, p=.015), indicating that the relationship between age 5 vocabulary and the likelihood of achieving ≥ grade 4 on the core subjects is moderated by household income. Here the pattern is clear, with children at the bottom of the vocabulary distribution in all income groups experiencing low probabilities of achieving ≥grade 4 on the core subjects. As

vocabulary skill increases, the probability of passing the benchmark also increases, and this increase is steeper for those in higher income groups.

Occupational Status

The relationship between age 5 vocabulary and achieving a grade 4 or above on the core subjects at the end of secondary school is moderated by occupational status (see Table 5.4). Compared to a model controlling for sex, ethnicity, EAL status, country, caregiver vocabulary, parent education, household income, wealth and relative neighbourhood deprivation, and occupational status and age 5 vocabulary as main effects, a model which also included an occupational status*age 5 vocabulary score interaction term increased the model fit (Dm(3, 359.66) = 4.57, p=.004), indicating that the relationship between age 5 vocabulary and the likelihood of achieving ≥ grade 4 on the core subjects is moderated by occupational status. Like with parent education, occupational status boosts the effects of vocabulary for attainment in the middle occupation groups and ameliorates the impact of a negative vocabulary for those whose parents have professional occupations, with a higher probability for achievement observed for those in this group compared to lower occupations, even at the bottom of the vocabulary distribution.

Wealth

The relationship between age 5 vocabulary and achieving a grade 4 or above on the core subjects at the end of secondary school is not moderated by household wealth (see Table 5.4). Compared to a model controlling for sex, ethnicity, EAL status, country, caregiver vocabulary, parent education, income, occupational status, and relative neighbourhood deprivation, and household wealth and age 5 vocabulary as main effects, a model which also included a household wealth*age 5 vocabulary score interaction term did not increase the model fit (Dm(4, 301.97) = 0.75, p=.557), indicating that the relationship between age 5 vocabulary and the likelihood of achieving ≥ grade 4 on the core subjects is not moderated by household wealth.

Relative Neighbourhood Deprivation

The relationship between age 5 vocabulary and achieving a grade 4 or above on the core subjects at the end of secondary school is not moderated by relative neighbourhood deprivation (see Table 5.4). Compared to a model controlling for sex, ethnicity, EAL status, country, caregiver vocabulary, parent education, income, occupational status, and wealth, and

relative neighbourhood deprivation and age 5 vocabulary as main effects, a model which also included a relative neighbourhood deprivation*age 5 vocabulary score interaction term did not increase the model fit (Dm(9, 15528) = 0.87, p=.528), indicating that the relationship between age 5 vocabulary and the likelihood of achieving ≥ grade 4 on the core subjects is not moderated by relative neighbourhood deprivation.

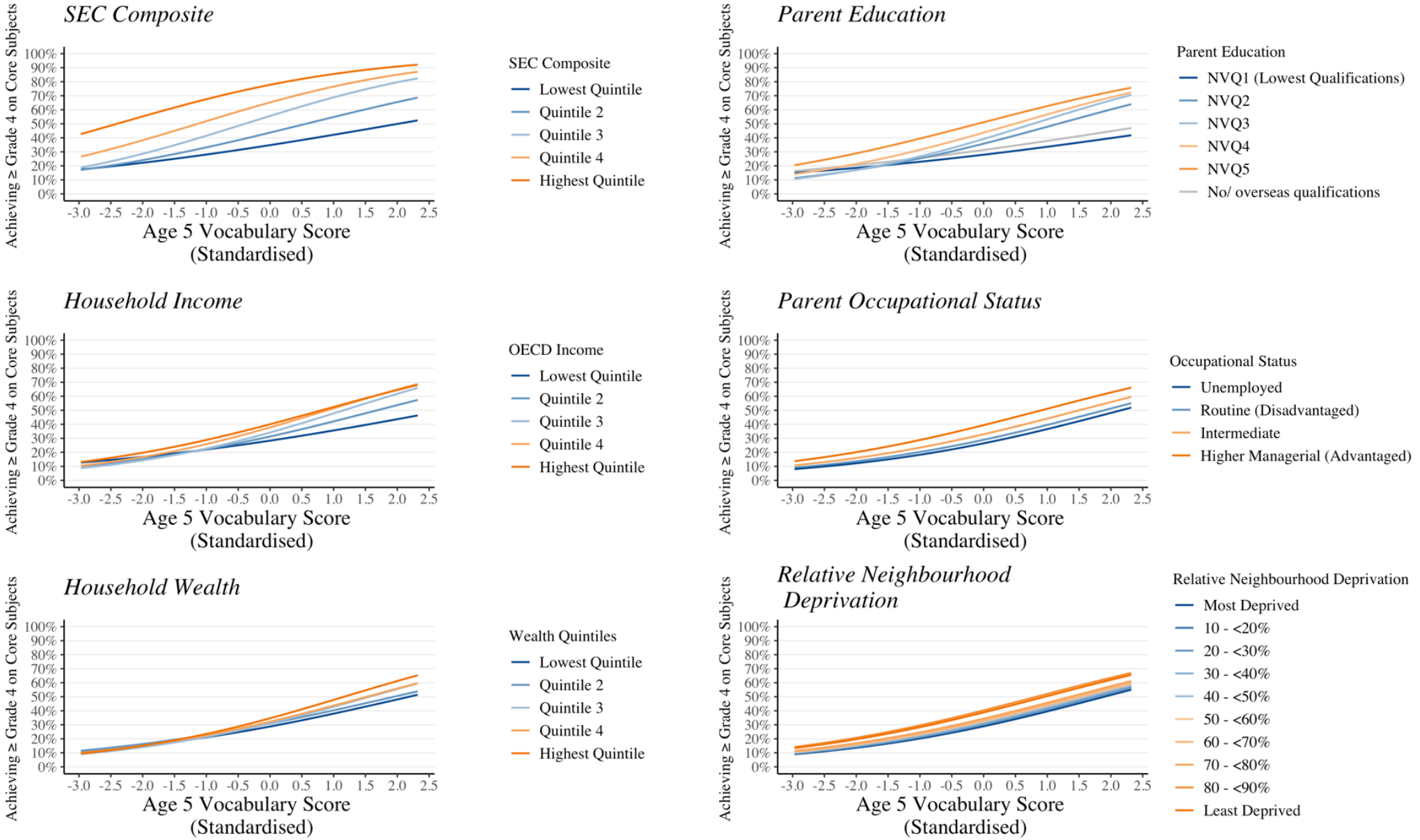
Overall, parent education, household income and occupational status moderated the relation between age 5 vocabulary and the likelihood of achieving ≥grade 4 on the core subjects. Wealth and relative neighbourhood deprivation did not moderate the relation between age 5 vocabulary and the likelihood of achieving ≥grade 4 on the core subjects.

Sensitivity Analyses

To ensure our moderation analyses with separate indicators of SEC as the moderator properly adjusted for confounding between the confounding variable, the predictor and the moderator, we conducted a series of sensitivity analyses, whereby we included interaction terms between potential confounders (remaining SEC variables) and between the predictor (vocabulary) and potential confounders and the moderator (each SEC variable in turn). For each analysis, a model with the interaction term was compared to a model without the interaction term, to establish whether there were any moderation effects when adjusting for confounding in this conservative way (see Appendix 3, Section 9).

This extremely conservative set of sensitivity analyses revealed that, over 25 imputed datasets, there were no moderation effects for individual SEC indicators. This suggests that the individual indicators of SEC are not separable in their interaction effects on the relation between vocabulary and educational attainment. The moderation of SEC is likely an additive effect of each SEC indicator, which likely have shared variance in their interaction terms.

Figure 5.2. Predicted probabilities of achieving the benchmark threshold for vocabulary, moderated by each SEC indicator



**Predicted probabilities when categorical potential confounders (sex, ethnicity, EAL status, country and remaining SEC indicators for individual indicator moderations set to reference levels and caregiver vocabulary score set to the mean.*
See table S9 for predicted probabilities values.

Table 5.4. Moderation analyses results for each SEC indicator moderator (vocabulary and indicator main effects and interaction terms)

	Variable	Composite SEC moderator (OR[95% CIs]	Parent education moderator (OR[95% CIs]	Income moderator (OR[95% CIs]	Occupational status moderator (OR[95% CIs]	Wealth moderator (OR[95% CIs]	Relative neighbourhood deprivation moderator (OR[95% CIs]
SEC Composite Moderator	Cohort Member Vocabulary (Naming Vocabulary Score)	1.63[1.53;1.74] * * * p<.001	1.31[1.07;1.60] * * p= .010	1.40[1.26;1.56] * * * p<.001	1.59[1.41;1.78] * * * p<.001	1.51[1.35;1.70] * * * p<.001	1.49[1.30;1.71] * * * p<.001
	Socioeconomic Circumstances (composite)	1.97[1.85;2.10] * * * p<.001					
	SEC*Vocabulary	1.09[1.03;1.14] * * p= .002					
Parent Education Moderator	Parent Education (NVQ1)		REFERENCE				
	Parent Education (None of these/overseas qualifications)		1.18[.87;1.60] p= .291				
	Parent Education (NVQ2)		1.44[1.15;1.80] * * p= .002				
	Parent Education (NVQ3)		1.66[1.30;2.12] * * * p<.001				
	Parent Education (NVQ4)		2.00[1.56;2.55] * * * p<.001				
	Parent Education (NVQ5)		2.70[1.90;3.82] * * * p<.001				

	Parent Education (NVQ1)*Age 5 Vocabulary	REFERENCE
	Parent Education (None of these/overseas qualifications)*Age 5 Vocabulary	1.02[.80;1.29] p= .872
	Parent Education (NVQ2)*Age 5 Vocabulary	1.26[1.01;1.58] * p= .041
	Parent Education (NVQ3)*Age 5 Vocabulary	1.35[1.05;1.74] * p= .018
	Parent Education (NVQ4)*Age 5 Vocabulary	1.29[1.04;1.61] * p= .021
	Parent Education (NVQ5)*Age 5 Vocabulary	1.23[.89;1.70] p= .211
Income Moderator	Income Quintile 1	REFERENCE
	Income Quintile 2	1.15[.97;1.37] p= .111
	Income Quintile 3	1.31[1.07;1.61] * * p= .009
	Income Quintile 4	1.54[1.26;1.88] * * * p<.001
	Income Quintile 5	1.69[1.29;2.21] * * * p<.001
	Income Quintile 1*Age 5 Vocabulary	REFERENCE

Income Quintile 2*Age 5 Vocabulary	1.14[1.00;1.30] p= .056
Income Quintile 3*Age 5 Vocabulary	1.26[1.08;1.47] * * p= .003
Income Quintile 4*Age 5 Vocabulary	1.24[1.08;1.44] * * p= .003
Income Quintile 5*Age 5 Vocabulary	1.18[.99;1.40] p= .059
Occupational Status (routine)	REFERENCE
Occupational Status (unemployed)	.83[.69;1.00] p= .053
Occupational Status (intermediate)	1.21[1.04;1.40] * p= .011
Occupational Status (higher managerial)	1.58[1.34;1.88] * * * p<.001
Occupational Status (routine)*Age 5 Vocabulary	REFERENCE
Occupational Status (unemployed)*Age 5 Vocabulary	.85[.74;.99] * p= .038
Occupational Status	1.10[.95;1.28] p= .217

Occupational Status Moderator

	(intermediate)*Age 5 Vocabulary		
	Occupational Status (higher managerial)*Age 5 Vocabulary	1.10[.95;1.27] p= .217	
	Wealth Quintile 1		REFERENCE
	Wealth Quintile 2		1.10[.92;1.33] p= .283
	Wealth Quintile 3		1.12[.91;1.37] p= .289
	Wealth Quintile 4		1.18[.96;1.45] p= .116
	Wealth Quintile 5		1.31[1.02;1.69] * p= .034
Wealth Moderator	Wealth Quintile 1*Age 5 Vocabulary		REFERENCE
	Wealth Quintile 2*Age 5 Vocabulary		1.00[.84;1.19] p= .997
	Wealth Quintile 3*Age 5 Vocabulary		1.10[.92;1.31] p= .289
	Wealth Quintile 4*Age 5 Vocabulary		1.08[.90;1.30] p= .423
	Wealth Quintile 5*Age 5 Vocabulary		1.14[.96;1.36] p= .139
Neighbourhood	Relative Neighbourhood Deprivation		REFERENCE

(most deprived decile)	
Relative Neighbourhood Deprivation (10 - <20%)	1.11[.92;1.34] p= .275
Relative Neighbourhood Deprivation (20 - <30%)	1.09[.89;1.33] p= .422
Relative Neighbourhood Deprivation (30 - <40%)	1.14[.93;1.41] p= .203
Relative Neighbourhood Deprivation (40 - <50%)	1.18[.97;1.42] p= .093
Relative Neighbourhood Deprivation (50 - <60%)	1.21[.98;1.49] p= .070
Relative Neighbourhood Deprivation (60 - <70%)	1.25[1.01;1.55] * p= .038
Relative Neighbourhood Deprivation (70 - <80%)	1.31[1.05;1.64] * p= .019
Relative Neighbourhood Deprivation (80 - <90%)	1.70[1.34;2.15] * * * p<.001
Relative Neighbourhood Deprivation	1.60[1.25;2.04] * * * p<.001

(least deprived decile)	
Relative Neighbourhood Deprivation (most deprived decile)*Age 5 Vocabulary	REFERENCE
Relative Neighbourhood Deprivation (10 - <20%)*Age 5 Vocabulary	1.07[.89;1.28] p= .466
Relative Neighbourhood Deprivation (20 - <30%)*Age 5 Vocabulary	1.04[.85;1.28] p= .698
Relative Neighbourhood Deprivation (30 - <40%)*Age 5 Vocabulary	.96[.77;1.19] p= .715
Relative Neighbourhood Deprivation (40 - <50%)*Age 5 Vocabulary	1.27[1.03;1.57] * p= .029
Relative Neighbourhood Deprivation (50 - <60%)*Age 5 Vocabulary	1.18[.95;1.47] p= .124
Relative Neighbourhood Deprivation (60 - <70%)*Age 5 Vocabulary	1.02[.82;1.27] p= .831

Relative Neighbourhood Deprivation (70 - <80%)*Age 5 Vocabulary	1.11[.89;1.39] p= .357
Relative Neighbourhood Deprivation (80 - <90%)*Age 5 Vocabulary	1.08[.83;1.39] p= .573
Relative Neighbourhood Deprivation (least deprived decile)*Age 5 Vocabulary	1.07[.83;1.37] p= .600

** Models adjusted for sex, ethnicity, EAL status, country, SEC indicators that are not the moderator variable of interest, and caregiver vocabulary scores.*

5.5. Discussion

In a series of pre-registered analyses, we investigated whether age 5 vocabulary skill predicted educational attainment, above SEC and caregiver vocabulary factors, at the end of secondary school. We also investigated whether any relation was moderated by childhood SEC. We used a large, nationally representative cohort of contemporary adolescents in the United Kingdom to investigate our research questions. Our analyses revealed that age 5 vocabulary predicted educational achievement, both in terms of achieving an important government benchmark of success (grade 4-9/ A*-C in English, Mathematics and Science subjects), and in terms of the level of achievement (regardless of pass/fail) in these core subjects. Higher age 5 vocabulary predicted an increased likelihood of achieving the benchmark, and higher average grades across the core subjects. This relation persisted when adjusting for SEC and caregiver vocabulary, and vocabulary was found to predict unique variance in educational attainment, above and beyond these factors. Further, the relation between age 5 vocabulary and achieving the benchmark of success at this level of education was moderated by SEC. However, the nature of this relation is not straightforward: the benefits conferred by strong age 5 vocabulary skills on educational attainment are greater in

the middle SEC group than in the highest SEC group, where educational attainment is high even in children with very low vocabulary scores. Nonetheless, the impact of SEC on the relation between vocabulary and educational attainment is greater in both these groups relative to the lowest SEC group, where educational attainment is poor, regardless of vocabulary.

Focusing on the first set of findings, it is clear that age 5 vocabulary predicts educational attainment at the end of compulsory schooling in the UK in this large, nationally representative sample. This is not surprising, since vocabulary lays the foundations for reading and mathematics, which are essential for later academic achievement (Ricketts et al., 2020; Slusser et al., 2019). Reading comprehension relies on knowledge of the meaning of words in a written text, and without this vocabulary skill, a reader will not be able to understand the meaning of the text, presenting a barrier to education (Elleman et al., 2019; Ricketts et al., 2020). It therefore follows that if a child has vocabulary difficulties, they may have reading difficulties which present as a barrier to accessing the curriculum, impacting their educational achievement. Relatedly, vocabulary has been shown to predict early educational attainment (Pace et al., 2019; Roulstone et al., 2010), which is likely to predict later educational achievement; for example, achievement at the end of primary school has been shown to be the strongest predictor of educational attainment in secondary school (Sutherland et al., 2015), presenting another mechanism by which vocabulary may affect education. For example, children are placed into ability sets and streams based on their prior educational achievement. This may coincide with poorer quality teaching and lower-level ability education in lower sets, that may not stimulate the same level of thinking as in higher sets and may even negatively impact educational attainment (Henry, 2015; Muijs & Dunne, 2010). Further, at GCSE level, some pupils in lower ability sets may be more likely to be entered into "Foundation Tier" examinations and therefore are restricted to a maximum grade C/ grade 4 in these examinations (although note that in the new GCSEs in England, English subjects are no longer tiered, and the new Maths GCSE foundation tier encompasses grades 1-5 (Ofqual, 2013).

Early vocabulary skill also predicts later vocabulary skill, and stability in vocabulary is experienced once children enter formal schooling school (Eadie et al., 2021; McKean, Reilly, et al., 2017), meaning those with initially high levels of vocabulary upon school entry will continue to have a greater vocabulary compared to those starting school with smaller vocabularies. As children progress through school, the need for proficient vocabulary knowledge becomes more vital as textbooks and reading materials become more complex

and academic (Nagy & Townsend, 2012; Schuth, Köhne, & Weinert, 2017), increasing demands on language ability. Individuals are increasingly likely to come across new words that are not used in their every-day lives, and a rich vocabulary will help them deduce the meaning of new words, based on the context of the words they already know (Elleman et al., 2019; Larsen & Nippold, 2007; Oxford University Press, 2020). These increasing demands are thought to highlight individual's vocabulary difficulties and individual differences in vocabulary ability mean that those with substantial vocabulary gaps will have a limited understanding of taught content and therefore difficulties in accessing the curriculum (Oxford University Press, 2020). As well as facilitating comprehension of oral and written communication, a strong vocabulary allows children to express their experiences, thoughts and knowledge, both in spoken conversation and in writing; indeed, a positive relationship between a strong vocabulary and the quality of written work has been shown to exist (Elleman et al., 2019; Graham et al., 2015). Future work would be therefore well placed to investigate the mediating role of late childhood and adolescent vocabulary in the relation between age 5 vocabulary and educational outcomes at the end of secondary school. It is worth noting that in exploratory analyses, we found age 5 vocabulary to similarly predict achievement in English, Mathematics and Science subjects, highlighting the importance of a strong vocabulary (and likely subject specific and academic vocabulary) in achievement across the curriculum, and not just in English subjects, which have a clear link to vocabulary.

Intervening on vocabulary difficulties early is thought to be important: since early vocabulary is the basis for later vocabulary skill, intervening early to lay the foundations for a strong vocabulary in adolescence is also likely to boost educational achievement at the end of secondary school. Further, those who leave primary school with poor literacy skills (which are based on vocabulary skill) are placed at a disadvantage when starting secondary school, and many never catch up with their peers (Department for Education, 2013). Relatedly, it has been argued that intervening during the first five years of a child's life ensures the highest rate of return to human capital, particularly for disadvantaged families. The earlier an investment in child development is made, the greater the return in that investment. Early intervention is key to producing a highly educated and skilled labour market and therefore improving social mobility (Heckman, 2011, 2012). Furthermore, attempts to narrow socioeconomic gaps in ability are more efficient than intervening later in childhood (Heckman, 2015). Although oral language interventions have been shown to improve reading and mathematics outcomes (Education Endowment Foundation, 2021), and a key intervention, NELI, is now being delivered at scale to reception classes in primary schools

across England (West et al., 2021), the long term effectiveness of such interventions remain largely unknown. Language interventions are often grounded on the idea that benefits will transfer to wider contexts outside of the intervention setting, and lead to maintained effects. In intervention studies that do include long term follow ups, it is thought that intervention effects are measured in terms of the same construct measured in the initial intervention, rather than a developmentally appropriate outcome (Law, Charlton, Dockrell, et al., 2017). The finding that, in a large nationally representative cohort of contemporary UK adolescents, age 5 vocabulary predicts educational outcomes, with large effect sizes, suggests that oral language interventions will plausibly have long lasting effects on educational outcomes at the end of secondary school. However, this support likely needs to continue throughout primary school and into adolescence as previous work (Thornton et al., 2021) has found that inequalities in vocabulary persist throughout childhood and into adolescence.

Caregiver vocabulary mediates the relation between factors of SEC and offspring vocabulary skill (Sullivan et al., 2021) and was found to predict educational outcomes above and beyond SEC factors, highlighting the possible role of genetics in this relation, since there is a heritable component of both vocabulary ability and educational attainment (Chow & Wong, 2021; Selzam et al., 2017); although note that a recent adoption study found input effects on vocabulary remained even in the absence of genetics (Coffey et al., 2021). The gene*environment hypothesis suggests that the influence of genetics may be supressed among lower SEC individuals (Gottschling et al., 2019; Scarr & Mccartney, 1983). Caregiver vocabulary is also indicative or parent's own abilities and may therefore reflect the ability of caregivers to engage in and help with schoolwork, which would benefit educational outcomes. SEC inequalities in vocabulary and in educational attainment are well established (Crenna-Jennings, 2018; Pace et al., 2017; Pensiero & Schoon, 2019; Sullivan et al., 2021; Thornton et al., 2021), so it is plausible that the effects of vocabulary on educational attainment are a by-product of SEC inequalities in vocabulary. However, age 5 vocabulary was found to predict educational outcomes over and above SEC and caregiver vocabulary factors: vocabulary has independent effects on educational outcomes that are not simply a reflection of inequalities in language ability. Vocabulary skill plays a long-lasting role in educational outcomes, beyond the initial inequalities in vocabulary. It is therefore plausible that focusing on improving the language abilities of all children will boost educational outcomes.

However, results from our moderation analyses suggest that the benefits of having a strong vocabulary at age 5 are greater for some groups more than others. Specifically, those in the middle SEC groups in early childhood experience greater benefits of having a large childhood vocabulary, as the steepest increase in probability of passing the threshold of ≥grade 4 on core subjects as vocabulary skill increases is observed among these groups. It is possible that parents of children in the middle SEC groups are able to exploit available resources to bolster the impact of vocabulary on educational attainment, with this effect becoming more apparent as SEC increases. For those in the highest SEC group, the probability of reaching this benchmark is already high, even at the lowest level of vocabulary skill, and continues to increase as vocabulary skill increases. This pattern of findings suggests that among those in the highest SEC quintile, sufficient resources and knowledge are accessible to compensate for vocabulary ability, whereas among other SEC groups, vocabulary ability plays an important role. However, the probability of achieving this level of achievement for those in the lowest SEC quintile is low, regardless of their vocabulary skill; even among those in this SEC group with vocabulary scores at the top of the distribution, the probability of passing this benchmark remains much lower than those from more advantaged groups. At the bottom of the vocabulary distribution, very poor levels of vocabulary at the age of 5 impede the likelihood of achieving ≥grade 4 on the core subjects at the end of secondary school at all levels of SEC, although this effect is attenuated in the highest SEC quintile. As vocabulary skill increases, so does the probability of passing this benchmark of attainment in all SEC groups, and inequalities in the effects of vocabulary on educational attainment are clear at all levels of vocabulary skill, with those in higher SEC groups being more likely to pass the benchmark threshold relative to those of a lower SEC with the same level of vocabulary skill.

The finding that across the distribution of SEC, those with very poor vocabulary skill at the age of 5 (which may be indicative of clinical levels of vocabulary difficulty) are unlikely to achieve ≥grade 4 on the core subjects, suggests that a functional level of vocabulary is necessary for educational attainment. Vocabulary forms the basis of reading and mathematics ability, which in turn are vital for accessing the curriculum. For example, as children progress through school, learning by reading becomes one of the main modes of education, and to be able to successfully obtain knowledge from reading text, one has to be able to understand the meaning of the words, rather than simply recognising the words in the text: those who struggle to understand the words they read (poor comprehenders) face difficulties in learning from reading (Ricketts et al., 2008). Poor comprehenders have also

been found to have poorer oral language scores, and to be at an educational disadvantage, compared to those who do not struggle with reading comprehension (Ricketts et al., 2014). Further, poor language is a shared risk factor for both reading and mathematics disorders, with both disorders being associated with verbal difficulties, and comorbid DLD and reading disorder, and DLD and maths disorder are thought to be high (Snowling et al., 2021). 20% of 15 year olds in England were incapable of accurately reading and understanding simple written texts (Jerrim & Shure, 2017), and since learning from reading becomes increasingly important throughout education, it is likely that these adolescents will have difficulties accessing the curriculum. Since vocabulary is vital for the initial development of reading and maths skills, it is not surprising that a functional level of vocabulary is required to increase the probability of achieving ≥grade 4 on the core subjects at the end of secondary school, regardless of family SECs.

It is worth noting that those with vocabulary at the bottom of the distribution, who are in the highest SEC quintile, are at an increased probability of reaching this benchmark, compared to their less advantaged SEC counterparts. Previous work has reported that children from disadvantaged backgrounds are both more likely to experience clinical levels of language difficulty, and are less likely to be referred for, or be in receipt of, speech and language therapy services (Bercow, 2018; Law et al., 2011; Roy et al., 2014). Furthermore, there is a lack of understanding and awareness about speech, language and communication needs, and a shortage of resources for the issue means that many children's difficulties go unnoticed, and the level of support received is "inadequate, ineffective and inequitable" (Bercow, 2018). Accessible information about speech, language and communication is necessary, but is often unavailable, with many parents reporting difficulties in accessing SLT services (Bercow, 2018), resulting in many children with speech, language and communication needs falling through the gaps. The highest SEC group is likely to be characterised by higher levels of parent education and household income, and higher levels of parent vocabulary. Those with the lowest levels of vocabulary in this SEC quintile are more likely than lower SEC quintiles to achieve ≥grade 4 on the core subjects; this could be because their parents may have an awareness of speech, language, and communication needs, and be well placed to access SLT services. High SEC families are also more likely to have a higher disposable income to provide resources, such as books, to aid vocabulary development (Duncan et al., 2017; Washbrook & Waldfogel, 2011). Being in the highest SEC quintile appears to be somewhat protective against the negative effects of poor vocabulary skill on later educational attainment, compared to lower SEC quintiles; however, this probability is

still below 50% for this group, indicating that at least a functional level of vocabulary is necessary.

However, having a functional level of vocabulary, within the average range of ability, appears to be only half of the story. Once vocabulary skill moves away from the bottom of the distribution, family SEC clearly plays an important role in the probability of whether cohort members achieved ≥grade 4 on the core subjects: regardless of vocabulary skill, the probability of reaching this benchmark is higher among higher SEC groups. Even for those with vocabulary skills at the top of the distribution, those in the most advantaged SEC quintile are almost 40% more likely to reach the benchmark than those in the lowest SEC quintile of the same ability. These differences are stark, and clearly show that ability can only get you so far in terms of educational success. In exploratory analyses, we investigated the moderating role of each SEC indicator separately (parent education, household income, occupational status, household wealth and relative neighbourhood deprivation). The overall moderation pattern remained for parent education and occupational status, although only parent education, household income and occupational status were significant moderators, indicating they may be driving the moderating effect observed by the composite indicator. However, conservative sensitivity analyses removed any moderating effect, suggesting shared variance among the individual indicators in how they moderate the relation between child vocabulary and later educational attainment, so no conclusion about any one indicator can be drawn here. The moderating effect based on the composite SEC factor also showed the largest effect sizes, which is not surprising, since this reflects a cumulative effect of all SEC indicators. Effect sizes for individual indicators were also large.

Generally speaking, a more advantaged SEC can benefit educational outcomes in multiple ways, and it is likely that the possible mechanisms by which SEC impacts educational attainment interact with each other since each indicator of SEC likely shares variance in the moderating relation between vocabulary and educational outcomes. For example, children of parents with more educational qualifications may benefit more from parental input with school work, due to their parents having the ability to do so: children from disadvantaged backgrounds experience less parental input with homework (Jerrim, 2017). Our moderation analyses adjusted for caregiver vocabulary skill, which likely reflects caregiver literacy level and therefore caregivers with greater vocabularies may be better placed to help their children with schoolwork. Similarly, disadvantaged students are less likely to benefit from private tuition, to attend high quality schools and to take part in extra-curricular activities, which have been linked to increased educational success (Dearden et al.,

2011; Donnelly et al., 2019; Kirby, 2016). It is also possible that non-academic routes may be more valued than academic qualifications among lower SEC groups, which may result in these pupils being disengaged in and undervaluing academic education (House of Commons, 2021). Those from a more disadvantaged background may hold fewer aspirations for their educational achievement than those from a higher SEC background , for example, in terms of pursuing post-16 education (Chowdry et al., 2009). It is also possible that those who had initially high vocabulary scores from lower SEC backgrounds suffered from summer learning loss, a phenomenon that refers to the decline in learning by disadvantaged pupils that is thought to occur during the summer holiday period, which has been observed in the US (Entwisle, & Olson, 2007; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996), and more recently in a sample of low SEC pupils in England and Scotland with regards to their spelling ability (Shinwell & Defeyter, 2017). Furthermore, recent Education Policy Institute research has suggested that throughout the course of the COVID-19 pandemic, all primary and secondary pupils have experienced learning losses, and this has been greater among disadvantaged pupils (Department for Education, 2021b). This speculation is a further possible explanation for the moderation effect observed here.

Furthermore, high quality early years education is important for vocabulary development and for later educational attainment, particularly among disadvantaged students (Becker, 2011; Department for Education, 2015; Schmerse, 2020). Attendance at early years settings has been the focus of government policy, with 15 hours free childcare access being introduced universally for all 3- and 4-year-olds in England in 2010, with this being extended to include 2-year-olds from disadvantaged family backgrounds in 2013. The government doubled this offer nationally for 3- and 4-year-old children of working parents in 2017 (the 30-hour childcare offer), making these children eligible for 30 hours per week of free childcare. “Working parents” are defined as those who earn, or expect to earn, the equivalent of working 16 hours a week at national minimum wage (which is around £6,000 a year and applies to both parents in the household, with an upper income limit of £100,000 per person), and accessing the free childcare is supposed to enable parents to be able to work (Paull & La Valle, 2018). However, a lack of awareness and confusion about eligibility for this offer appears to have presented as a barrier for many people, and those who are unemployed or who earn below the minimum income threshold are ineligible (Paull & La Valle, 2018; Social Mobility Commission, 2019). This means that those who are the most disadvantaged and may benefit the most from early childcare and education in terms of their vocabulary development and educational attainment are prevented from accessing free childcare. Furthermore, the

quality of early years education is inconsistent across settings (Gambaro et al., 2015) and the early years workforce is faced with staff shortages, a lack of qualifications among staff, and difficulties retaining qualified staff (Bonetti, 2019; Ceeda, 2018). Regarding a lack of qualifications among staff, in 2018, ~25% of staff in this workforce had a degree level qualification, compared to 93% of teaching staff. Further, 24.4% of early years staff have GCSE qualifications or below (Bonetti, 2019). Low qualifications among staff negatively impacts the quality of early childcare provision received. The workforce faces difficulties in staff retention due to a lack of career progression and further education opportunities, yet skill development of the workforce is important for the delivery of high quality early education (Bonetti & Brown, 2018). Since children are spending increasing amounts of time in early years settings, and high quality provision has important benefits, particularly for disadvantaged children, it is important that high quality, consistent provision is offered and accessible to all children, so that all children can start school having had the same opportunities to improve their language ability and experience the benefits of early childhood education on their later educational outcomes. Although universal early childcare was offered in 2013, when MCS2001 cohort members were already in adolescence and so did not benefit from such offers, previous research with this cohort has shown benefits for vocabulary for disadvantaged cohort members (Becker, 2011). Inequalities in access to quality early childcare provision is a further possible mechanism by which SEC can impact on later educational attainment.

Barriers to social mobility are present at every stage of a child's education, and later when they enter the workforce (Social Mobility Commission, 2019). Importantly, achievement at the end of secondary school likely perpetuates the lack of social mobility even more, therefore reducing gaps in attainment at this stage are vital for a socially mobile society and ensuring post-16 destinations for these pupils are not hampered. For example, in England, Technical Education in the form of Further Education is the most common post-16 destination for disadvantaged pupils (Department for Education, 2018c), yet a funding requirement of post-16 routes is that pupils must have ≥grade 4 (or grade C) in English and Mathematics, or have to resit these qualifications or enter functional skills qualifications (Gov.Uk, 2021). However, disadvantaged pupils are more likely to fail a second time around, or preform worse compared to more advantaged students of a similar original grade (Department of Education, 2017), limiting access to post-16 opportunities for disadvantaged pupils. Further Education colleges face difficulties retaining staff, as a result of competition of higher incomes from industry, lack of further training opportunities and a lack of qualified

staff (Department for Education, 2018a), negatively affecting the quality of education received. Furthermore, apprenticeships at the degree and higher degree level, an alternative to the academic route of university attendance, offer a cheaper, loan-free alternative to obtaining a degree, and students are often paid to gain the qualification, with a guaranteed job once the qualification is achieved. Such routes would benefit disadvantaged pupils, who are more likely to drop out of university if they are successful in gaining a place due to a myriad of factors. Yet peers from more advantaged backgrounds are more likely to benefit, due to being more likely to have the entry requirements to degree level apprenticeships, such as GCSEs in at least English and Mathematics (Social Mobility Commission, 2019).

We used vocabulary measured at age 5 as our main predictor, and it is possible, as outlined above, that adolescent vocabulary mediates the relation between age 5 vocabulary and educational outcomes at the end of secondary school. Further, the development of language throughout childhood and into adolescence is also likely patterned by SEC. Indeed, as reported in Chapter 4, inequalities in vocabulary persist and possibly widen throughout education. It is possible that those with more advantaged SECs have more opportunities to expand their vocabulary throughout childhood and into adolescence, and therefore those who have the same age 5 vocabulary skill from different SEC backgrounds could have different trajectories of vocabulary development throughout school. It is plausible that schooling can exacerbate and widen inequalities in language ability. For example, the availability of books and vocabulary-rich online content may be higher among higher SEC children. 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. As children get older, interactions with friends is the main means of vocabulary acquisition (Oxford University Press, 2020). Children are more likely to form friendships with those who are similar to them, therefore pupils from disadvantaged backgrounds may be more likely to form friendships with each other. It is thought that the need to 'code switch' between the language used with friends and that required in formal education (which is more congruent with the language use of higher SEC children) can disadvantage those whose language use is most different from that used in school, such as those from more deprived backgrounds, in terms of their educational achievement (Bernstein, 1964; Oxford University Press, 2020; Spencer et al., 2013). Therefore, it is possible that, despite high levels of vocabulary skill at age 5, those in lower SEC groups have a lower vocabulary skill later in childhood. Higher SEC groups likely have more opportunities to reach their full potential, and this is reflected in their increased

likelihood of achieving ≥grade 4 on the core subjects at the end of secondary school. In contrast, if vocabulary ability is stable relative to peers once children start school, those who start school with high levels of vocabulary may retain good vocabulary skill throughout school, but due to their low SECs, lack the resources necessary to reach this benchmark; however, vocabulary ability alone is not enough to reach the benchmark. We did not specifically look at mechanisms of how SEC affects educational outcomes here, but it is evident that SEC is important in order to reach a functional level of education at the end of secondary school, regardless of ability.

Implications

We found age 5 vocabulary ability to uniquely predict educational attainment at the end of secondary school, even after controlling for SEC and caregiver vocabulary, indicating that vocabulary skill when children enter the education system plays a substantial role for the rest of their educational careers. Since good vocabulary skill in early childhood is evidently crucial for educational attainment, even at the end of secondary school, interventions that aim to improve vocabulary are well placed to improve educational outcomes. Attainment at this stage of education is vital for social mobility, with failing to achieve ≥grade 4 (or ≥grade C) in English and Mathematics (and sometimes, additionally Science) presenting as a barrier in post-16 destinations, including subject choice at A-Level and university entry, in addition to many areas of the labour market (Lupton et al., 2021; The Children's Comissioner, 2019). Therefore, targeting the vocabulary skill of children as they enter the education system will likely benefit them as they prepare for post-16 education transitions. It is possible that vocabulary support throughout educational careers is necessary, especially for those from lower SEC backgrounds, so that those who are initially high achievers do not fall behind due to their lack of resources. Closing the word gap early in childhood is one of the Department for Education's four key ambitions for increasing social mobility through education, with the aim of making sure there is outreach particularly to disadvantaged communities who will benefit the most (Department for Education, 2017d), further highlighting the importance of vocabulary in educational attainment.

Regardless of *how* SEC affects education, it is clear that age 5 vocabulary ability alone is not enough to succeed in education. Britain claims to be a meritocratic society, which places emphasis on ability and effort, rather than family background or circumstances. However, if this were the case, those with a high vocabulary would experience high

probabilities of achieving the important government benchmark of ≥grade 4 on core subjects, regardless of SEC. In England, the government is rumoured to impose a minimum level of GCSE achievement on eligibility for student loans by 2023, making students without a grade 4/C on English and Maths ineligible for a student loan (Fazackerley, 2021). Disadvantaged pupils, regardless of ability, are less likely to reach this benchmark, and are also the students who arguably need a student loan the most to attend university. This has massive implications for social mobility, since those from disadvantaged areas are less likely to achieve this benchmark, regardless of ability, and will therefore be unable to attend university, will be unable to access highly paid jobs, and thus will remain in a disadvantaged SEC. Educational attainment is already a driver of social mobility, with poor educational attainment presenting as a barrier preventing disadvantaged people moving out of poverty (ESRC, 2012), and educational attainment is cited as the key factor causing poor children to become poor adults (DWP, 2014). Despite policies aimed at increasing social mobility, such as the introduction of the 30-hours childcare entitlement, these often fail to reach those who would benefit from them the most and have disproportionately benefitted those who are more affluent (Social Mobility Commission, 2019). The introduction of further policies, such as eligibility requirements of ≥grade 4 on English and Mathematics for student loans, will likely disproportionately impede the chances of disadvantaged pupils, and will ensure that social mobility remains stagnant.

Caregiver vocabulary was found to be a uniquely important factor in predicting educational outcomes, above and beyond family SECs. Caregiver vocabulary is a likely reflection of their wider literacy abilities, which may influence how able parents are to assist their children with schoolwork, in addition to being more aware of possible entitlements to benefits (such as free school meal eligibility), making them better equipped than less-able parents to aid in the educational attainment of their children. Targeting caregiver vocabulary and ability is a likely means to improve educational attainment and thus increase social mobility, for example through widespread adult education, which is currently sparse and concentrated among more affluent adults (Social Mobility Commission, 2019).

Limitations and strengths

There are some limitations to this research that need to be kept in mind when interpreting findings. Our educational attainment outcomes are based on self-reported qualifications: cohort members were presented with a list of qualifications and asked to

indicate which they currently had; this included, GCSEs, iGCSEs, National Four, National Five and BTECs, and how many qualifications they had. Cohort members were asked to indicate the subjects they studied and the grade they achieved, with subjects being presented to them in a list format. All possible subjects were presented and as can be seen from Table 1, there were many different options for each of the core subjects, particularly science subjects. Due to the self-reported nature of the data, we have no way of knowing for sure the objective subjects the students had taken: for example, students who sat the same combination of science subjects may have reported these in different ways – such as "Science" or "Combined Science". We therefore decided to include any of the reported core subjects. Furthermore, although we had information about BTEC qualifications available to us, we did not have the level of data needed to look at these qualifications in detail, for example, we did not have access to information about whether cohort members had a diploma, extended certificate or certificate BTEC qualification. This information would have enabled us to convert BTEC qualifications into the equivalent number of GCSE points. We used the available data to the best of our ability to include BTEC qualifications in our outcome variable. Future research could use linked National Pupil Database data (which is only available for England) to investigate whether or not the results found here replicate with a more objective version of qualification data. Relatedly, as aforementioned, different qualifications are taken in England, Wales, Scotland and Northern Ireland at the end of secondary school. We harmonised educational attainment in each country to the best of our ability. Furthermore, there are school level factors which are important for GCSE attainment, such as the type of school attended, which we did not have access to (Social Mobility Commission, 2019). Future work with linked National Pupil Database data could investigate the role of school type in the relation between vocabulary and educational attainment.

Finally, as with any longitudinal data analysis, missing data had to be accounted for. Those of a disadvantaged SES tend to be underrepresented in subsequent sweeps of cohort studies (Elliot & Shepherd, 2006; Mostafa & Wiggins, 2014), and it is possible that those with poor vocabulary may not respond to later sweeps. Further, it is possible that cohort members may not have reported grades they felt were poor and may therefore not have reported all of their qualifications, so our estimates could be underestimates of the relation. However, our analyses were sample and attrition weighted and we used multiple imputations with a rich set of auxiliary indicators to account for missing data, which is considered to be a "best effort" approach (Little & Rubin, 2002).

Despite these limitations, the strengths of this research lie in the large, nationally representative birth cohort used, which has rich data from England, Wales, Scotland and Northern Ireland, where previous work on education has often focused on individual countries. We had qualification data for contemporary adolescents, some of whom were among the first to take the new GCSEs in England. Although different qualifications are taken in each of these countries and each country has their own policy on education, which needs to be kept in mind, achievement at this stage of education is vital in each country, and we conducted sensitivity analyses on each country separately to assess whether our findings were robust in each country. Moderation analyses were only significant in England when doing this sensitivity check, so caution should be taken when interpreting the main analyses for Scotland, Wales, and Northern Ireland. It is possible that we did not have enough data from these separate countries to detect an effect. Another major strength of this work is that we were able to include a measure of caregiver vocabulary in our analyses, which likely reflects both the environment cohort members are exposed to and genetics, both of which are important for vocabulary development and educational attainment. We were also able to use a rich set of objective SEC indicators and a composite measure of SEC, constructed from five separate indicators to provide a more comprehensive estimate of SEC effects.

5.6. Conclusion

Overall, we found that in a large, nationally representative cohort across England, Wales, Scotland and Northern Ireland, early vocabulary predicts unique variance in educational attainment at the end of secondary school, an important point in education for post-16 outcomes, and for social mobility. This was the case even after adjusting for SEC and caregiver vocabulary factors. Although socioeconomic inequalities already exist in early vocabulary, vocabulary skill has a long-lasting effect beyond these initial inequalities. Thus, targeting vocabulary as a means of improving educational attainment and increasing social mobility is well founded. Further, although not investigated here, vocabulary support throughout childhood and into adolescence can only bolster these benefits further, especially when children transition into secondary education. However, we also found SEC to moderate the relation between age 5 vocabulary and later educational outcomes, such that regardless of early vocabulary ability, SEC differentials in attainment at the end of secondary school are apparent, with the benefits of a high vocabulary increasing as a function of SEC. Despite claims that the United Kingdom is a meritocratic society, clearly the family circumstances

individuals are born into still predict outcomes, regardless of skill or ability. For social mobility to be achieved, all children need to be party to the same opportunities, regardless of their background, yet clearly not all children are maximising their potential, as evidenced by our pattern of results. Whilst targeting vocabulary ability is an important part of the story, there is clearly no simple solution to improving the educational outcomes of all children, and it is important that well-intended policies reach those who would benefit from them the most.

Chapter 6 : General Discussion.

The overall aim of this thesis was to revisit socioeconomic inequalities in child language ability and to investigate any cascading effects of such differences on two important outcomes in adolescence: mental health and educational attainment. Chapter-specific discussions and interpretations of findings can be found in the relevant empirical chapters (Chapters Three to Five). This general discussion presents a reflection on the key findings of each empirical chapter taken together to provide an insight into what we have learned about the likely social value of improving language skills, and the wider implications in terms of parenting interventions that focus on promoting early language skills to improve wider functioning. Three key themes emerge: 1) the impact of vocabulary on adolescent outcomes (Chapters 3 and 5); 2) changes in British society over recent decades, as suggested by our cross-cohort comparisons (Chapters 3 and 4); and 3) the importance of vocabulary (and language skill more broadly) for social mobility (Chapters 4 and 5). Future directions and alternative avenues of research that were beyond the scope of this thesis are also discussed. Reflections on the research process, datasets and measures used are highlighted, before concluding the thesis with a summary of the unique contributions and the overall take home message.

6.1.The impact of vocabulary on adolescent outcomes

Findings from Chapter 3 suggest that in the older BCS1970 cohort, age 5 vocabulary is not predictive of adolescent self-reported internalising symptoms, whereas in the contemporary MCS2001 cohort, better age 5 vocabulary skill predicted more adolescent self-reported internalising symptoms, albeit with small effect sizes. When considering parent report, age 5 vocabulary is negatively related to adolescent self-reported symptoms, with poorer vocabulary skill predicting poorer mental health. This reporter effect is discussed at length in Chapter 3, and so is not addressed here. Overall, when considering self-reported internalising symptoms, our findings indicate that there would be little value of targeting vocabulary as a means of improving adolescent internalising mental health: where a relation does exist, this has a positive direction of effect, which is very small. However, it is worth considering possible reasons for the positive direction of effect. One such explanation could be related to academic pressure, which may have increased in recent years. Schoolwork, examinations and feeling pressured are commonly reported stressors among adolescents (Gray et al., 2011). Results of high stakes national tests and public examinations inform

school performance league tables, allowing judgements to be made about the pupil and the school, such as parental decisions about school quality founded on examination results (Allen & Burgess, 2014; Amoako, Quainoo, & Adams, 2019). Such tests and examinations may contribute to these pressures (Bradbury, 2020). Additionally, it is possible that those who have previously performed well at school may face pressure at home to continue performing to a high standard. Those who have a higher language ability in their childhood may be more likely to achieve to a higher level academically than those with lower language abilities, and therefore may face increased pressures to succeed in a contemporary environment. Furthermore, those with high language ability may possess such skills because their parents are from a higher SEC, may value success more than some lower SEC communities (see below), and may therefore place pressure on their children to succeed at school.

A clear and possible pathway presents itself here for the relation between language and adolescent internalising mental health, through pressures for high educational attainment, particularly because we found a clear link between age 5 vocabulary and achievement at the end of secondary school in Chapter 5. However, it is worth emphasising that although we observed an association where better age 5 vocabulary skill predicted poorer adolescent internalising symptoms, our effect sizes were small. Recent work with the MCS2001 sought to investigate the relations between wellbeing and performance in Key Stage 2 tests (taken by children at the end of primary school in England), and determine whether the wellbeing of pupils who sat the tests (cohort members in England) differed from pupils who do not sit such formal examinations (cohort members in Wales and Scotland) (Jerrim, 2021). This work found no evidence to suggest that taking these tests negatively impacted wellbeing, and no relations between wellbeing and performance. These findings, based on the same cohort in this thesis (MCS2001), could highlight why our effect sizes were small: although the pathway is plausible, in this cohort, school pressures may not impact internalising mental health, and therefore there is a positive but negligible association between language and internalising symptoms.

Alternatively, the observed relationship could be related to aspirations for educational achievement. Of 17-year-old students with similar low examination scores, those with typical language development stated that they were dissatisfied with their achievements, whereas those with impaired language said they were satisfied (Conti-Ramsden et al., 2009). Those with impaired language may set their sights lower than their typically developing peers. The current pattern of results may be due to those who have better language abilities aiming for higher grades than those with lower language abilities, and their mental health being

negatively impacted if they do not reach this standard. Indeed, Gustafsson et al (2010) concluded that depression may be related to not achieving aspired grades, despite spending time and effort on school work. Further, mental health literacy, which relates to the understanding of mental health (of both procuring and sustaining positive mental health, and of mental health disorders and their treatments) and to the decrease of mental health-related stigma and to the increase of effective help-seeking (Kutcher et al., 2016), may be higher among individuals with stronger vocabularies. If one is better at language in general, they may be better able to articulate and express a greater willingness to share their feelings, due to better understanding their emotions and having the vocabulary to describe them. These interpretations are merely speculation, yet findings suggest caution is needed with proposals for interventions to improve language skills as a means of improving internalising mental health, as the positive relation may indicate that such interventions could exacerbate such symptoms in adolescence. However, effect sizes are negligible, so it is unlikely that interventions aimed at improving language ability will have the desired effects when considering mental health in adolescence.

Turning to the second adolescent outcome I focussed on, educational attainment at the end of secondary school (Chapter 5), findings clearly highlight the value of improving language skills to bolster educational achievement: age 5 vocabulary skill is strongly positively related to attainment at this important stage of education. This is not surprising, since vocabulary lays the foundations for reading and mathematics, which are essential for later academic achievement (Ricketts et al., 2020; Slusser et al., 2019). The finding that, in a large nationally representative cohort of contemporary UK adolescents, age 5 vocabulary predicts educational outcomes, with large effect sizes, suggests that oral language interventions will plausibly have long lasting effects on educational outcomes at the end of secondary school.

Mental health and educational attainment are considered to be two key outcomes in adolescence. Intense, costly efforts to design and implement parenting interventions successfully at scale to improve both language ability and wider functioning has been the focus of research and policy makers in recent years (Bercow, 2018). Findings from this thesis suggest that such efforts are well placed when considering educational achievement and are potentially a good means of improving achievement in important public examinations that pave the way for the rest of one's educational and labour market career. However, in terms of improving adolescent internalising mental health, early language interventions are unlikely to have much, if any, effect, as there is no strong indication that language affects wellbeing in

the first two decades of life. There are plausible reasons why one may expect vocabulary to be predictive of mental health, leading researchers and policy makers to conclude that early language skill should be an important clinical focus for improving mental health (Bercow, 2018; Salmon et al., 2016). For example, vocabulary and narrative skill are claimed to be important for understanding the emotions and mental states of the self and others, and for regulating one's behaviour, both of which predict mental health (Robson et al., 2020; Salmon et al., 2016; Trentacosta & Fine, 2010). Language also facilitates social interaction and could be a major determinant of our ability to relate to others and maintain relations with them, which likely supports mental health. However, despite confirming the widely held notion that intervening on vocabulary skill will improve educational attainment, findings from this thesis do not support claims that this will also improve internalising mental health.

It is interesting that early vocabulary skill is important for educational attainment, but not for internalising mental health, since both are key adolescent outcomes. In the United Kingdom and United States, the increase in meritocratic beliefs in recent decades, which place emphasis on individual agency, skill and hard work in becoming successful in the labour market, has likely increased the importance of educational attainment, particularly higher education, with education being the focus of politicians such as Tony Blair and Barack Obama to achieve social mobility (Sandel, 2020). Such a focus on education, combined with the globalisation of labour markets and the economy becoming more knowledge-based, both of which increases the value of educational qualifications, undermines those without qualifications, such as those of a lower SEC, and devalues the work they produce (Sandel, 2020). If one were to conceptualise success on the basis of educational attainment and labour market success, then interventions to improve early language ability are well placed to improve the overall functioning and long-term outcomes of individuals. According to the world health organisation, poor mental health is a major world problem (World Health Organization, 2019), highlighting the importance of promoting positive mental health outcomes. Although language skill is clearly important for educational attainment, the same value is not observed in improving language skills to promote mental health outcomes in adolescence. However, BCS1970 data indicate that lower vocabulary scores at age 5 were predictive of poorer mental health at age 34, but not age 16. This pattern may indicate the emergence of a pathway, whereby language indirectly influences mental health outcomes, via its effects on educational attainment. This could reflect a need to monitor the mental health of those with poor language abilities after they leave school. However, the size of the effect on

age 34 mental health was small enough to conclude public health interventions may not be warranted in the context of targeting language to improve mental health specifically.

Education and the skills that are associated with educational success carry a large weight among high SEC communities, yet to many other communities, school achievement is not viewed as the height of success to strive towards (Rogoff et al., 2017). Different societies and cultures may value education differently: non-academic routes may be more valued than academic qualifications among lower SEC groups (House of Commons, 2021) and those from a more disadvantaged background may hold fewer aspirations for their educational achievement than those from a higher SEC background , for example, in terms of pursuing post-16 education (Chowdry et al., 2009). Individuals from lower SEC backgrounds, who are likely to have lived in the same area across multiple generations and therefore experience less geographical mobility (Kelly, 2013), may hold different aspirations to remain and work in the same area, and therefore value education less than individuals from higher SEC backgrounds (House of Commons, 2021), and some parents from White working-class backgrounds with no or low level educational qualifications may attach low value and priorities to education, and simply view sending their children to school as something they are required to do by law, which in turn may affect children's own engagement with education (Demie & Lewis, 2011, 2014).

The weight that is placed on school achievement in the labour market presents a barrier to accessing socioeconomic resources. Since the education system favours a specific dialect, which is characteristic of the middle class (Labov, 1969a), those who do not possess this dialect are disadvantaged in terms of their educational achievement and employment prospects, biasing the entire system against them. Whilst possessing these skills and the specific dialect necessary for school success, not doing so at the start of formal schooling should not be viewed as a deficit but rather something that can be learned and that can build on existing skills and practises; viewing the different skills of non-mainstream populations as a deficit penalises these individuals in terms of restricted access to socioeconomic resources, and thus contributes to stagnant social mobility (Rogoff et al., 2017). Instead of seeing lower SEC individuals who do not possesses these skills as deficient and in need of intervention, they should be viewed as individuals who still have the capacity to learn new skills, and this should be addressed in a cumulative way that adds to their existing skills, rather than undermining or eradicating existing skills (Rogoff et al., 2017; Sperry et al., 2019a). Such a strengths-based approach to teaching which recognises the abilities of different groups, may enable these pupils to succeed in school. This has been shown to be beneficial in a study of

children of LatinX immigrants in the US: teachers modified their classrooms so that they were in line with cultural practises, and in a subsequent 3 year follow up, the majority of these students were successful in assessments, with higher pass rates compared to pupils from the same communities whose classrooms had not been modified (Aidair et al., 2017). Rather than viewing the skills and abilities that children bring to school as flawed and in need of intervention, recognising these skills and playing to the strengths of specific communities may increase their school success. While this may be difficult in practise, particularly in schools that serve many different communities, this highlights flaws in the current system, which prioritise the dialect, skills and practises of those from a high SEC background.

The perception of what constitutes a problem may differ between communities; for example, a family who places less value on educational attainment may not view language as being disordered or in need of intervention, if their socioemotional wellbeing is intact. Taking a more holistic approach of individuals, and considering factors beyond educational achievement, such as wellbeing and mental functioning, paints a different picture about the value of improving early language skills: in terms of promoting internalising symptoms, improving the language ability of all children is not of an immediate concern. The prevalence of internalising difficulties is large, both in the UK and globally: for example, almost 20% of around 30,000 adolescents (aged 11-14) reported experiencing internalising difficulties, and depression is one of the main causes of disability worldwide (Deighton et al., 2018; World Health Organization, 2021). Given the scale of internalising mental health difficulties globally, and the fact that mental health is fundamental for social interactions with others and the ability to live an enjoyable life, the promotion of mental health can, and should be, viewed as a priority for individuals and wider societies across the globe (World Health Organization, 2018). This highlights the need for the mental health of adolescents to be taken into account, as well as educational attainment, when considering the long-term impact of early childhood language interventions. Although it is plausible to assume that early language skills make a good candidate for improving the wellbeing of adolescents, which would be attractive in that one would be able to kill two birds (educational attainment and wellbeing) with one stone (language interventions), findings from this thesis indicate that even where associations between early language and adolescent mental health exist, they are negligible. At least when considering language skill across the general population of children, in large, nationally representative cohorts of UK children, early ability appears to have little effect on adolescent internalising mental health, and therefore a public health intervention to improve the language skills of all children cannot be expected to have a significant impact on

adolescent internalising mental health. Instead, focussing efforts to improve the wellbeing of adolescents on providing universal access to mental health interventions, such as school-based interventions, may be more effective than early language interventions.

6.2.Changes in British society over recent decades

Chapters 3 and 4 also consist of cross-cohort comparisons between two cohorts born 30 years apart: the BCS1970 and the MCS2001. When considering the impact of SEC on vocabulary, and of early vocabulary ability on later outcomes, it is important to consider the social context in which the individual is developing. Important societal changes have occurred in the UK in the decades separating the two cohorts. Adopting a cross-cohort comparison approach provides insight into how such changes may be reflected in findings, and whether such changes have different implications for children born in different generations — how vocabulary impacts adolescent internalising symptoms (Chapter 3), or how different indicators of SEC influence vocabulary development over time (Chapter 4). For example, in Chapter 3, we found no relation between age 5 vocabulary and self-reported adolescent internalising symptoms in the BCS1970, while in the MCS2001, a small but positive relation between the two emerged. The prevalence of internalising symptoms has increased over recent decades (Bor et al., 2014; Patalay & Gage, 2019), a rise which coincides with increases in awareness of mental health: for example, the first Mental Health Awareness Week took place in 2001 (Mental Health Foundation, 2022), resulting in mental health being a more commonly addressed and openly discussed topic in modern society. This increase in mental health awareness is a societal change that may help explain the finding that there was no relation between vocabulary and adolescent self-reported internalising symptoms in the older BCS1970 cohort. Perhaps the cross-cohort difference concerning self-reported internalising symptoms in the older cohort could be a result of cohort members not being as willing or able to express their feelings or emotions, which could differ based on language ability: for example, those with lower language skill may not have the vocabulary to describe their feelings or emotions, and this could have been exacerbated if there is a lack of mental health awareness in the older cohort. Interestingly, a negative relation between vocabulary and internalising symptoms emerged in this cohort in adulthood: poorer age 5 vocabulary predicted poorer internalising mental health when cohort members were aged 34. This may plausibly be through the importance of educational qualifications for success in the labour market, with those without the language skills necessary to succeed in the education

system being disadvantaged in terms of employment, impacting mental health. The MCS2001 cohort are currently in their late teens-early twenties, so a cross-cohort comparison of adulthood internalising symptoms cannot be made at present, but given the increased importance on educational qualifications in the labour market and the economy becoming more knowledge based, it will be interesting to see if a similar pattern emerges in the contemporary cohort, whereby poor language negatively effects adulthood wellbeing.

The contrasting findings between the two cohorts indicate important cohort differences surrounding the role of language ability in adolescent internalising symptoms. Social norms, for example in terms of the acceptability and awareness of internalising symptoms, are likely to have changed between cohorts, and the relation between academic achievement and pressure to succeed has likely increased with the rise in the importance of gaining educational qualifications (Sandel, 2020). Language ability is likely to play an important role in this relation. Further, it has been suggested that meritocratic beliefs, which paint the individual as responsible for their success through hard work and talent, may have increased during the period that separates the two cohorts: since 1990, the proportion of Harvard students who believe they are wholly responsible for their success has increased (Sandel, 2020), perhaps leading to greater competition and therefore increased pressures during adolescence to take part in extracurricular activities, private tutor sessions and obtain good grades in order to receive a place at prestigious universities. The cohort difference observed in this thesis is the first step towards investigating how social norms and pressures may have changed between the BCS1970 and MCS2001 cohorts, and how the role of language ability in predicting mental health has changed between these two time points.

Changes to the education system and occupational structure of the UK are discussed at length in Chapter 4. The fact that occupational status as an indicator of SEC was a stronger predictor of vocabulary throughout childhood and into adolescence in the older BCS1970, and the increased benefit of having parents with university level qualifications in the MCS2001 points to societal changes, such as the move towards an hourglass economy, and the increased importance of educational qualifications in the labour market (Holmes & Mayhew, 2012; Sandel, 2020). The educational system has also experienced considerable change over recent decades, particularly in terms of the compulsory school leaving age, the proportions of individuals remaining in post-compulsory education and increases in university attendance (Amaranayake et al., 2000; Bolton, 2012). However, while more MCS2001 parents have university-level qualifications compared to parents of the BCS1970 cohort, still only a quarter of cohort members had parents who held such qualifications.

These societal changes reflect changes in the meanings of indicators of SEC over historical time, highlighting the importance of considering multiple measures to gain a holistic picture of the impact of SEC on vocabulary. Occupational status is becoming less valuable as a predictor, while parental university level qualifications are more clearly associated with better child language in contemporary society. 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. Alternatively, it might be that the relative importance of the various proximal causal mechanisms themselves is changing with time.

The overall finding of the cross-cohort comparison in Chapter 4 suggested little change in inequalities in vocabulary over historical time, despite changes in the meanings of SEC indicators, and despite decades of policy to reduce these inequalities. Whilst the changes in UK society between the BCS1970 and MCS2001 may have impacted the relation between vocabulary and adolescent self-reported internalising mental health, inequalities in vocabulary have generally persisted, *despite* societal changes. This finding supports proposals that interventions to lift the language skills of more disadvantaged children need to be ambitious, and scaled up considerably (Greenwood et al., 2020; Wake et al., 2012). Our cross-cohort comparison of mental health outcomes suggests that the relation between language and the outcome of internalising mental health is not clear, and in terms of promoting internalising symptoms, improving the language ability of children is not of an immediate concern when considering the outcome of mental health. The main goals of this thesis were to explore the extent to which differences in early vocabulary are associated with SEC, and the extent to which they predict positive or negative outcomes in adolescence. Cross-cohort comparisons provided further insight into these goals by allowing us to begin to hypothesise and understand the processes by which different social contexts impacted these relations. When evaluating the efficacy of intervening on early vocabulary development as a means of improving outcomes, it is important to consider the influence of the economic and social contexts in which different cohorts of people grow up in. For example, in future generations, different patterns may emerge as a result of the COVID-19 pandemic, with

inequalities in vocabulary outcomes perhaps being wider as a result of device poverty (i.e., limited access to electronic devices; this is discussed in more detail in section 6.4).

6.3.The importance of vocabulary for social mobility

Chapters 4 and 5 both highlight the importance of vocabulary for social mobility in contemporary UK society. Findings reported in Chapter 4 indicate that inequalities in language ability widen between the ages of 3 and 5, and then persist throughout childhood and into adolescence, supporting arguments for testing early interventions that seek to avoid inequalities becoming entrenched before children access formal schooling. Findings from Chapter 5 highlight the importance of strong early vocabulary skills for educational success at the end of secondary school, and importantly, those with more advantaged SECs in early childhood experience greater benefits of having a large childhood vocabulary, compared to those from more disadvantaged SECs with the same level of childhood vocabulary. The benefits conferred by strong early childhood vocabulary skills on educational attainment increases as one's SEC becomes more advantaged, with this increase being strongest among middle SEC groups relative to the highest SEC group, where educational attainment is high even in children with very low vocabulary scores. The impact of SEC on the relation between vocabulary and educational attainment is greater in both groups relative to the lowest SEC group, where educational attainment is poor, regardless of vocabulary. Once vocabulary skill moves away from the bottom of the distribution, family SEC clearly plays an important role in achieving ≥ grade 4 on the core subjects: regardless of vocabulary skill, the probability of reaching this benchmark is higher among higher SEC groups. Even for those with vocabulary skills at the top of the distribution, those in the most advantaged SEC quintile are almost 40% more likely to reach the benchmark than those in the lowest SEC quintile of the same ability.

These differences are stark, and clearly suggest that ability can only get you so far in terms of educational success. This brings us back to the notion of a meritocratic society — the assumption that you can make it if you work hard and possess the talent to do so, and that success is based on talent and skill. A meritocratic society would not eradicate inequalities, but these would be based on individual's skills and talent (their merits), rather than the circumstances in which they were born into. However, findings from Chapter 5 demonstrate that this is not the case, as possessing a strong vocabulary only benefits achievement up until a certain point, beyond which, childhood SEC is more influential. Such meritocratic beliefs can be damaging. Viewing success as something that is earnt through hard work and talent

leads to blame and negative feelings towards those who do not succeed, and yet talent appears to be only half of the picture. For example, university educated individuals in the UK, the Netherlands and Belgium were more biased against and disliked less educated people, and the less educated people had negative opinions of themselves (Kuppens et al., 2018). Beyond these issues, which are discussed at length in *The Tyranny of Merit* (Sandel, 2020), it is clear from our findings that skill and talent alone do not equate to educational success. If this were the case, those with a high vocabulary would experience high probabilities of achieving the important government benchmark of ≥grade 4 on core subjects, regardless of SEC. In England, plans to introduce an eligibility criteria based on GCSE grades for student loans by 2023 are likely to make social mobility even more stagnant (Fazackerley, 2021). The people who attend university will already be more likely to be from more affluent backgrounds, meaning the education system further entrenches inequalities based on SEC background, rather than reducing them, since those from disadvantaged areas are less likely to achieve this benchmark, regardless of ability, and will therefore be unable to attend university, will be unable to access highly paid jobs, and thus more likely to remain in a disadvantaged SEC.

Taken together, findings from Chapters 4 and 5 indicate that whilst interventions aiming to improve vocabulary are well placed to improve educational outcomes in adolescence, they are not the sole answer to improving educational attainment. Although vocabulary skill in early childhood predicted educational attainment at the end of secondary school, above and beyond SEC and caregiver vocabulary factors, SEC moderated this relation, such that regardless of early vocabulary ability, SEC differentials in attainment at the end of secondary school are apparent. Even if vocabulary skill is improved through intervention, childhood SEC is important and children from lower SEC families will not have the same experiences and opportunities as higher SEC children, even if language skills are equal. This claim is supported by the fact that in Chapter 5, cohort members with the same level of language skill in the highest SEC group had higher odds of achieving ≥ grade 4 on the core subjects than those in the lowest SEC group. Chapter 4 findings showed that inequalities in vocabulary remain persistently wide throughout primary school and into secondary school and may even widen again at the age of 14. This indicates that the development of language throughout childhood and into adolescence is likely to be patterned by SEC, for example through the provision of private tutors, help with homework, and the availability of books (Jerrim, 2017; Kirby, 2016). This is in addition to the SEC inequalities that emerge in the early childhood years (McGillion et al., 2017), and therefore vocabulary

support throughout childhood and into adolescence can only bolster these benefits further, especially as children transition into secondary education. However, it is evident that existing policies, such as universal and quality childcare provision in England are failing to reach those who would benefit from them the most, and quality is inconsistent across early childcare settings (Gambaro et al., 2015).

In addition to increasing the quality of early years educational settings, interventions such as NELI, which have shown to be effective in promoting language skills of primary school children in England at scale (West et al., 2021), and testing ways of providing sustained support that is not burdensome for families (e.g. the BBC's UK-wide Tiny Happy People programme (*Tiny Happy People*, 2021)) are needed to prevent inequalities in language ability becoming entrenched before children start formal schooling. Such interventions are likely to have long lasting impacts on educational attainment at the end of secondary school, and thus for improving social mobility. It is important that support continues for those who continue to need it, such as those from low SEC backgrounds, who are disadvantaged in terms of their attainment, even if they have initially high vocabulary skills. Although intervening to support early vocabulary development is likely to benefit children in terms of their educational outcomes, and placing vocabulary as one of the four key ambitions for increasing social mobility through education (Department for Education, 2017d) is well founded, this alone is unlikely to achieve social mobility, given the findings of moderation analyses in Chapter 5 that beyond a functional level of vocabulary, early childhood SEC appears to be more important for educational outcomes. However, as vocabulary development in childhood and adolescence is likely to be patterned by SEC (Chapter 4), continuing support throughout formal schooling may help achieve these aims.

Findings from Chapters 4 and 5 have important implications for the difference-deficit debate (see Chapter 1). It has been claimed that the different language skills of those from a lower SEC background may be seen as a deficit if they negatively impact educational outcomes, which has long-lasting effects for labour market success and outcomes in adulthood (Hoff, 2013). Findings from this thesis indicate wide inequalities in vocabulary from the age of three (Chapter 4), which clearly have long lasting impacts on educational attainment (Chapter 5), and therefore indicate that early individual differences in vocabulary, which are clearly associated with SEC, regardless of indicator, throughout development, and across historical time, impact educational attainment. From the difference-deficit perspective then, such SEC differences may be seen as a deficit. However, it is important to keep in mind, as mentioned in section 6.1, that not all individuals equally value education, and

placing such a strong emphasis on educational attainment detracts attention away from things that may be valued more by lower SEC households. For example, as evidenced by Chapter 3, individual differences in vocabulary at age 5 are not related to internalising mental health (and where they are, there is a negligible, positive association), and thus in this respect, would not be considered a deficit.

6.4.Future directions and alternative avenues

Several opportunities for future research present themselves, such as alternative avenues of investigating the overall aims, which were beyond the scope of this thesis. First, in Chapter 3, the focus was solely on internalising mental health symptoms in adolescence. We did not focus on externalising symptoms (behavioural problems) due to the high prevalence of internalising symptoms and the negative impact these can have on future functioning, both during adolescence and into adulthood (Clayborne et al., 2019). Further, this chapter consisted of a cross-cohort comparison, and adequate measures of self-reported externalizing symptoms are not available in the BCS1970. However, a different pattern of results may have emerged had we considered externalising symptoms, as research suggests a relation between language and externalising symptoms (Chow & Wehby, 2018), which may be stronger than the relation between language and internalising problems, for example through self-regulation of behaviour. There may also be an overrepresentation of people with poor language skills in the criminal justice system (Bryan, 2004). Exploring the relation between early language ability and externalising symptoms will provide insight into whether early language interventions are valuable for improving behavioural functioning.

In Chapter 3, an additional analysis of age 34 internalising symptoms in the BCS1970 indicated a negative relation between age 5 language ability and adulthood internalising mental health emerged, such that poor language in childhood predicted poorer mental health. MCS2001 cohort members are currently aged 21-22, with the age 22 sweep data collection currently underway. As these cohort members continue to transition into adulthood and enter the labour market, it will be important to see whether a similar relation between early vocabulary and adulthood mental health emerges in the MCS2001. Since in adolescence, a very small but positive relation was present, if a negative relation emerges in adulthood, this may indicate that it is through the detrimental effects poor language can have on educational attainment (Chapter 5) and thus the longer-term effects on employment prospects and social mobility, that early language affects adulthood mental health. Age 17 MCS2001 data was released after the publication of Chapter 3 in *Child Development*. However, this adolescent

age is similar to that of the BCS1970 mental health outcome (age 16), so future work could also investigate the relation at this later adolescent age, as cohort members prepared for public examinations, and post-16 transitions. Further, given the counterintuitive nature of our findings in Chapter 3, additional research on large population samples investigating the relation between vocabulary and internalising mental health are warranted.

During the age 14 sweep of the MCS2001, genetic information via saliva from cohort members, and their natural parents who were present in the home and available at the time of interview was obtained, providing DNA samples from 9259 cohort members, 8898 mothers and 5179 fathers, with 4533 genetic trios (Fitzsimons, Moulton, et al., 2020). Vocabulary skill is thought to be partly heritable (Chow & Wong, 2021), and there is also a genetic basis for educational attainment (Shakeshaft et al., 2013) and internalising mental health (Thompson et al., 2017). In Chapter 5, we aimed to control for the genetic component of vocabulary skill by adjusting for caregiver vocabulary. However, the MCS2001 provides a unique opportunity to thoroughly investigate the genetic basis for vocabulary, since it is the only nationally representative population based cohort in the UK to contain DNA samples on cohort members, their mothers and their fathers (Fitzsimons, Moulton, et al., 2020). Genome wide association analyses study a set of genetic variants across individuals, to see if specific variants are associated with a trait (such as language ability or educational attainment). From such analyses, polygenetic risk scores can be created, which are a summary estimate of the effect of genetic variants, which can then be used to assess the genetic basis of traits and predict specific phenotypes based on different genetic profiles (Choi et al., 2020; von Stumm & Plomin, 2021). Using the genetic data now available in the MCS2001 to investigate the extent to which vocabulary skill in this cohort is genetically determined, and the role this heritability has in the relation between SEC indicators and vocabulary or the relation between vocabulary and education, presents an exciting avenue for future research. Understanding the role that genetics plays in these relations may provide further insight into the efficacy of parenting interventions early in childhood.

One of the aims of this thesis was to assess the value of early childhood vocabulary in predicting internalising mental health and educational attainment in adolescence, in two large, nationally representative cohorts. Research among DLD populations suggests a relation between poor language and poorer internalising mental health, although note that this research almost exclusively focuses on parent-reported internalising symptoms (see Chapter 3), and between disordered language and educational attainment (for example, Conti-Ramsden, Durkin, Toseeb, Botting, & Pickles, 2018). The focus of this thesis was on

language skill across the distribution, beyond clinical populations, to investigate the efficacy of global provision to improve language ability in early childhood. We therefore did not specifically focus on disordered language, although exploratory analyses in Chapter 3 considered those below a specific cut off to indicate disordered language (1SD and 1.5SD below the mean on language measures). It is possible that attrition between sweeps may have been high among those with very poor language skills, and therefore those at the very bottom of the language distribution may be underrepresented in our samples, and this may explain the lack of a relation between language and internalising mental health in Chapter 3. However, as noted in Chapter 3, when using the more conservative cut off of 1.5SD below the mean, 8% of MCS2001 cohort members were classed as having vocabulary difficulties (1204 cohort members), which maps on to national prevalence levels for DLD, which are estimated to be around 7.5% (Norbury et al., 2016). Speech, language and communication needs represent the most common difficulty among pupils requiring special educational needs and disabilities (SEND) support in English schools (Department for Education, 2021a). Furthermore, language difficulties are common among many other types of SEND, such as autism and ADHD (Georgiou & Spanoudis, 2021; Mueller & Tomblin, 2012; Tomas & Vissers, 2019), and there is an association between disadvantage and SEND, with higher proportions of pupils with SEND being eligible for free school meals, compared to pupils without SEND (Department for Education, 2021a). Pupils with SEND are also more likely to experience poor internalizing mental health (BOND, 2015; Department for Education, 2018b). We did not investigate SEND status in this thesis, and given the association between disadvantage and SEND status, and comorbidities of language difficulties and other SEND, an alternative avenue that would be interesting to explore would be to specifically focus on these groups.

This work presented in this thesis was largely conducted during the COVID-19 pandemic. The impacts of the pandemic on the issues raised in this thesis, such as inequalities in language ability (Chapter 4), and the moderating role of SEC in the relation between early vocabulary and educational attainment at the end of secondary school (Chapter 5) should be explored in future research, where quality data is available. It is likely that the pandemic will have impacted these relations in several ways. In March 2020, schools and daycares across the UK closed to all pupils, except those who were children of key workers, and many parents transitioned to working from home, whilst also having to look after, and even home school, their children. The effects of this first lockdown on vocabulary development in toddlers aged 8-36 months from 13 countries has been investigated, and findings indicate that

toddlers vocabulary size (based on parent reported CDI scores) increased more during the lockdown period, relative to normative data collected before the onset of the pandemic (Kartushina et al., 2022). This could indicate that the lockdown may actually have been beneficial to vocabulary development, providing parents with more opportunities for high quality interactions, and being more aware of words their children already know and adapting input accordingly. However, this sample was skewed towards highly educated parents, so this beneficial effect may only be present in higher SEC families, who were already providing high quality input prior to the pandemic. Screen time during lockdown was also higher compared to pre-pandemic levels, particularly among the lower SEC participants (although the sample was a relatively high SEC sample), and increased screen time was associated with less vocabulary gains (Bergmann et al., 2022). However, since these samples were vastly parents from high SEC backgrounds, it is likely that SEC effects reported in these analyses are underestimates, and it is likely that those from low SEC backgrounds may have different trajectories as a result of the lockdowns.

When schools closed in March 2020, teaching rapidly changed to remote learning, with a reliance on access to computers, laptops or tablets and the internet. The digital divide, which refers to the gap between those who have access to such technology and those who do not, is likely to have widened as a result of the pandemic (Baker et al., 2020; Office for National Statistics, 2019a). Pupils from disadvantaged backgrounds were likely to have experienced more disruptions to their education, as a result of not having adequate access to technology to participate in remote, online schooling, and school closures may have widened the achievement gap (Cullinane & Montacute, 2020; Education Endowment Foundation, 2020). Further, parents from disadvantaged backgrounds may feel less confident in home schooling during the pandemic, than those from more advantaged backgrounds (Cullinane & Montacute, 2020). Children from low SEC backgrounds are likely to experience the biggest benefits of attending high quality daycare settings, perhaps because quality interactions at daycare compensate for lower quality input experienced in the home (Becker, 2011; Schmerse, 2020). Our findings from Chapter 5 allude to the stark SEC effects on education, regardless of vocabulary skill, and the closure of daycare settings and schools as a result of the pandemic may have widened inequalities in language ability and educational attainment, by exacerbating the inequalities and different learning environments experienced by children discussed at length throughout this thesis. The findings presented in this thesis may not apply to children born during the pandemic, for whom the quality of the home learning environment may be more important than ever.

Prior to the onset of the COVID-19 pandemic, I planned to use linked National Pupil Database (NPD) data for the MCS2001 in Chapter 5. This would have allowed me to use objective, official reports of GCSE results, rather than relying on self-reported qualification data. Although the use of self-reported data allowed me to investigate my research questions across the United Kingdom (NPD data is only available for England), there were several challenges with using self-reported data (see Chapter 5). Due to the sensitive nature of NPD data, it has to be accessed via an Institutional PC via a secure lab online and is not downloaded onto the PC. Due to the lockdowns in 2020, I was unable to attend the university to access this data, and although alternative arrangements were put in place by the UKDS to allow access to this data from home through remote access to an institutional computer, institutional access was still required to enable this. Delays in receiving NPD data made this plan no longer feasible given the time constraints to complete my PhD. However, future work conducting the same analyses presented in Chapter 5 using NPD data instead of self-reported qualification data, would be a valuable opportunity and would establish the extent to which our findings were a result of using self-reported outcome data. NPD data would also enable the consideration of the role of prior attainment, such as achievement in Key Stage 2 tests at the end of primary school, as a possible pathway for the relations observed in Chapter 5.

This thesis focuses exclusively on cohort studies from the United Kingdom, which are nationally representative of the population at the time they began, meaning they include a disproportionate number of individuals of a White ethnicity and a largely English-speaking background (as migrants from relevant generations are not included over time). Differences in cultural practices can lead to differences in the extent to which individual differences in language ability are recognized as deficits, which tends to be culturally defined (Norbury & Sparks, 2013). Interventions aimed at improving language ability, for example practices in speech and language therapy provision, can conflict with some cultural norms, such as those which focus on contingent talk (Norbury & Sparks, 2013). Further, since language and dialect use is culturally defined, it is important to understand the dialects of different communities in the relevant contexts (Sperry et al., 2019a). Future work should investigate the role of language ability in predicting these adolescent outcomes cross-culturally and cross-nationally to see if the same conclusions hold. This would have important implications for the value of intervening to improve early language skills, and any such interventions will need to be culturally appropriate. Ethnographic research of language and dialect use in context of different cultural groups will likely be beneficial for this purpose.

6.5.Reflections

The analyses presented in this thesis consist of secondary data analysis of two national UK birth cohort studies. Undertaking these analyses presented a steep but enjoyable learning curve at the beginning of my PhD, including learning the programming language R, and handling and cleaning multiple large, complex datasets. The research process I have engaged in has also involved open science at every stage, through pre-registration, open code, and freely available data from the UK Data Service. Throughout the duration of my PhD, several challenges associated with secondary data analysis arose, which are discussed in this section.

One reflection that is worth making concerns the measures administered to cohort members. All cohort members are given the same questionnaires, and questions may be interpreted differently by different cohort members, based for example on their SEC background or ability. It is thought that most standardised tests of language ability, for example, are developed and normed on middle class samples, and are therefore less suited to lower SEC populations. However, the British Ability Scales utilised in the MCS2001 utilised a representative norming sample in terms of eligibility for free school meals, parent education and ethnicity (Connelly, 2013). Self-reported mental health measures rely on self-awareness and the ability to recall recent feelings and experiences (Wigelsworth et al., 2010). It may be reasonable to assume that individuals with poor language ability may interpret items on these measures differently, for example related to how they are feeling when completing the survey if they have difficulties engaging in decontextualised talk. However, evidence indicates that individual items of self-report measures function similarly across different groups, including those with special educational needs or from disadvantaged backgrounds (Deighton et al., 2013; Patalay et al., 2014), and children and adolescents with DLD are capable reporters of their own experiences (Gough Kenyon et al., 2021; Owen et al., 2004; Palikara et al., 2009). Cohort studies consist of many measures and questions about different aspects of an individual's life and therefore provide a valuable opportunity to investigate the impact of early life experiences on later outcomes, and possible pathways that may exist for these relations. We did not consider pathways between vocabulary and adolescent internalising mental health, or vocabulary and educational attainment, as such questions were beyond the scope of this thesis, which sought to investigate the existence of a relation between vocabulary and these outcomes. Whilst it would be interesting to explore mediators of any relation, the essential first question is, for example, whether a relation between language and

internalising symptoms does indeed hold beyond populations with clinically significant language delay. This question must first be addressed before any mediation analyses could be conducted.

In cohort studies, measures are administered without a specific purpose in mind and can be used by multiple disciplines to investigate different research questions. Measures employed by cohort studies vary, both between cohorts and within cohorts, such as the measures of vocabulary and mental health administered to cohort members, and analyses are restricted to what has been collected in a given cohort. For example, Chapter 4 looked at the impact of different indicators of SEC on vocabulary throughout childhood and into adolescence. An initial analysis plan was to assess the rate of change in vocabulary ability within individuals, by conducting multi-level modelling with vocabulary score at each time point nested within the individual. However, given the different vocabulary measures used at each age (see Chapter 2 for details) and the fact that we did not have a vocabulary measure available for all childhood age points (at age 7, data was collected regarding reading, but not vocabulary, ability), we decided that the available data did not adequately allow us to address inequalities in the rate of change in vocabulary development for individual cohort members. The developmental appropriateness of such vocabulary tests is a likely reason for why different measures were used at each age. I focussed this research question on whether or not one's position in the language distribution changes at each age, and how much of this is a function of SEC; because I standardised my vocabulary outcome measures, my coefficients were directly comparable, allowing me to assess relative vocabulary skill at each developmental time point of interest.

Another reflection regarding the availability of data within cohort studies concerns the use of public examination grades as the outcome variable in Chapter 5. We used self-reported grades (see Chapter 5 for a discussion of the challenges this presented), operationalised as two different outcome variables: a binary variable of whether or not cohort members achieved an important government benchmark, and overall performance regardless of pass or fail. However, recent work from the Education Endowment Foundation evaluating the use of GCSE grades as the outcome variable in intervention studies, instead of the raw marks which underlie these grades, suggests that using grades, which are a summary of multiple marks, result in a loss of power when evaluating the efficacy of education interventions (Smith et al., 2021), necessitating larger samples to detect an intervention effect. Although we did not evaluate interventions, this finding could indicate our results are an underestimation of the true effect of language on educational attainment. Further, GCSE grades are awarded based

on grade boundaries, which differ from subject to subject, and year to year, based on the overall national performance in the public examinations (Pearson Qualifications, 2022). This means that in any given subject, fewer marks may equate to a higher grade in that subject, and the same marks may equate to a lower grade in another subject. If marks on different examinations were available in conjunction with grades, a deeper understanding of the relation between vocabulary and educational attainment could be gained. However, we only had the overall grades available to us. Finally, Chapter 5 almost exclusively focuses on academic attainment. Although alternatives to GCSEs or N5s were considered (Science BTECs), we did not look at the relation between vocabulary and more vocational routes and qualifications, which may be more valued or pursued more by those of a lower SEC. However, the level of data available on vocational qualifications available in the self-reported MCS2001 qualifications data was not detailed enough to allow us to include them in our analyses.

In the empirical chapters of this thesis (Chapters 3 to 5), we considered one cohort member per family. Chapters 3 and 4 used the 7th edition of the third MCS2001 sweep, which is when the measure of age 5 vocabulary was collected (this variable was used for sample selection in Chapters 3 and 5) (Centre for Longitudinal Studies, 2017). The previous edition of the MCS2001 data included cohort members who were singletons, twins and triplets available under End User License (the standard agreement to access data from the UK data service). However, in 2020, the latest edition of MCS2001 data was released (Centre for Longitudinal Studies., 2021). In addition to converting all existing MCS2001 data to long format, triplets were no longer available under End User License, meaning they were removed from the data. The latest edition of the data was used in Chapter 5, meaning the number of age 5 vocabulary responses available under End User License was smaller than what was available in the previous 2 chapters. However, there are only 10 sets of triplets in the MCS2001, meaning the vocabulary responses was only reduced by 10 cohort members. A major challenge that presented itself throughout the course of my PhD concerned cross-cohort comparisons. As noted above, different cohorts measure things in different ways, and the meaning of some measures may change over time, such as SEC as discussed in Chapter 4. This meant that available data had to be harmonised so that valid comparisons between the BCS1970 and MCS2001 could be made. Data harmonisation is the process of making data from different sources (such as different cohorts) more similar to improve comparability between cohorts (O'Neill et al., 2020). The harmonisation of SEC indicators in the BCS1970 and MCS2001 is discussed in Chapter 4, which involved creating similar parent education

and occupational status variables in both cohorts. In both cases, a new scale was created for each cohort, that was made up of similar responses, to create a comparable variable. However, it must be noted that despite extensive efforts to harmonise my variables, historical changes regarding occupational status and parent education (see Chapter 4) make it difficult to definitively compare results across the two cohorts, and such differences should be kept in mind when interpreting the results of this thesis. Extensive data harmonisation also occurred in Chapter 5, since different public examinations are taken in England, Scotland, Wales, and Northern Ireland (see Chapter 5 for a discussion). A lot of thought went into how best to conceptualise attainment at the end of secondary school given these differences. The decision to use a binary variable of whether or not cohort members achieved ≥grade 4/C on English, Mathematics and Science subjects meant that we could calculate these for GCSEs and N5s separately, and then combine these into one variable. Similarly, average grades across these subjects were calculated for GCSEs and N5s separately and standardised before multiple imputation, meaning they were standardised to the relevant populations. This was important, as the range of grades for GCSEs was 1 (lowest) – 9 (highest), whereas for N5s, the highest grade was a grade A (5), with possible grades being A, B, C, D, or no award. Given the extensive and complicated harmonisation process that took place for the outcome variable in Chapter 5, the decision was made to drop the cross-cohort element of this chapter, which would have required further harmonisation procedures. On this matter, it is also worth noting that educational attainment data in the BCS1970 is in the form of a retrospective question asked when cohort members were aged 26, about which qualifications, if any, they had achieved, either at school or since leaving school.

The final reflection that I will make is a positive one, which considers the contribution of the work in this thesis to open science, which refers to making knowledge transparent and replicable, and increases the integrity of results, as people are deterred from p-hacking (running different analyses until a significant finding emerges) and hypothesising after the results are known (Allen & Mehler, 2019). There is concern in the literature that many published findings are difficult to replicate (Ioannidis, 2018; Iqbal et al., 2016), and published literature is likely to have suffered from publication bias, whereby a certain pattern of results is more likely to be published than others: for example, statistically significant findings are more likely to be published than null results (Jerrim & De Vries, 2017). Further, results that a contrary to expectations are unlikely to be published, and can lead to researchers selectively reporting results, or research not being accepted by peer review. An example of this occurred in this thesis concerning the findings of Chapter 3. There is a general belief that language is

important for internalising mental health, supported by research on populations with DLD suggesting such a link. However, as discussed in Chapter 3, the majority of these studies used measures of parent-reported symptoms. A recent meta-analysis on the relation concluded there was a negative relation between language ability and internalising mental health (Hentges et al., 2021), however, a closer look at the studies included in this meta-analysis revealed that measures of parent report were utilised, or correlations between language and mental health were evident when different research questions were investigated. Our finding that good early vocabulary skill predicted poorer internalising mental health in adolescence received resistance from peer reviewers and audiences at academic conferences, because it was against expectations, and indeed was opposite to our hypotheses. A previous study with MCS2001 data found a similar positive relation between childhood cognitive ability and adolescent internalising symptoms (Patalay & Fitzsimons, 2018). However, it is surprising that a greater amount of similar findings with the MCS2001 do not exist, given the extensive use of both the cognitive and mental health measures used in research with this cohort. This presents the possibility that previous research has engaged in the selective reporting of results.

All analyses contained within this thesis were pre-registered with clear statistical analysis plans. Such open science practises can increase confidence in our findings. The confidence in the findings presented in this thesis is further increased, due to the rigorous sensitivity analyses carried out for each chapter (Jerrim & De Vries, 2017), investigating different ways of dealing with missing data (complete case analyses) and different ways of conceptualising key variables (for example, a binary measure of language and a binary outcome of mental health in Chapter 3, given the counter-intuitive findings). Furthermore, making the finished papers available as pre-prints at the time of journal submission meant that our findings were in the published domain, protecting against publication bias. Finally, making all programming code used in data cleaning and analyses further enhances transparency in the research process, and enables reproducibility of findings. All code used in this thesis has been made openly available, either on the Open Science Framework, or on GitHub, and all data used is openly available on the UK Data Service. Making the entire research process of this thesis transparent and reproducible demonstrates a commitment to open science, and is a major strength of this thesis.

6.6.Conclusion

The overall aim of this thesis was to examine the extent to which differences in early vocabulary are associated with SEC and predict positive or negative outcomes in adolescence, to provide insight into whether intervening to support early vocabulary development is likely to benefit children in terms of their educational outcomes and wellbeing. Taking advantage of two large, nationally representative British birth cohorts, we addressed several important gaps in the literature, and each chapter makes important contributions.

In Chapter 3, although findings were surprising and counterintuitive, we found that in a contemporary cohort, better age 5 vocabulary predicted poorer self-reported adolescent internalising mental health, although effect sizes were small. I found important reporter effects, emphasising the value of self-reported measures and the need to listen to individuals when they report on their mental health. The findings of Chapter 3 also led to the discovery that much of the research on the relation between language and mental health to date focuses on parent-reported measures of mental health, calling attention to the need for more work with self-reported outcomes, in order to advance the field.

In Chapter 4, I found substantial individual differences in child and adolescent language were explained by several SEC indicators each making their own unique contribution, most notably parent education, income and occupational status. Inequalities persist from ages 3 to 14 years, with SEC indicators explaining 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. These inequalities appear to be stable over historical time, despite increases in policies aimed at reducing them. Findings of this chapter highlighted the need to focus on the widening of inequalities as children enter compulsory education and as they prepare to leave it, and the importance of sustained support.

In Chapter 5, I found early vocabulary to predict unique variance in educational attainment at the end of secondary school, above and beyond SEC and caregiver vocabulary factors: although socioeconomic inequalities already exist in early vocabulary, vocabulary skill has a long-lasting effect beyond these initial inequalities. These findings were apparent in all four countries of the UK, despite differences in education systems. I also found SEC to moderate the relation between age 5 vocabulary and later educational outcomes, such that regardless of early vocabulary ability, SEC differentials in attainment at the end of secondary school are apparent.

Taken together, the key findings of this thesis show that, whilst early language interventions are well placed to improve educational attainment, where associations between

early language and adolescent mental health exist, they are negligible, making it unlikely that interventions aimed at improving language ability will have the desired effects when considering mental health in adolescence. Furthermore, early language interventions will only likely improve educational attainment to a certain extent, as moderation analyses in Chapter 5 suggested that once vocabulary skill moves away from the bottom of the distribution, family SEC clearly plays an important role in educational achievement. Therefore, interventions to reduce socioeconomic inequalities, such as reduced poverty, caregiver lifelong learning, and efforts to make sure that interventions reach those who will benefit most, are also necessary, otherwise there is the risk that early interventions may unintentionally widen inequalities. Given that inequalities in vocabulary are persistent throughout childhood and into adolescence (Chapter 4), those from lower SEC backgrounds will likely benefit from continuous support throughout school, both in terms of their vocabulary ability and subsequently their educational attainment.

Good language ability, as measured by standardised tests, is important for educational attainment but not internalising mental health. It could be that the system is inherently flawed, and benefits those who speak a dialect characteristic of the middle class, placing them at an unrivalled advantage in their education and labour market chances. The same dialect that allows those from a higher SEC background to prosper in education is likely not necessary for good mental health, hence we do not see the same pattern of results when predicting adolescent internalising symptoms. Although early childhood language interventions are well placed to improve educational outcomes and thus social mobility to an extent, educational attainment is not the whole story of a positive adolescence, and is not valued equally by all. In order to improve wider functioning in adolescence, we also need to directly target internalising mental health and structural inequalities.

References

Aidair, J. K., Colegrove, K. S. S., & McManus, M. E. (2017). How the word gap argument negatively impacts young children of LatinX immigrants' conceptualisations of learning. *Harvard Educational Review*, *87*(3), 309–334.

Akhtar, N., & Jaswal, V. K. (2013). Deficit or difference? Interpreting diverse developmental paths: an introduction to the special section. *Developmental Psychology*, *49*(1), 1–3. https://doi.org/10.1037/a0029851

Alcock, K. J., Meints, K., & Rowland, C. F. (2020). *The UK Communicative Development Inventories: Words and Gestures* (pp. 1–12). J&R Press.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review*, *72*(2), 167–180. https://doi.org/10.1177/000312240707200202

Alexander, R. (2020). *A Dialogic Teaching Companion*. Routledge.

Allen, C., & Mehler, D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLOS Biology*, *17*(5), e3000246. https://doi.org/10.1371/journal.pbio.3000246

Allen, R., & Burgess, S. (2014). *Parental Choice of School : Secondary Data Analysis*. *January*.

Amaranayake, N., Church, J., Hill, C., Jackson, J., Jackson, V., Myers, C., Saeed, Z., Shipsey, C., & Tony Symmonds. (2000). *Social trends 30*.

Ambridge, B., & Lieven, E. V. M. (2011). Child Language Acquisition. Contrasting Theoretical Approaches. In *Understanding Linguistic Fieldwork*. Cambridge University Press.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed.)*.

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders DSM-5. In *Encyclopedia of Applied Psychology, Three-Volume Set* (Fifth). American Psychiatric Association. https://doi.org/10.1016/B0-12-657410-3/00457-8

Amoako, I., Quainoo, E., & Adams, F. (2019). High-Stakes Test Accountability: A Controversial Issue in Educational Measurement. *International Journal of Latest Research in Humanities and Social Science*, *2*(10), 94–96.

Anders, J., & Henderson, M. (2019). Socioeconomic Inequality and Student Outcomes in English Schools. In L. Volante, S. Schnepf, J. Jerrim, & D. Klinger (Eds.),

Socioeconomic Inequality and Student Outcomes. (Vol. 4, pp. 41–62). Springer. https://doi.org/10.1007/978-981-13-9863-6_3

Anders, J., Henderson, M., Moulton, V., & Sullivan, A. (2018). The role of schools in explaining individuals' subject choices at age 14. *Oxford Review of Education*, *44*(1), 75–93. https://doi.org/10.1080/03054985.2018.1409973

Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary Development: A Morphological Analysis. *Monographs of the Society for Research in Child Development*, *58*(10), i. https://doi.org/10.2307/1166112

Angold, A., Costello, E. J., & Messer, S. C. (1995). The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. In *International Journal of Methods in Psychiatric Research* (Vol. 5, pp. 237–249).

Anthony, J. L., Davis, C., Williams, J. M., & Anthony, T. I. (2014). Preschoolers' oral language abilities: A multilevel examination of dimensionality. *Learning and Individual Differences*, *35*(October), 56–61. https://doi.org/10.1016/j.lindif.2014.07.004

APA. (2007). Report of the APA Task Force on Socioeconomic Status. In *American Psychological Association Task Force on Socioeconomic Status* (Vol. 23).

Arriaga, R. I., Fenson, L., Cronan, T., & Pethick, S. J. (1998). Scores on the MacArthur Communicative Development Inventory of children from lowand middle-income families. *Applied Psycholinguistics*, *19*(2), 209–223. https://doi.org/10.1017/s0142716400010043

Asmussen, K., Law, J., Charlton, J., Acquah, D., Brims, L., Pote, I., & Mcbride, T. (2018). *Key competencies in early cognitive development Things, people, numbers and words*. *December*. https://www.eif.org.uk/report/key-competencies-in-early-

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

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*(1), 40–49. https://doi.org/10.1002/mpr.329

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

Baker, C., Hutton, G., Christie, L., & Wright, S. (2020). *COVID-19 and the digital divide*.

Banks, J., Blundell, R., & Smith, J. P. (2004). Wealth Portfolios in the United Kingdom and the United States. In D. A. Wise (Ed.), *Perspectives on the Economics of Aging* (Issue June, pp. 205–246). University of Chicago Press.

Barnett, S. E., Levickis, P., McKean, C., Letts, C., & Stringer, H. (2021). Validation of a measure of parental responsiveness: Comparison of the brief Parental Responsiveness Rating Scale with a detailed measure of responsive parental behaviours. *Journal of Child Health Care*. https://doi.org/10.1177/1367493521996489

Bassok, D., Finch, J. E., Lee, R., Reardon, S. F., & Waldfogel, J. (2016). Socioeconomic Gaps in Early Childhood Experiences. *AERA Open*, *2*(3), 233285841665392. https://doi.org/10.1177/2332858416653924

Bassok, D., & Galdo, E. (2016). Inequality in Preschool Quality? Community-Level Disparities in Access to High-Quality Learning Environments. *Early Education and Development*, *27*(1), 128–144. https://doi.org/10.1080/10409289.2015.1057463

Bates, E., & Goodman, J. C. (1997). On the Inseparability of Grammar and the Lexicon: Evidence from Acquisition, Aphasia and Real-time Processing. *Language and Cognitive Processes*, *12*(5–6), 507–584. https://doi.org/10.1080/016909697386628

Bates, E., Thal, D., Finlay, B., Clancy, B., Origins, P. D.-, & Int, N. S. F. (2003). Early Language Development and Its Neural Correlates. In I. Rapin & S. Segalowitz (Eds.), *Handbook of Neuropsychology, Vol. 6, Child Neurology (2nd edition)* (Vol. 22343). Elsevier.

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.

Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., Escobar, M., & Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(1), 75–82. https://doi.org/10.1097/00004583-200101000-00019

Benton, T. (2015). Can we do better than using 'mean GCSE grade' to predict future outcomes? An evaluation of Generalised Boosting Models. *Oxford Review of Education*, *41*(5), 587–607. https://doi.org/10.1080/03054985.2015.1074563

Bercow, J. (2008). The Bercow Report. *Communication*, 139. http://www.mendeley.com/research/bercow-report/

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

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bergmann, C., Dimitrova, N., Alaslani, K., Almohammadi, A., Alroqi, H., Aussems, S., Barokova, M., Davies, C., Gonzalez-Gomez, N., Gibson, S. P., Havron, N., Horowitz-Kraus, T., Kanero, J., Kartushina, N., Keller, C., Mayor, J., Mundry, R., Shinskey, J., & Mani, N. (2022). Young children's screen time during the first COVID-19 lockdown in 12 countries. *Scientific Reports*, *12*(1), 1–15. https://doi.org/10.1038/s41598-022-05840-5

Berman, R. A. (2007). Language Use Across Adolescence. In M. (Eds. . Hoff, E., Shatz (Ed.), *Handbook of Language Development.* Blackwell Publishing, London.

Bernstein, B. (1964). Elaborated and Restricted Codes : Their Social Origins and Some Consequences. *American Anthropological Association*, *66*(6), 55–69.

Bernstein, B. (1970). Social class, language and socialization. In *Theoretical Studies Towards a Sociology of Language* (pp. 144–158). https://doi.org/10.4324/9780203014035-18

Bernstein, B. (1971). *Class, codes and control: Theoretical Studies Towards a Sociology of Language*. Routledge & Kegan Paul.

Bishop, D. V. M. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., Baird, G., Bauer, A., Bellair, J., Boyle, C., Brownlie, E., Carter, G., Clark, B., Clegg, J., Cohen, N., Conti-Ramsden, G., Dockrell, J., Dunn, J., Ebbels, S., … Whitehouse, A. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Blackburn, R. M., & Jarman, J. (1993). *Changing Inequalities in Access to British Universities*. *19*(2), 197–215.

Blackwell, D. L. (1998). Marital Homogamy in the United States: The Influence of Individual and Paternal Education. *Social Science Research*, *27*(2), 159–188. https://doi.org/10.1006/ssre.1998.0618

Blanden, J., & Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, *51*(2), 230–249. https://doi.org/10.1111/j.0036-9292.2004.00304.x

Blanden, J., & Machin, S. (2010). Intergenerational inequality in early years assessments. In K. Hansen, H. Joshi, & S. Dex (Eds.), *Children of the 21st Century: The First Five Years* (Vol. 2, Issue 0, pp. 153–168). Bristol University Press. https://doi.org/10.2307/j.ctt9qgrfx.17

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*(6), 1461–1476. https://doi.org/10.1017/S0142716416000060

Blossfeld, H. P., & Buchholz, S. (2009). Increasing Resource Inequality among Families in Modern Societies : The Mechanisms of Growing Educational Homogamy , Changes in the Division of Work in the Family and the Decline of the Male Breadwinner Model. *Journal of Comparative Family Studies*, *40*(4), 603–616.

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

BOND. (2015). *Children and Young People with Learning Disabilities - Understanding their Mental Health*.

Bonetti, S. (2019). *The early years workforce in England*. Education Policy Institute.

Bonetti, S., & Brown, K. (2018). *Structural elements of quality of early years provision: A review of the evidence*. Education Policy Institute.

Bor, W., Dean, A. J., Najman, J., & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian and New Zealand Journal of Psychiatry*, *48*(7), 606–616. https://doi.org/10.1177/0004867414533834

Bornstein, M. H., Cote, L. R., Haynes, O. M., Hahn, C. S., & Park, Y. (2010). Parenting knowledge: Experiential and sociodemographic factors in European American Mothers of Young Children. *Developmental Psychology*, *46*(6), 1677–1693. https://doi.org/10.1037/a0020677

Bornstein, M. H., Hahn, C.-S., & Putnick, D. L. (2016a). Long-term stability of core language skill in children with contrasting language skills. *Developmental Psychology*, *52*(5), 704–716. https://doi.org/10.1037/dev0000111

Bornstein, M. H., Hahn, C.-S., & Suwalsky, J. T. D. (2013). Language and Internalizing and Externalizing Behavioral Adjustment: Developmental Pathways from Childhood to Adolescence. *Developmental Psychology*, *25*(3), 857–878. https://doi.org/10.1017/S0954579413000217.

Bornstein, M. H., Hahn, C. S., & Putnick, D. L. (2016b). Stability of core language skill

across the first decade of life in children at biological and social risk. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(12), 1434–1443. https://doi.org/10.1111/jcpp.12632

Bornstein, M. H., Hahn, C. S., Putnick, D. L., & Suwalsky, J. T. D. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development*, *85*(4), 1346–1356. https://doi.org/10.1111/cdev.12192

Bornstein, M. H., & Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *New Directions for Child and Adolescent Development*, *1989*(43), 49–61. https://doi.org/10.1002/cd.23219894306

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and Me: Familiar Names Help Launch Babies Into Speech-Stream Segmentation. *Psychological Science*, *16*(4), 298–304. https://doi.org/10.1111/j.0956-7976.2005.01531.x

Bourdieu, P. (2000). The Production and Reproduction of Legitimate Language (1982). In *The Routledge Language and Cultural Theory Reader* (pp. 467–477).

Boutwell, B. B., Meldrum, R. C., & Petkovsek, M. A. (2017). General intelligence in friendship selection: A study of preadolescent best friend dyads. *Intelligence*, *64*(July), 30–35. https://doi.org/10.1016/j.intell.2017.07.002

Bradbury, A. (2020). Pressure, anxiety and collateral damage: The headteachers' verdict on SATs. *More than a Score*, 1–60. www.moerthanascore.org.uk

Brenne, E., & Rimehaug, T. (2019). Pragmatic language impairment general and specific associations to mental health symptom dimensions in a child psychiatric sample. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *7*(1), 1–10. https://doi.org/10.21307/sjcapp-2019-001

Brimer, M. A., & Dunn, L. M. (1962). Manual for the English Picture Vocabulary Tests: Test 1 (age Range 5: 0-8: 11), Test 2 (age Range 7: 0-11: 11). *Educational Evaluation Enterprises.*

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

Brooks-Gunn, J., Han, W.-J., & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: The NICHD study of early child care. National Institute of Child Health and Human Development. *Child Development*, *73*(4), 1052–1072. https://doi.org/10.1111/1467-8624.00457

Bryan, K. (2004). Preliminary study of the prevalence of speech and language difficulties in young offenders. *International Journal of Language and Communication Disorders*,

39(3), 391–400. https://doi.org/10.1080/13682820410001666376

Buckley, B. (2003). *Children's Communication Skills From birth to five years*. Routledge.

Bukodi, E., Bourne, M., & Betthäuser, B. (2017). Wastage of talent?: Social origins, cognitive ability and educational attainment in Britain. *Advances in Life Course Research*, *34*, 34–42. https://doi.org/10.1016/j.alcr.2017.09.003

Bukodi, E., Bourne, M., Betthäuser, B. A., Goldthorpe, J. H., Bukodi, E., Betthäuser, B. A., & Goldthorpe, J. H. (2018). 'Persistence of the social': The role of cognitive ability in mediating the effects of social origins on educational attainment in Britain. *Research in Social Stratification and Mobility*, *58*(July), 11–21. https://doi.org/10.1016/j.rssm.2018.09.001

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference.* Springer-Verlag New York, Inc.

Burston, K., Cleary, A., Colin, G., Michelmore, O., Sheriffs, P., Conolly, A., Dangerfiled, P., Knibbs, S., Rose, N., & Tietz, S. (2017). *Millennium Cohort Study Sixth Sweep (MCS6)-Technical Report*. *2017*(February), 133.

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

Callanan, M., & Waxman, S. (2013). Commentary on special section: deficit or difference? Interpreting diverse developmental paths. *Developmental Psychology*, *49*(1), 80–83. https://doi.org/10.1037/a0029741

Campbell, T. (2015). Stereotyped at seven? Biases in teacher judgement of pupils' ability and attainment. *Journal of Social Policy*, *44*(3), 517–547. https://doi.org/10.1017/S0047279415000227

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

Ceeda. (2018). *About early years: Early Years Sector Skills Survey*. https://www.aboutearlyyears.co.uk/media/1317/about-early-years-sector-skills-survey.pdf

Centre for Longitudinal Studies. (2021). *Millennium Cohort Study: Third Survey, 2006. [data collection]. 8th Edition.* https://doi.org/http://doi.org/10.5255/UKDA-SN-5795-5

Centre for Longitudinal Studies. (2017). *Millennium Cohort Study: Third Survey, 2006. [data collection]. 7th Edition.* UK Data Service. SN: 5795. https://doi.org/http://doi.org/10.5255/UKDA-SN-5795-4

Centre for Longitudinal Studies. (2020). *Millennium Cohort Study: User Guide (Surveys 1-*

5): Vol. 9th ed (Issue August). https://cls.ucl.ac.uk/wp-content/uploads/2020/09/MCS1-5_User_Guide_ed9_2020-08-07.pdf

Centre for Longitudinal Studies, U. I. of E., & Ipsos. (2019). *Millennium Cohort Study Seventh Sweep (MCS7) Technical Report*. *November*, 1–141. http://www.cls.ioe.ac.uk/shared/get-file.ashx?id=3284&itemtype=document

Centre for Social Justice. (2013). *The Causes of Educational Failure* (Issue September). www.centreforsocialjustice.org.uk

Chaplin-Gray, J., Gatenby, R., Simmonds, N., & Huang, Y. (2010). *Millennium Cohort Study Sweep 2 - Technical Report*. *44*(0), 1–98.

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

Cherng, H. Y. S., Calarco, J. M. C., & Kao, G. (2013). Along for the Ride: Best Friends' Resources and Adolescents' College Completion. *American Educational Research Journal*, *50*(1), 76–106. https://doi.org/10.3102/0002831212466689

Chiat, S., & Roy, P. (2008). Early phonological and sociocognitive skills as predictors of later language and social communication outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(6), 635–645. https://doi.org/10.1111/j.1469-7610.2008.01881.x

Choi, S. W., Mak, T. S. H., & O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols*, *15*(9), 2759–2772. https://doi.org/10.1038/s41596-020-0353-1

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

Chow, J. C., & Wehby, J. H. (2018). Associations Between Language and Problem Behavior: a Systematic Review and Correlational Meta-analysis. *Educational Psychology Review*, *30*(1), 61–82. https://doi.org/10.1007/s10648-016-9385-z

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: Analysis using linked administrative data. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *176*(2), 431–457. https://doi.org/10.1111/j.1467-985X.2012.01043.x

Chowdry, H., Crawford, C., Dearden, L., Joyce, R., Sibieta, L., Sylva, K., & Washbrook, E. (2010). Poorer children's educational attainment: How important are attitudes and behaviour? *Joseph Rowntree Foundation*, 1–72.

Chowdry, H., Crawford, C., & Goodman, A. (2009). *Drivers and Barriers to Educational Success: Evidence from the Longitudinal Study of Young People in England*.

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(1), 72–79. https://doi.org/10.1016/j.jaac.2018.07.896

Clegg, J., & Ginsborg, J. (2006). *Language and Social Disadvantage*. John Wiley & Sons.

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

Coffey, J. R., Shafto, C. L., Geren, J. C., & Snedeker, J. (2021). The effects of maternal input on language in the absence of genetic confounds: Vocabulary development in internationally adopted children. *Child Development*, 1–17. https://doi.org/10.1111/cdev.13688

Conger, R. D., & Conger, K. J. (2002). Resilience in Midwestern families: Selected findings from the first decade of a prospective, longitudinal study. *Journal of Marriage and Family*, *64*(2), 361–373. https://doi.org/10.1111/j.1741-3737.2002.00361.x

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

Conley, D. (2001). Capital for College : Parental Assets and Postsecondary Schooling. *Sociology of Education*, *74*(1), 59–72.

Connelly, R. (2013). *Millennium Cohort Study data note 2013/1: Interpreting test scores*. *September*, 1–20.

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

Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(5), 516–525. https://doi.org/10.1111/j.1469-7610.2007.01858.x

Conti-Ramsden, G., Durkin, K., Simkin, Z., & Knox, E. (2009). Specific language impairment and school outcomes. I: Identifying and explaining variability at the end of compulsory education. *International Journal of Language and Communication Disorders*, *44*(1), 15–35. https://doi.org/10.1080/13682820801921601

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language and Communication Disorders*, *53*(2), 237–255. https://doi.org/10.1111/1460-6984.12338

Conti-Ramsden, G., Mok, P. L. H., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities*, *34*(11), 4161–4169. https://doi.org/10.1016/j.ridd.2013.08.043

Conway, L. J., Levickis, P. A., Mensah, F., McKean, C., Smith, K., & Reilly, S. (2017). Associations between expressive and receptive language and internalizing and externalizing behaviours in a community-based prospective study of slow-to-talk toddlers. *International Journal of Language and Communication Disorders*, *52*(6), 839–853. https://doi.org/10.1111/1460-6984.12320

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227–268. https://doi.org/10.3102/00346543066003227

Corlett, A., Clarke, S., Mccurdy, C., Rahman, F., & Whittaker, M. (2019). *The Living Standards Audit 2019*. *44*(July). https://www.resolutionfoundation.org/app/uploads/2019/07/Living-Standards-Audit-2019.pdf

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

Crawford, C., Duckworth, K., Vignoles, A., & Wyness, G. (2011). *Young people ' s education and labour market choices aged 16/ 17 to 18/19*.

Crawford, C., Macmillan, L., & Vignoles, A. (2017). When and why do initially high-achieving poor children fall behind? *Oxford Review of Education*, *43*(1), 88–108. https://doi.org/10.1080/03054985.2016.1240672

Crenna-Jennings, W. (2018). *Key drivers of the disadvantage gap*. Education Policy Institute.

Croll, P. (1995). Early Linguistic Attainment, Family Background and Performance in 16 + Examinations. *Educational Studies*, *21*(1), 13–28. https://doi.org/10.1080/0305569950210102

Cullinane, C., & Montacute, R. (2020). COVID-19 and social mobility impact brief #1:

School shutdown. *The Sutton Trust*, *April*, 1–11.

Dailey, S., & Bergelson, E. (2021). Language input to infants of different socioeconomic statuses: A quantitative meta-analysis. *Developmental Science*, *November*, 1–13. https://doi.org/10.1111/desc.13192

Daviss, W. B., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(9), 927–934. https://doi.org/10.1111/j.1469-7610.2006.01646.x

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

Deighton, J., Lereya, T., Patalay, P., Casey, P., Humphrey, N., & Wolpert, M. (2018). *Results from the first HeadStart annual survey Executive summary*. *January*.

Deighton, J., Tymms, P., Vostanis, P., Belsky, J., Fonagy, P., Brown, A., Martin, A., Patalay, P., & Wolpert, M. (2013). The Development of a School-Based Measure of Child Mental Health. *Journal of Psychoeducational Assessment*, *31*(3), 247–257. https://doi.org/10.1177/0734282912465570

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

Demie, F., & Lewis, K. (2011). White working class achievement: An ethnographic study of barriers to learning in schools. *Educational Studies*, *37*(3), 245–264. https://doi.org/10.1080/03055698.2010.506341

Demie, F., & Lewis, K. (2014). *Raising the achievement of white working class pupils : barriers and school strategies*.

Department for Business Innovation and Skills. (2012). *Skills for life Survey report* (Issue 81). http://www.nationalnumeracy.org.uk/sites/default/files/the_2011_skills_for_life_survey_-_december_2012.pdf

Department for Children Schools and Families. (2009). *Every Child a Talker: Guidance for Consultants and Early Language Lead Practitioners*.

Department for Education. (2013). Raising Ambitions and Standards for Primary Schools. In *Education of disadvantaged children, Children and young people and Schools*. https://www.gov.uk/government/news/raising-ambitions-and-standards-for-primary-schools

Department for Education. (2014). *The national curriculum in England: Framework document* (Issue December).

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*.

Department for Education. (2017d). *Unlocking Talent, Fulfilling Potential: A plan for improving social mobility through education*.

Department for Education. (2018a). College staff survey 2018: Research report. *Department for Education*, *November*.

Department for Education. (2018b). *Mental health and behaviour in schools*.

Department for Education. (2018c). *Post-16 education: outcomes for disadvantaged students England*. *November*.

Department for Education. (2018d). SFR 01/2018: Revised GCSE and equivalent results in England, 2016 to 2017. *Department for Education*, *January*, 1–53. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/604312/KS4_shadow_measures_FINAL.pdf%0Ahttps://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2016-to-2017

Department for Education. (2019). *Latest headline data for pupils at the end of key stage 4*. *2019*, 1–17. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/863815/2019_KS4_revised_text.pdf

Department for Education. (2021a). *Special educational needs in England*. https://explore-education-statistics.service.gov.uk/find-statistics/special-educational-needs-in-england

Department for Education. (2021b). Understanding Progress in the 2020/21 Academic Year.

Gov.Uk, *October 2021*, 1–40. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/962330/Learning_Loss_Report_1A_-_FINAL.pdf

Department of Education. (2017). Revised A level and other 16-18 results 2017/2018. *Government Report*, *January*.

Diesendruck, G. (2008). Mechanisms of Word Learning. In E. Hoff & M. Shatz (Eds.), *Blackwell Handbook of Language Development* (pp. 257–276). https://doi.org/10.1002/9780470757833.ch13

Dockrell, J. E., Lindsay, G., & Palikara, O. (2011). Explaining the academic achievement at school leaving for pupils with a history of language impairment: Previous academic achievement and literacy skills. *Child Language Teaching and Therapy*, *27*(2), 223–237. https://doi.org/10.1177/0265659011398671

Dodgeon, B., Elliott, J., Hancock, M., Johnson, J., Parsons, S., & Shepherd, P. (2006). *British Cohort Study (1970) 2004-2005 follow up* (Issue August).

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N., & Kurs-Lasky, M. (2000). Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research*, *42*(6), 1432–1443. https://doi.org/10.1044/jslhr.4206.1432

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

Donnelly, M., Lazetic, P., Sandoval-Hernández, A., Kameshwara, K. K., & Whewall, S. (2019). *An Unequal Playing Field: Extra-Curricular Activities, Soft Skills and Social Mobility*.

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*.

Douglas Willms, J., & Somers, M. A. (2001). Family, Classroom, and School Effects on Children's Educational Outcomes in Latin America. *School Effectiveness and School Improvement*, *12*(4), 409–445. https://doi.org/10.1076/sesi.12.4.409.3445

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.

Durham, R. E., Farkas, G., Hammer, C. S., Bruce Tomblin, J., & Catts, H. W. (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, *25*(4), 294–305. https://doi.org/10.1016/j.rssm.2007.03.001

Durkin, K., & Conti-Ramsden, G. (2010). Young people with specific language impairment: A review of social and emotional functioning in adolescence. *Child Language Teaching and Therapy*, *26*(2), 105–121. https://doi.org/10.1177/0265659010368750

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*.

Eadie, P., Bavin, E. L., Bretherton, L., Cook, F., Gold, L., Mensah, F., Wake, M., & Reilly, S. (2021). Predictors in Infancy for Language and Academic Outcomes at 11 Years. *Pediatrics*, *147*(2). https://doi.org/10.1542/peds.2020-1712

Education Endowment Foundation. (2020). *Impact of school closures on the attainment gap: Rapid Evidence Assessment*. *June*, 19.

Education Endowment Foundation. (2021). *Teaching and Learning Toolkit: Oral language interventions*. https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/oral-language-interventions

Education Scotland. (2021). *What is Curriculum for Excellence?* https://education.gov.scot/education-scotland/scottish-education-system/policy-for-scottish-education/policy-drivers/cfe-building-from-the-statement-appendix-incl-btc1-5/what-is-curriculum-for-excellence

Einarsdóttir, J. T., Björnsdóttir, A., & Símonardóttir, I. (2016). The predictive value of preschool language assessments on academic achievement: A 10-year longitudinal study of icelandic children. *American Journal of Speech-Language Pathology*, *25*(1), 67–79. https://doi.org/10.1044/2015_AJSLP-14-0184

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, C. D., Smith, P., & McCulloch, K. (1997). *British Ability Scales–Second Edition (BAS II): Technical manual.* 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

Ellwood-Lowe, M. E., Foushee, R., & Srinivasan, M. (2022). What causes the word gap? Financial concerns may systematically suppress child-directed speech. *Developmental Science*, *25*(1). https://doi.org/10.1111/desc.13151

Enders, C. K. (2010). *Applied Missing Data Analysis*. The Guilford Press.

Erikson, R., & Goldthorpe, J. H. (1991). *The constant flux: A study of class mobility in industrial societies.* Clarendon Press.

Eshel, N., Daelmans, B., Cabral De Mello, M., & Martines, J. (2006). Responsive parenting: Interventions and outcomes. *Bulletin of the World Health Organization*, *84*(12), 992–998. https://doi.org/10.2471/BLT.06.030163

ESRC. (2012). Education vital for social mobility. *Economic and Social Research Council*. https://esrc.ukri.org/news-events-and-publications/evidence-briefings/education-vital-for-social-mobility/

Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology*, *35*(4), 1020–1023. https://doi.org/10.1037/0012-1649.35.4.1020

Fabozzi, F. J., Focardi, S. M., Rachev, S. T., & Arshanapalli, B. G. (2014). Appendix E: Model Selection Criterion: AIC and BIC. *The Basics of Financial Econometrics*, *41*(1979), 399–403. https://doi.org/10.1002/9781118856406.app5

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, *33*(3), 464–497. https://doi.org/10.1016/j.ssresearch.2003.08.001

Fazackerley, A. (2021). Fears half of poorer pupils in England could be barred from university. *The Guardian*.

Feinstein, L. (2003). Inequality in the early cognitive development of British children in the

1970 cohort. *Economica*, *70*(277), 73–97. https://doi.org/10.1111/1468-0335.t01-1-00272

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).

Fergusson, D. M., Horwood, L. J., Ridder, E. M., & Beautrais, A. L. (2005). Subthreshold Depression in Adolescence and Mental Health Outcomes in Adulthood. *Arch Gen Psychiatry*, *62*(1), 66–72. https://doi.org/10.1001/archpsyc.62.1.66

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

Field, F. (2010). *The Foundation Years: Preventing Poor Children be- coming Poor Adults, The Report of the Independent Review on Poverty and Life Chances*. The Stationary Office.

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*.

Fitzsimons, E., Haselden, L., Smith, K., Gilbert, E., Calderwood, L., Agalioti- Sgompou, V., Veeravalli, S., Silverwood, R., & Ploubidis, G. (2020). *Millennium Cohort Study: Age 17 Sweep (MCS7) User guide* (2nd editio, Issue December). UCL Centre for Longitudinal Studies. https://cls.ucl.ac.uk/wp-content/uploads/2020/09/MCS1-5_User_Guide_ed9_2020-08-07.pdf

Fitzsimons, E., Moulton, V., Hughes, D. A., Neaves, S., Ho, K., Hemani, G., Timpson, N., Calderwood, L., Gilbert, E., & Ring, S. (2020). *Collection of DNA Samples and Genetic Data at Scale in the UK Millennium Cohort Study. CLS working paper number 2020/7*. UCL Centre for Longitudinal Studies.

Forrest, C. L., Gibson, J. L., Halligan, S. L., & St Clair, M. C. (2020). A Cross-Lagged Analysis of Emotion Regulation, Peer Problems, and Emotional Problems in Children With and Without Early Language Difficulties: Evidence From the Millennium Cohort Study. *Journal of Speech, Language, and Hearing Research*, *63*(4), 1227–1239. https://doi.org/https://doi.org/10.1044/2020_JSLHR-19-00188

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

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–

694. https://doi.org/10.1017/S0305000916000209

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

Galindo-Rueda, F., Marcenaro-Gutierrez, O., & Vignoles, A. (2004). The Widening Socio-economic Gap in UK Higher Education. *National Institute Economic Review*, *190*, 70–82. https://doi.org/https://doi.org/10.1177/002795010419000108

Gallop, K., Anderson, F., & Bram, A. (2013). *Millennium Cohort Study Fifth Sweep (MCS5) Technical Report*. *May*.

Galobardes, B., Lynch, J., & Smith, G. D. (2007). Measuring socioeconomic position in health research. *British Medical Bulletin*, *81–82*(1), 21–37. https://doi.org/10.1093/bmb/ldm001

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Smith, G. D. (2006a). Indicators of socioeconomic position (part 1). *Journal of Epidemiology and Community Health*, *60*(1), 7–12. https://doi.org/10.1136/jech.2004.023531

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Smith, G. D. (2006b). Indicators of socioeconomic position (part 2). *Journal of Epidemiology and Community Health*, *60*(2), 95–101. https://doi.org/10.1136/jech.2004.028092

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

Gatt, D., Baldacchino, R., & Dodd, B. (2020). Which measure of socioeconomic status best predicts bilingual lexical abilities and how? A focus on four-year-olds exposed to two majority languages. *Journal of Child Language*, *47*(4), 737–765. https://doi.org/10.1017/S0305000919000886

Geoffroy, M. C., Côté, S. M., Borge, A. I. H., Larouche, F., Séguin, J. R., & Rutter, M. (2007). Association between nonmaternal care in the first year of life and children's receptive language skills prior to school entry: The moderating role of socioeconomic status. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *48*(5), 490–497. https://doi.org/10.1111/j.1469-7610.2006.01704.x

Georgiou, N., & Spanoudis, G. (2021). Developmental language disorder and autism:

Commonalities and differences on language. *Brain Sciences*, *11*(5). https://doi.org/10.3390/brainsci11050589

Gibb, N. (2015). *The purpose of education*. Education Reform Summit. https://www.gov.uk/government/speeches/the-purpose-of-education

Gibbons, S. (2012). Valuing schooling through house prices. *CentrePiece*, *17*(2), 4.

Goldthorpe, J. H., Llewellyn, C., & Payne, G. (1980). *Class structure and social mobility in modern Britain.* Clarendon Press.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language Matters: Denying the Existence of the 30-Million-Word Gap Has Serious Consequences. *Child Development*, *90*(3), 985–992. https://doi.org/10.1111/cdev.13128

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

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Gottschling, J., Hahn, E., Beam, C. R., Spinath, F. M., Carroll, S., & Turkheimer, E. (2019). Socioeconomic status amplifies genetic effects in middle childhood in a large German twin sample. *Intelligence*, *72*(November 2018), 20–27. https://doi.org/10.1016/j.intell.2018.11.006

Gough Kenyon, S. M., Palikara, O., & Lucas, R. M. (2021). Consistency of Parental and Self-Reported Adolescent Wellbeing: Evidence From Developmental Language Disorder. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.629577

Gov.Uk. (2021). *16 to 19 funding: maths and English condition of funding*. https://www.gov.uk/guidance/16-to-19-funding-maths-and-english-condition-of-funding#summary

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

Graham, S., Harris, K. R., & Santangelo, T. (2015). Research-Based Writing Practices and the Common Core. *The Elementary School Journal*, *115*(4), 498–522. https://doi.org/10.1086/681964

Gray, J., Galton, M., McLaughlin, C., Clarke, B., & Symonds, J. (2011). *The supportive school*. Cambridge Scholars Publishing.

Green, F., Anders, J., Henderson, M., & Henseke, G. (2020). Private benefits? External

benefits? Outcomes of private schooling in 21stcentury britain. *Journal of Social Policy*, *49*(4), 724–743. https://doi.org/10.1017/S0047279419000710

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. (2012). Dimensions and Boundaries: Comparative Analysis of Occupational Structures Using Social Network and Social Interaction Distance Analaysis. *Sociological Research Online*, *17*(2).

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

Gustafsson, J.-E., Westling, A. M., Alin Åkerman, B., Eriksson, C., Eriksson, L., Fischbein, S., Granlund, M., Gustafsson, P., Ljungdahl, S., Ogden, T., & Persson, R. S. (2010). *School, learning and mental health: A systematic review*.

Hallam, S., & Parsons, S. (2013). Prevalence of streaming in UK primary schools: Evidence from the Millennium Cohort Study. *British Educational Research Journal*, *39*(3), 514–544. https://doi.org/10.1080/01411926.2012.659721

Hansen, K. (2012). *Millennium Cohort Study First, Second, Third and Fourth Surveys A Guide to the Datasets (Seventh Edition)* (Issue February).

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

Harel, O. (2009). The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation. *Journal of Applied Statistics*, *36*(10), 1109–1118. https://doi.org/10.1080/02664760802553000

Hart, B., & Risley, T. R. (1995a). 42 American Families. In *Meaningful Differences in the Everyday Experience of Young American Children* (pp. 53–74).

Hart, B., & Risley, T. R. (1995b). *The Early Catastrophe: The 30 Million Word Gap by Age 3.* Brookes.

Harwell, M., Maeda, Y., Bishop, K., & Xie, A. (2017). The Surprisingly Modest Relationship Between SES and Educational Achievement. *Journal of Experimental Education*, *85*(2),

197–214. https://doi.org/10.1080/00220973.2015.1123668

Hayiou-Thomas, M. E., Dale, P. S., & Plomin, R. (2012). The etiology of variation in language skills changes with development: A longitudinal twin study of language from 2 to 12years. *Developmental Science*, *15*(2), 233–249. https://doi.org/10.1111/j.1467-7687.2011.01119.x

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

Heckman, J. J. (2011). The economics of inequality: The value of early childhood education. *American Educator*, 31–36. http://files.eric.ed.gov/fulltext/EJ920516.pdf

Heckman, J. J. (2012). The Heckman Equation. *Www.Heckmanequation.Org*, 1–2. www.heckmanequation.org

Heckman, J. J. (2015). *Four Big Benefits of Investing in Early Childhood Development*.

Hegedus, A. (2018). *Evaluating the relationships between poverty and school performance*. *October*, 1–15.

Henderson, M., Anders, J., Green, F., & Henseke, G. (2020). Private schooling, subject choice, upper secondary attainment and progression to university. *Oxford Review of Education*, *46*(3), 295–312. https://doi.org/10.1080/03054985.2019.1669551

Henderson, M., Sullivan, A., Anders, J., & Moulton, V. (2018). Social Class, Gender and Ethnic Differences in Subjects Taken at Age 14. *Curriculum Journal*, *29*(3), 298–318. https://doi.org/10.1080/09585176.2017.1406810

Henry, L. (2015). The effects of ability grouping on the learning of children from low income homes: a systematic review. *The STeP Journal*, *2*(3), 70–87.

Hentges, R. F., Devereux, C., Graham, S. A., & Madigan, S. (2021). Child Language Difficulties and Internalizing and Externalizing Symptoms: A Meta-Analysis. *Child Development*, *92*(4), 1–25. https://doi.org/10.1111/cdev.13540

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-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, *19*(4), 603–629. https://doi.org/10.1017/s0142716400010389

Hoff, E. (2003). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*, *74*(5), 1368–1378. https://doi.org/10.1111/1467-8624.00612

Hoff, E. (2013). Interpreting the Early Language Trajectories of Children from Low SES and Language Minority Homes: Implications for Closing Achievement Gaps. *Developmental Psychology*, *49*(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

Hollingshead, A. B. (1975). *Four factor index of social status.* Unpublished manuscript, Yale University, New Haven, CT.

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

Horvat, E. M. N., Weininger, E. B., & Lareau, A. (2003). From Social Ties to Social Capital: Class Differences in the Relations between Schools and Parent Networks. *American Educational Research Journal*, *40*(2), 319–351. https://doi.org/10.3102/00028312040002319

House of Commons. (2021). *The forgotten: how White working-class pupils have been let down, and how to change it* (Issue June).

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

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. Cognitive psychology, 61(4), 343-365. *Cognitive Psychology*, *61*(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Independent Schools Council. (2018). *ISC Census and Annual Report 2016*. http://www.isc.co.uk/media/3182/isc_census_2016_final.pdf

Institute for Fiscal Studies. (2017). Living standards, poverty and inequality in the UK: 2020. In *Institute For Fiscal Studies*. https://www.ifs.org.uk/publications/14901

Institute of Child Health. (1970). *The 1970 birth cohort*.

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

completion questionnaire.

Ioannidis, J. P. A. (2018). Why most published research findings are false. *Getting to Good: Research Integrity in the Biomedical Sciences*, *2*(8), 2–8. https://doi.org/10.1371/journal.pmed.0020124

Iqbal, S. A., Wallach, J. D., Khoury, M. J., Schully, S. D., & Ioannidis, J. P. A. (2016). Reproducible Research Practices and Transparency across the Biomedical Literature. *PLOS Biology*, *14*(1), 1–13. https://doi.org/10.1371/journal.pbio.1002333

Jenkins, R. (2006). *Pierre Bourdieu*. Taylor & Francis.

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

Jerrim, J. (2021). National tests and the wellbeing of primary school pupils: new evidence from the UK. *Assessment in Education: Principles, Policy and Practice*, *28*(5–6), 507–544. https://doi.org/10.1080/0969594X.2021.1929829

Jerrim, J., & De Vries, R. (2017). The limitations of quantitative social science for informing public policy. *Evidence and Policy*, *13*(1), 117–133. https://doi.org/10.1332/174426415X14431000856662

Jerrim, J., Greany, T., & Perera, N. (2016). *Educational disadvantage: how does England compare?*

Jerrim, J., & Shure, N. (2017). *Achievement of 15-year-olds in England : PISA 2015 national report*. *December*, 246. https://www.gov.uk/government/publications/pisa-2015-national-report-for-england

Jerrim, J., & Vignoles, A. (2013). Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *176*(4), 887–906. https://doi.org/10.1111/j.1467-985X.2012.01072.x

Jez, S. J. (2014). The Differential Impact of Wealth Versus Income in the College-Going Process. *Research in Higher Education*, *55*(7), 710–734. https://doi.org/10.1007/s11162-014-9332-0

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-year follow-up of children with and without speech-language impairments: Family, educational, occupational, and quality of life outcomes. *American Journal of Speech-Language Pathology*, *19*(1), 51–65. https://doi.org/10.1044/1058-0360(2009/08-0083)

Johnston, D. W., Propper, C., Pudney, S. E., & Shields, M. A. (2014). The income gradient in childhood mental health: All in the eye of the beholder? *Journal of the Royal Statistical*

Society. Series A: Statistics in Society, *177*(4), 807–827. https://doi.org/10.1111/rssa.12038

Jordan, J. A., & Coulter, L. (2017). Individual Differences in Speech and Language Ability Profiles in Areas of High Deprivation. *Child Care in Practice*, *23*(4), 417–430. https://doi.org/10.1080/13575279.2016.1188759

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

Joshi, H., & Fitzsimons, E. (2016). The UK millennium cohort study: The making of a multipurpose resource for social science and policy. *Longitudinal and Life Course Studies*, *7*(4), 409–430. https://doi.org/10.14301/llcs.v7i4.410

Kalmijn, M. (1991a). Shifting Boundaries: Trends in Religious and Educational Homogamy. *American Sociological Review*, *56*(6), 786–800. https://doi.org/https://doi.org/10.2307/2096256

Kalmijn, M. (1991b). Status Homogamy in the United States. *American Journal of Sociology*, *97*(2), 496–523. https://doi.org/10.1086/229786

Karagiannaki, E. (2017). The effect of parental wealth on children's outcomes in early adulthood. *Journal of Economic Inequality*, *15*(3), 217–243. https://doi.org/10.1007/s10888-017-9350-1

Kartushina, N., Mani, N., AKTAN-ERCIYES, A., Alaslani, K., Aldrich, N. J., Almohammadi, A., Alroqi, H., Anderson, L., Andonova, E., Aussems, S., Babineau, M., Barokova, M., Bergmann, C., Cashon, C. H., Custode, S., Carvalho, A. de, Dimitrova, N., Dynak, A., Farah, R., … Mayor, J. (2022). COVID-19 first lockdown as a window into language acquisition: associations between caregiver-child activities and vocabulary gains. *Language Development Research*, *2*(1), 1–36. https://doi.org/https://doi.org/10.34842/abym-xv34

Kaushal, N., Magnuson, K., & Waldfogel, J. (2011). How is Family Income related to Investments in Children ' s Learning ? In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 187–206). New York, NY:Russell Sage Foundation.

Keep, E., & Mayhew, K. (2004). The economic and distributional implications of current policies on higher education. *Oxford Review of Economic Policy*, *20*(2), 298–314. https://doi.org/10.1093/oxrep/grh017

Keister, L. (2018). *Income and wealth are not highly correlated : here is why and what it means*. Work In Progress. http://www.wipsociology.org/2018/10/29/income-and-wealth-are-not-highly-correlated-here-is-why-and-what-it-means/

Keller, M. C. (2014). Gene × environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, *75*(1), 18–24. https://doi.org/10.1016/j.biopsych.2013.09.006

Kelly, B. (2013). The process of socio-economic constraint on geographical mobility: England 1991 to 2008. *CCSR Working Paper*, *2013–08*.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*(2), 184–189. https://doi.org/10.1001/archpsyc.60.2.184

Ketelaars, M. P., Cuperus, J., Jansonius, K., & Verhoeven, L. (2010). Pragmatic language impairment and associated behavioural problems. *International Journal of Language and Communication Disorders*, *45*(2), 204–214. https://doi.org/10.3109/13682820902863090

Kidd, E., & Donnelly, S. (2020). Individual Differences in First Language Acquisition. *Annual Review of Linguistics*, *6*(1), 319–340. https://doi.org/10.1146/annurev-linguistics-011619-030326

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

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

Kim, S. won, Cho, H., & Kim, L. Y. (2019). Socioeconomic Status and Academic Outcomes in Developing Countries: A Meta-Analysis. *Review of Educational Research*, *89*(6), 875–916. https://doi.org/10.3102/0034654319877155

Kirby, P. (2016). Shadow schooling: private tuition and social mobility in the UK. *The Sutton Trust*, *September*.

Kornrich, S. (2016). Inequalities in Parental Spending on Young Children: 1972 to 2010. *AERA Open*, *2*(2), 1–12. https://doi.org/10.1177/2332858416644180

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

Koshy, V., Brown, J., Jones, D., & Portman Smith, C. (2013). Exploring the views of parents of high ability children living in relative poverty. *Educational Research*, *55*(3), 304–

320. https://doi.org/10.1080/00131881.2013.825167

Kupersmidt, J. B., DeRosier, M. E., & Patterson, C. P. (1995). Similarity as the basis for children's friendships: the roles of sociometric status, aggressive and withdrawn behaviour, academic achievement and demographic characteristics. *Journal of Social and Personal Relationships*, *12*(3), 439–452. https://doi.org/https://doi.org/10.1177/0265407595123007

Kuppens, T., Spears, R., Manstead, A. S. R., Spruyt, B., & Easterbrook, M. J. (2018). Educationism and the irony of meritocracy: Negative attitudes of higher educated people towards the less educated. *Journal of Experimental Social Psychology*, *76*(August 2016), 429–447. https://doi.org/10.1016/j.jesp.2017.11.001

Kutcher, S., Wei, Y., & Coniglio, C. (2016). Mental health literacy: Past, present, and future. *Canadian Journal of Psychiatry*, *61*(3), 154–158. https://doi.org/10.1177/0706743715616609

Labov, W. (1969a). *A Study of Non-Standard English*. 555–562.

Labov, W. (1969b). *The Logic of Non-Standard English*.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. https://doi.org/10.1037//0033-295x.104.2.211

Landry, S. H., Zucker, T. A., Montroy, J. J., Hsu, H. Y., Assel, M. A., Varghese, C., Crawford, A., & Feil, E. G. (2021). Replication of combined school readiness interventions for teachers and parents of head start pre-kindergarteners using remote delivery. *Early Childhood Research Quarterly*, *56*, 149–166. https://doi.org/10.1016/j.ecresq.2021.03.007

Landry, S. H., Zucker, T. A., Taylor, H. B., Swank, P. R., Williams, J. M., Assel, M., Crawford, A., Huang, W., Clancy-Menchetti, J., Lonigan, C. J., Phillips, B. M., Eisenberg, N., Spinrad, T. L., De Villiers, J., De Villiers, P., Barnes, M., Starkey, P., & Klein, A. (2014). Enhancing early child care quality and learning for toddlers at risk: The responsive early childhood program. *Developmental Psychology*, *50*(2), 526–541. https://doi.org/10.1037/a0033494

Landry, S. H., Zucker, T. A., Williams, J. M., Merz, E. C., Guttentag, C. L., & Taylor, H. B. (2017). Improving school readiness of high-risk preschoolers: Combining high quality instructional strategies with responsive training for teachers and parents. *Early Childhood Research Quarterly*, *40*, 38–51. https://doi.org/10.1016/j.ecresq.2016.12.001

Language and Reading Research Consortium. (2015). The Dimensionality of Language

Ability in Young Children. *Child Development*, *86*(6), 1948–1965. https://doi.org/10.1111/cdev.12450

Larsen, J. A., & Nippold, M. A. (2007). Morphological Analysis in School-Age Children: Dynamic Assessment of a Word Learning Strategy. *Language, Speech, and Hearing Services in Schools*, *38*(3), 201–212. https://doi.org/10.1044/0161-1461(2007/021)

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

Law, J., McBean, K., & Rush, R. (2011). Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage. *International Journal of Language and Communication Disorders*, *46*(6), 657–664. https://doi.org/10.1111/j.1460-6984.2011.00036.x

Law, J., Reilly, S., & Snow, P. C. (2013). Child speech, language and communication need re-examined in a public health context: A new direction for the speech and language therapy profession. *International Journal of Language and Communication Disorders*, *48*(5), 486–496. https://doi.org/10.1111/1460-6984.12027

Law, J., Rush, R., King, T., Westrupp, E., & Reilly, S. (2018). Early Home Activities and Oral Language Skills in Middle Childhood: A Quantile Analysis. *Child Development*, *89*(1), 295–309. https://doi.org/10.1111/cdev.12727

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes James. *Journal of Speech, Language, and Hearing Research*, *52*, 1401–1416. https://doi.org/1092-4388/09/5206-1401

Law, J., Todd, L., Clark, J., Mroz, M., & Carr, J. (2013). Early language delays in the UK. In *Save The Children*.

http://www.mcri.edu.au/media/214393/law_et_al_2013_early_language_delays_in_the_uk.pdf

Lehrl, S., Ebert, S., Blaurock, S., Rossbach, H. G., & Weinert, S. (2019). Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school. *School Effectiveness and School Improvement*, *3453*. https://doi.org/10.1080/09243453.2019.1618346

Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the Developmental Relationship Between Oral Language Skills and Reading Comprehension: It's Simple, But Complex. *Child Development*, *89*(5), 1821–1838. https://doi.org/10.1111/cdev.12861

Lessard, L. M., & Juvonen, J. (2019). Cross-class friendship and academic achievement in middle school. *Developmental Psychology*, *55*(8), 1666–1679. https://doi.org/10.1037/dev0000755

Levickis, P., Reilly, S., Girolametto, L., Ukoumunne, O. C., & Wake, M. (2018). Associations between maternal responsive linguistic input and child language performance at age 4 in a community-based sample of slow-to-talk toddlers. *Child: Care, Health and Development*, *44*(5), 776–783. https://doi.org/10.1111/cch.12600

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

Little, R., & Rubin, D. (2019). *Statistical analysis with missing data*. Wiley-Blackwell.

Lonigan, C. J., & Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of children with typical development in preschool through fifth grade. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2185–2198. https://doi.org/10.1044/2017_JSLHR-L-15-0402

Lupton, R., Thomson, S., Velthuis, S., & Unwin, L. (2021). *Moving on from initial GCSE "failure": Post-16 transitions for "lower attainers" and why the English education system must do better* (Vol. 4, Issue February, pp. 1–140). Nuffield Foundation. https://www.research.manchester.ac.uk/portal/files/187105835/FINAL_main_report_for_publishing.pdf

Lusardi, A., Schneider, D., & Tufano, P. (2011). Financially fragile households: Evidence and implications. *Brookings Papers on Economic Activity*, *1*, 83–134. https://doi.org/10.1353/eca.2011.0002

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

Machin, S., McNally, S., & Ruiz-Valenzuela, J. (2020). Entry through the narrow door: The costs of just failing high stakes exams. *Journal of Public Economics*, *190*(April). https://doi.org/10.1016/j.jpubeco.2020.104224

Machin, S., & Vignoles, A. (2004). Educational inequality: The widening socio-economic gap. *Fiscal Studies*, *25*(2), 107–128. https://doi.org/10.1111/j.1475-5890.2004.tb00099.x

Magnuson, K. A., Meyers, M. K., Ruhm, C. J., & Waldfogel, J. (2004). Inequality in preschool education and school readiness. *American Educational Research Journal*, *41*(1), 115–157. https://doi.org/10.3102/00028312041001115

Maguire, M. J., Schneider, J. M., Middleton, A. E., Ralph, Y., Lopez, M., Ackerman, R. A., & Abel, A. D. (2018). Vocabulary knowledge mediates the link between socioeconomic status and word learning in grade school. *Journal of Experimental Child Psychology*, *166*, 679–695. https://doi.org/10.1016/j.jecp.2017.10.003

Malacarne, T. (2017). Rich Friends, Poor Friends: Inter–Socioeconomic Status Friendships in Secondary School. *Socius: Sociological Research for a Dynamic World*, *3*, 237802311773699. https://doi.org/10.1177/2378023117736994

Malin, J. L., Cabrera, N. J., & Rowe, M. L. (2014). Low-income minority mothers' and fathers' reading and children's interest: Longitudinal contributions to children's receptive vocabulary skills. *Early Childhood Research Quarterly*, *29*(4), 425–432. https://doi.org/10.1016/j.ecresq.2014.04.010

Marmot, M. (2010). *Fair society, healthy lives: The Marmot Review*.

Marulis, L. M., & Neuman, S. B. (2013). How Vocabulary Interventions Affect Young Children at Risk: A Meta-Analytic Review. *Journal of Research on Educational Effectiveness*, *6*(3), 223–262. https://doi.org/10.1080/19345747.2012.755591

Mayhew, K., Deer, C., & Dua, M. (2004). The move to mass higher education in the UK: Many questions and some answers. *Oxford Review of Education*, *30*(1), 65–82. https://doi.org/10.1080/0305498042000190069

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

McKean, C., Reilly, S., Bavin, E. L., Bretherton, L., Cini, E., Conway, L., Cook, F., Eadie, P., Prior, M., Wake, M., & Mensah, F. (2017). Language outcomes at 7 years: Early

predictors and co-occurring difficulties. *Pediatrics*, *139*(3). https://doi.org/10.1542/peds.2016-1684

McKean, C., Wraith, D., Eadie, P., Cook, F., Mensah, F., & Reilly, S. (2017). Subgroups in language trajectories from 4 to 11 years: the nature and predictors of stable, improving and decreasing language trajectory groups. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1081–1091. https://doi.org/10.1111/jcpp.12790

McLaughlin, K. A., & King, K. (2016). Developmental Trajectories of Anxiety and Depression in Early Adolescence. *J Abnorm Child Psychol*, *43*(2), 311–323. https://doi.org/10.1007/s10802-014-9898-1

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

Mental Health Foundation. (2022). *Mental Health Awareness Week*. https://www.mentalhealth.org.uk/campaigns/mental-health-awareness-week

Miller, A. B., Machlin, L., McLaughlin, K. A., & Sheridan, M. A. (2020). Deprivation and psychopathology in the Fragile Families Study: A 15-year longitudinal investigation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.13260

Mol, S. E., & Bus, A. G. (2011). To Read or Not to Read: A Meta-Analysis of Print Exposure From Infancy to Early Adulthood. *Psychological Bulletin*, *137*(2), 267–296. https://doi.org/10.1037/a0021890

Moll, K., Snowling, M. J., Göbel, S. M., & Hulme, C. (2015). Early language and executive skills predict variations in number and arithmetic skills in children at family-risk of dyslexia and typically developing controls. *Learning and Instruction*, *38*, 53–62. https://doi.org/10.1016/j.learninstruc.2015.03.004

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and Diversity: Simulating Early

Word Learning Environments. *Cognitive Science*, *42*, 375–412. https://doi.org/10.1111/cogs.12592

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-Month-Old Children With Larger Oral Vocabularies Display Greater Academic and Behavioral Functioning at Kindergarten Entry. *Child Development*, *86*(5), 1351–1370. https://doi.org/10.1111/cdev.12398

Mostafa, T., & Ploubidis, G. (2017). *Millennium Cohort Study Sixth Survey 2015-2016: Technical report on response*. *February*. https://doc.ukdataservice.ac.uk/doc/8156/mrdoc/pdf/mcs6_report_on_response.pdf

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

Mueller, K. L., & Tomblin, J. B. (2012). Examining the comorbidity of language impairment and attention-deficit/ hyperactivity disorder. *Topics in Language Disorders*, *32*(3), 228–246. https://doi.org/10.1097/TLD.0b013e318262010d

Muijs, D., & Dunne, M. (2010). Setting by ability - or is it? A quantitative study of determinants of set placement in english secondary schools. *Educational Research*, *52*(4), 391–407. https://doi.org/10.1080/00131881.2010.524750

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In *The nature of vocabulary acquisition.* (pp. 19–35). Lawrence Erlbaum Associates, Inc.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Neuman, S. B., & Celano, D. (2001). Access to Print in Low-Income and Middle-Income Communities: An Ecological Study of Four Neighborhoods. *Reading Research Quarterly*, *36*(1), 8–26. https://doi.org/10.1598/rrq.36.1.1

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

Noble, C., Cameron-Faulkner, T., Jessop, A., Coates, A., Sawyer, H., Taylor-Ims, R., & Rowland, C. F. (2020). The impact of interactive shared book reading on children's language skills: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*, *63*(6), 1878–1897. https://doi.org/10.1044/2020_JSLHR-19-00288

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Noble, M., Wright, G., Dibben, C., Smith, G., McLennan, D., Anttila, C., Barnes, H., Mokhtar, C., Noble, S., Avenell, D., Gardner, J., Covizzi, I., & Lloyd, M. (2004). *The English indices of deprivation 2004 (revised)*. *2004*.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Norbury, C. F., & Sparks, A. (2013). Difference or disorder? Cultural issues in understanding neurodevelopmental disorders. *Developmental Psychology*, *49*(1), 45–58. https://doi.org/10.1037/a0027446

Norbury, C. F., Vamvakas, G., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A. (2017). Language growth in children with heterogeneous language disorders: a population study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1092–1105. https://doi.org/10.1111/jcpp.12793

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

OECD. (2006). *Encouraging Quality in Early Childhood Education and Care (ECEC): Research Brief: Qualifications, Education and Professional Develoment Matter*. 1–8.

OECD. (2015). *In It Together: Why Less Inequality Benefits All*. OECD Publishing, Paris. https://doi.org/10.1787/9789264235120-en

OECD. (2016a). Building Skills for All: A Review of England. *Policy Insights from the Survey of Adult Skills*, 1–110. https://www.oecd.org/unitedkingdom/building-skills-for-all-review-of-england.pdf

OECD. (2016b). Skills Matter. In *OECD Skills Studies*. OECD Publishing. https://doi.org/10.1787/9789264258051-en

Office for National Statistics. (2015). Chapter 3: Equivalised income. In *Compendium: Family Spending: 2015. A report on the Living Costs and Food Survey 2014*.

Office for National Statistics. (2019a). Exploring the UK's digital divide - Office for National Statistics. *Office for National Statistics*, 1–24.

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

Office of Population Censuses and Surveys. (1970). *Classification of Occupations*. HMSO.

Ofqual. (2013). *Reforms to GCSEs in England from 2015*. *November 2013*.

Ofqual. (2017). *Statement from the qualification regulators on changes to GCSEs , AS and A levels*. *August 2017*, 1–21.

Ofqual. (2018). *Grading new GCSEs*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/719124/Grading_new_GCSEs25.6.2018.pdf

AS and A levels in England , Wales and Northern Ireland GCSEs in England , Wales and Northern Ireland, (2017).

Ofsted. (2013). *Unseen children: access and achievement 20 years on Evidence report*.

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

Orr, A. J. (2003). Black-White Differences in Achievement: The Importance of Wealth. *Sociology of Education*, *76*(4), 281–304. https://doi.org/10.2307/1519867

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.

Owen, R., Hayett, L., & Roulstone, S. (2004). Children's views of speech and language therapy in school: Consulting children with communication difficulties. *Child Language Teaching and Therapy*, *20*(1), 55–73. https://doi.org/10.1191/0265659004ct263oa

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

Oxford University Press. (2020). *Bridging the Word Gap at Transition: The Oxford Language Report 2020*.

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

https://doi.org/10.1016/j.ecresq.2018.04.001

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

Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: A multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement*, *19*(1), 21–49. https://doi.org/10.1080/09243450801936845

Palikara, O., Lindsay, G., & Dockrell, J. E. (2009). Voices of young people with a history of specific language impairment (SLI) in the first year of post-16 education. *International Journal of Language and Communication Disorders*, *44*(1), 56–78. https://doi.org/10.1080/13682820801949032

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*(4), 763–782. https://doi.org/10.1111/j.1467-8624.2005.00876.x

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

Patalay, P., Deighton, J., Fonagy, P., Vostanis, P., & Wolpert, M. (2014). Clinical validity of the Me and My School questionnaire: A self-report mental health measure for children and adolescents. *Child and Adolescent Psychiatry and Mental Health*, *8*(1), 1–7. https://doi.org/10.1186/1753-2000-8-17

Patalay, P., & Fitzsimons, E. (2018). Development and predictors of mental ill-health and wellbeing from childhood to adolescence. *Social Psychiatry and Psychiatric Epidemiology*, *53*(12), 1311–1323. https://doi.org/10.1007/s00127-018-1604-0

Patalay, P., & Gage, S. H. (2019). Changes in millennial adolescent mental health and health-related behaviours over 10 years: a population cohort comparison study. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyz006

Paull, G., & La Valle, I. (2018). *Evaluation of the first year of the national rollout of 30 hours free childcare.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/740168/Evaluation_of_national_rollout_of_30_hours_free-childcare.pdf

Pearson Qualifications. (2022). *Understanding marks and grades*.

Pensiero, N., & Schoon, I. (2019). Social inequalities in educational attainment: The changing impact of parents' social class, social status, education and family income,

England 1986 and 2010. *Longitudinal and Life Course Studies*, *10*(1), 87–108. https://doi.org/10.1332/175795919X15468755933380

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.

Petitclerc, A., Côté, S., Doyle, O., Burchinal, M., Herba, C., Zachrisson, H. D., Boivin, M., Tremblay, R. E., Tiemeier, H., Jaddoe, V., & Raat, H. (2017). Who uses early childhood education and care services? Comparing socioeconomic selection across five western policy contexts. *International Journal of Child Care and Education Policy*, *11*(1). https://doi.org/10.1186/s40723-017-0028-8

Pew Research Center. (2020). *Most Americans say there is too much economic inequality in the U.S., but fewer than half call it a top priority*.

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

Pfeffer, F. T., & Killewald, A. (2018). Generations of advantage. Multigenerational correlations in family wealth. *Social Forces*, *96*(4), 1411–1442. https://doi.org/10.1093/sf/sox086

Pfeffer, F. T., & Schoeni, R. F. (2016). How wealth inequality shapes our future. *Rsf*, *2*(6), 2–22. https://doi.org/10.7758/rsf.2016.2.6.01

Plewis, I. (2007a). Non-response in a birth cohort study: The case of the millennium cohort study. *International Journal of Social Research Methodology*, *10*(5), 325–334. https://doi.org/10.1080/13645570701676955

Plewis, I. (2007b). *The Millennium Cohort Study. Technical report on Sampling (4th Edition)*. Centre for Longitudinal Studies.

Power, C., Atherton, K., Strachan, D. P., Shepherd, P., Fuller, E., Davis, A., Gibb, I., Kumari, M., Lowe, G., Macfarlane, G. J., Rahi, J., Rodgers, B., & Stansfeld, S. (2007). Life-course influences on health in British adults: Effects of socio-economic position in childhood and adulthood. *International Journal of Epidemiology*, *36*(3), 532–539. https://doi.org/10.1093/ije/dyl310

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

Public Health England. (2020b). *Best start in speech , language and communication : Supporting evidence*.

Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing*

Research, *41*(3), 688–700. https://doi.org/10.1044/jslhr.4103.688

Redmond, S. M., & Rice, M. L. (2002). Stability of behavioral ratings of children with SLI. *Journal of Speech, Language, and Hearing Research*, *45*(1), 190–201. https://doi.org/10.1044/1092-4388(2002/014)

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P., & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from early language in Victoria study. *Pediatrics*, *126*(6). https://doi.org/10.1542/peds.2010-0254

Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science and Medicine*, *90*, 24–31. https://doi.org/10.1016/j.socscimed.2013.04.026

Rescorla, L. A., Ginzburg, S., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., Bilenberg, N., Bird, H., Chahed, M., Dobrean, A., Döpfner, M., Erol, N., Hannesdottir, H., Kanbayashi, Y., Lambert, M. C., Leung, P. W. L., Minaei, A., Novik, T. S., Oh, K. J., … Verhulst, F. C. (2013). Cross-Informant Agreement Between Parent-Reported and Adolescent Self-Reported Problems in 25 Societies. *Journal of Clinical Child and Adolescent Psychology*, *42*(2), 262–273. https://doi.org/10.1080/15374416.2012.717870

Reznick, J. S., & Goldfield, B. A. (1992). Rapid Change in Lexical Development in Comprehension and Production. *Developmental Psychology*, *28*(3), 406–413. https://doi.org/10.1037/0012-1649.28.3.406

Ricketts, J., Bishop, D. V. M., & Nation, K. (2008). Investigating orthographic and semantic aspects of word learning in poor comprehenders. *Journal of Research in Reading*, *31*(1), 117–135. https://doi.org/10.1111/j.1467-9817.2007.00365.x

Ricketts, J., Lervåg, A., Dawson, N., Taylor, L. A., & Hulme, C. (2020). Reading and Oral Vocabulary Development in Early Adolescence. *Scientific Studies of Reading*, *24*(5), 380–396. https://doi.org/10.1080/10888438.2019.1689244

Ricketts, J., Sperring, R., Nation, K., Tighe, E. L., & State, F. (2014). *Educational attainment in poor comprehenders*. *5*(May), 1–11. https://doi.org/10.3389/fpsyg.2014.00445

Ridge, K. E., Weisberg, D. S., Ilgaz, H., Hirsh-Pasek, K. A., & Golinkoff, R. M. (2015). Supermarket Speak: Increasing Talk Among Low-Socioeconomic Status Families. *Mind, Brain, and Education*, *9*(3), 127–135. https://doi.org/10.1111/mbe.12081

Rigby, D. (2017). *Double trouble. a review of the relationship between UK poverty and economic inequality*.

Roberts, M. Y., & Kaiser, A. P. (2011). The effectiveness of parent-implemented language

interventions: A meta-analysis. *American Journal of Speech-Language Pathology*, *20*(3), 180–199. https://doi.org/10.1044/1058-0360(2011/10-0055)

Roberts, M. Y., & Kaiser, A. P. (2012). Assessing the effects of a parent-implemented language intervention for children with language impairments using empirical benchmarks: A pilot study. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1655–1670. https://doi.org/10.1044/1092-4388(2012/11-0236)

Robson, D. A., Allen, M. S., & Howard, S. J. (2020). Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review. *Psychological Bulletin*, *146*(4), 324–354. https://doi.org/10.1037/bul0000227

Rodriguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Development*, *82*(4), 1058–1075. https://doi.org/10.1111/j.1467-8624.2011.01614.x

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

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-Million-Word Gap: Children's Conversational Exposure Is Associated With Language-Related Brain Function. *Psychological Science*, *29*(5), 700–710. https://doi.org/10.1177/0956797617742725

Rose, D. (1995). *Official Social Classifications in the UK. Social Research Update.* United Kingdom: Department of Sociology, University of Surrey.

Rose, D., & Pevalin, D. J. (2001). The National Statistics Socio- economic Classification: unifying official and sociological approaches to the conceptualisation and measurement of social clasS. In *ISER Working Paper Series, No. 2001-04*. University of Essex, Institute for Social and Economic Research (ISER).

Rose, D., & Pevalin, D. J. (2005). *The National Statistics Socio-economic Classification: Origins, Development and Use*.

Rosenberg, R. (2012). *MCS2 : Guide to Derived Variables*. *August*.

Rothwell, D. W., & Han, C. K. (2010). Exploring the relationship between assets and family stress among low-income families. *Family Relations*, *59*(4), 396–407. https://doi.org/10.1111/j.1741-3729.2010.00611.x

Roulstone, S., Law, J., Rush, R., Clegg, J., & Peters, T. (2010). Investigating the role of

language in children's early educational outcomes. In *Department for Education*.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(1), 185–205. https://doi.org/10.1017/S0305000907008343

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech 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., Denmark, N., Harden, B. J., & Stapleton, L. M. (2016). The Role of Parent Education and Parenting Knowledge in Children's Language and Literacy Skills among White, Black, and Latino Families. *Infant and Child Development*, *25*(2), 198–220. https://doi.org/10.1002/icd.1924

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in Early Gesture Explain. *Science*, *323*(13), 951–953. https://doi.org/DOI: 10.1126/science.1167025

Rowe, M. L., Leech, K. A., & Cabrera, N. (2017). Going Beyond Input Quantity: Wh-Questions Matter for Toddlers' Language and Cognitive Development. *Cognitive Science*, *41*, 162–179. https://doi.org/10.1111/cogs.12349

Rowe, M. L., Pan, B. A., & Ayoub, C. (2005). Predictors of Variation in Maternal Talk to Children: A Longitudinal Study of Low-Income Families. *Parenting: Science and Practice*, *5*(3), 259–283. https://doi.org/10.1207/s15327922par0503_3

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The Pace of Vocabulary Growth Helps Predict Later Vocabulary Skill. *Child Development*, *83*(2), 508–525. https://doi.org/10.1111/j.1467-8624.2011.01710.x

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

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2003). Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech. *Journal of Child Language*, *30*(3), 609–635. https://doi.org/10.1017/S0305000903005695

Roy, P., Chiat, S., & Dodd, B. (2014). *Language and Socioeconomic Disadvantage: From Research to Practice*. London, UK: City University London.

RStudio Team. (2020). *RStudio: Integrated Development for R. RStudio, PBC., Boston, MA*.

RStudio, PBC. http://www.rstudio.com/

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

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* Wiley.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. Longman Publishing Group.

Sadler, K., Vizard, T., Ford, T., Marcheselli, F., Pearce, N., Mandalia, D., Davis, J., Brodie, E., Forbes, N., Goodman, A., Goodman, R., & McManus, S. (2018). *Mental Health of Children and Young People in England, 2017*.

Salmon, K., O'Kearney, R., Reese, E., & Fortune, C. A. (2016). The Role of Language Skill in Child Psychopathology: Implications for Intervention in the Early Years. *Clinical Child and Family Psychology Review*, *19*(4), 352–367. https://doi.org/10.1007/s10567-016-0214-1

Sandel, M. J. (2020). *Tyranny of merit : what's become of the common good?* Allen Lane.

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

Schneider, D., Hastings, O. P., & LaBriola, J. (2018). Income Inequality and Class Divides in Parental Investments. *American Sociological Review*, *83*(3), 475–507. https://doi.org/10.1177/0003122418772034

Schneider, S. (2011). Measuring Educational Attainment. In *Survey question bank: Topic Overview*. https://onlinelibrary.wiley.com/doi/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

Schoon, I., Parsons, S., Rush, R., & Law, J. (2010). Children's Language Ability and Psychosocial Development: A 29-Year Follow-up Study. *Pediatrics*, *126*(1), e73–e80. https://doi.org/10.1542/peds.2009-3282

Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction*, *49*, 157–165. https://doi.org/10.1016/j.learninstruc.2017.01.005

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

Scottish Government. (2004). *Scottish Index of Multiple Deprivation (SIMD) 2004*. https://data.gov.uk/dataset/535abdb8-b3d6-4802-ab9a-7646c417158c/scottish-index-of-multiple-deprivation-simd-2004

Selzam, S., Krapohl, E., Von Stumm, S., O'Reilly, P. F., Rimfeld, K., Kovas, Y., Dale, P. S., Lee, J. J., & Plomin, R. (2017). Predicting educational achievement from DNA. *Molecular Psychiatry*, *22*(2), 267–272. https://doi.org/10.1038/mp.2016.107

Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Rimfeld, K., Krapohl, E., Haworth, C. M. A., Dale, P. S., & Plomin, R. (2013). Strong genetic influence on a UK nationwide test of educational achievement at the end of compulsory education at age 16. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0080341

Shanks, T. R. W. (2007). The Impacts of Household Wealth on Child Development. *Journal of Poverty*, *11*(2), 93–116. https://doi.org/10.1300/J134v11n02_05

Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. *Journal of Abnormal Child Psychology*, *34*(3), 379–391. https://doi.org/10.1007/s10802-006-9027-x

Shaw, B., Baars, S., Menzies, L., Parameshwaran, M., & Allen, R. (2017). *Low income pupils' progress at secondary school*. https://www.tandfonline.com/doi/full/10.1300/J134v11n02_05

Shaw, M., Galobardes, B., Lawlor, D. A., Lynch, J., Wheeler, B., & Smith, G. D. (2005). Northern Ireland Multiple Deprivation Measure (NIMDM) 2005. *The Handbook of Inequality and Socioeconomic Position*, 120–123. https://doi.org/10.2307/j.ctt1t892cc.48

Shinwell, J., & Defeyter, M. A. (2017). Investigation of Summer Learning Loss in the UK—Implications for Holiday Club Provision. *Frontiers in Public Health*, *5*, 1–7. https://doi.org/10.3389/fpubh.2017.00270

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453. https://doi.org/10.3102/00346543075003417

Slusser, E., Ribner, A., & Shusterman, A. (2019). Language counts : Early language mediates the relationship between parent education and children's math ability. *Developmental Science*, *22*(3). https://doi.org/10.1111/desc.12773

Smith, B., Morris, S., & Armitage, H. (2021). *Guidance for evaluators and funders on using GCSE performance as an outcome measure*.

Smith, M. K. (1941). *Measurement of the size of general english vocabulary through the elementary grades and high school*. *24*, 311–345.

Snowling, M. J., Adams, J. W., Bishop, D. V. M., & Stothard, S. E. (2001). Educational attainments of school leavers with a preschool history of speech-language impairments. *International Journal of Language & Communication Disorders*, *36*(2), 173–183. https://doi.org/10.1080/13682820010019892

Snowling, M. J., Bishop, D. V. M., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(8), 759–765. https://doi.org/10.1111/j.1469-7610.2006.01631.x

Snowling, M. J., & Hulme, C. (2020). Annual Research Review: Reading disorders revisited – the critical importance of oral language. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *5*, 635–653. https://doi.org/10.1111/jcpp.13324

Snowling, M. J., Moll, K., & Hulme, C. (2021). Language difficulties are a shared risk factor for both reading disorder and mathematics disorder. *Journal of Experimental Child Psychology*, *202*, 105009. https://doi.org/10.1016/j.jecp.2020.105009

Social Mobility Commission. (2019). *Social mobility in Great Britain - state of the nation 2018 to 2019*. Social Mobility Commission.

Sohr-Preston, S. L., Scaramella, L. V., Martin, M. J., Neppl, T. K., Ontai, L., & Conger, R. (2013). Parental Socioeconomic Status, Communication, and Children's Vocabulary Development: A Third-Generation Test of the Family Investment Model. *Child Development*, *84*(3), 1046–1062. https://doi.org/10.1111/cdev.12023

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

Spencer, S., Clegg, J., & Stackhouse, J. (2013). Language, social class and education: Listening to adolescents' perceptions. *Language and Education*, *27*(2), 129–143. https://doi.org/10.1080/09500782.2012.760585

Spencer, S., Clegg, J., Stackhouse, J., & Rush, R. (2017). Contribution of spoken language and socio-economic background to adolescents' educational achievement at age 16 years. *International Journal of Language and Communication Disorders*, *52*(2), 184–

196. https://doi.org/10.1111/1460-6984.12264

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019a). Language Does Matter: But There is More to Language Than Vocabulary and Directed Speech. *Child Development*, *90*(3), 993–997. https://doi.org/10.1111/cdev.13125

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019b). Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Child Development*, *90*(4), 1303–1318. https://doi.org/10.1111/cdev.13072

St Clair, M. C., Forrest, C. L., Yew, S. G. K., & Gibson, J. L. (2019). Early Risk Factors and Emotional Difficulties in Children at Risk of Developmental Language Disorder: A Population Cohort Study. *Journal of Speech, Language, and Hearing Research*, *62*(8), 2750–2771. https://doi.org/10.1044/2018_JSLHR-L-18-0061

Stats Wales. (2005). *WIMD 2005*. https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation/Archive/WIMD-2005

National Survey for Wales 2017-18, (2018).

Stolarova, M., Brielmann, A. A., Wolf, C., Rinker, T., Burke, T., & Baayen, H. (2016). Early Vocabulary in Relation to Gender, Bilingualism, Type, and Duration of Childcare. *Advances in Cognitive Psychology*, *12*(3), 130–144. https://doi.org/10.5709/acp-0192-6

Strand, S. (2011). The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal*, *37*(2), 197–229. https://doi.org/10.1080/01411920903540664

Strand, S. (2014). Ethnicity, gender, social class and achievement gaps at age 16: Intersectionality and "getting it" for the white working class. *Research Papers in Education*, *29*(2), 131–171. https://doi.org/10.1080/02671522.2013.767370

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

Suskind, D. L., Leffel, K. R., Graf, E., Hernandez, M. W., Gunderson, E. A., Sapolich, S. G.,

Suskind, E., Leininger, L., Goldin-Meadow, S., & Levine, S. C. (2016). A parent-directed language intervention for children of low socioeconomic status: a randomized controlled pilot study. *Journal of Child Language*, *43*(2), 366–406. https://doi.org/10.1017/S0305000915000033

Suskind, D. L., Leung, C. Y. Y., Webber, R. J., Hundertmark, A. C., Leffel, K. R., Suskind, E., Hernandez, M. W., & Graf, E. (2018). Development of the Survey of Parent/Provider Expectations and Knowledge (SPEAK). *First Language*, *38*(3), 312–331. https://doi.org/10.1177/0142723717737691

Sutherland, A., Ilie, S., & Vignoles, A. (2015). Factors associated with achievement: Key stage 4. In *Department for Education*. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/473673/RR407_-_Factors_associated_with_achievement_-_key_stage_4.pdf

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, *72*(3), 748–767. https://doi.org/10.1111/1467-8624.00313

Tamis-LeMonda, C. S., & Rodriguez, E. T. (2008). Parents' Role in Fostering Young Children's Language and Literacy Development. *The Encyclopedia on Early Childhood Development*, 1–10.

Thabrew, H., Stasiak, K., Bavin, L. M., Frampton, C., & Merry, S. (2018). Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents. *International Journal of Methods in Psychiatric Research*, *27*(3), 1–9. https://doi.org/10.1002/mpr.1610

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *Lancet*, *379*, 117–130. https://doi.org/10.1016/S0140-6736(11)60871-4

The Children's Comissioner. (2019). *Briefing: The Children Leaving School with Nothing*. 1–11. https://doi.org/https://www.childrenscommissioner.gov.uk/wp-content/uploads/2019/09/cco-briefing-children-leaving-school-with-nothing.pdf

The Children's Society. (2019). *The Good Childhood Report*.

the National Parent Forum of Scotland. (n.d.). *Progression in the Senior Phase*.

The Sutton Trust. (2011). Improving the impact of teachers on pupil achievement in the UK – interim findings. In *The Sutton Trust*. http://www.suttontrust.com/researcharchive/improving-impact-teachers-pupil-achievement-uk-interim-findings/

The Sutton Trust. (2019). *Making the Grade 2019*.

Thompson, E. J., Kazantseva, A., & Gaysina, D. (2017). Internalizing Psychopathology across the Life Course: From Genes and Environment to Gene-Environment Interaction. *Psychopathology Review*, *4*(1), 26–51. https://doi.org/10.5127/pr.038415

Thornton, E., Matthews, D., Patalay, P., & Bannard, C. (2021). Tracking the relation between different dimensions of socio-economic circumstance and vocabulary across developmental and historical time. *PsyArXiv*. https://doi.org/10.31234/osf.io/bu3px

Thorpe, K., Rutter, M., & Greenwood, R. (2003). Twins as a natural experiment to study the causes of mild language delay: I: Design; Twin-singleton differences in language, and obstetric risks. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(3), 326–341. https://doi.org/https://doi.org/10.1111/1469-7610.00126

Tickell, D. C. (2011). *The Early Years: Foundations for life, health and learning*.

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

Tomas, E., & Vissers, C. (2019). Behind the Scenes of Developmental Language Disorder: Time to Call Neuropsychology Back on Stage. *Frontiers in Human Neuroscience*, *12*, 1–10. https://doi.org/10.3389/fnhum.2018.00517

Tomasello, M., & Farrar, M. J. (1986). Joint Attention and Early Language. *Child Development*, *57*(6), 1454–1463.

Tomblin, J. B., & Zhang, X. (2006). The Dimensionality of Language Ability in School-Age Children. *Journal of Speech, Language, and Hearing Research*, *49*, 1193–1208.

Treanor, M. (2017). Can we put the "poverty of aspiration" myth to bed now? *Centre for Research on Families and Relationships*, 1–6. https://dspace.stir.ac.uk/bitstream/1893/26654/1/Can we put the poverty of aspirations myth to bed now.pdf

Trentacosta, C. J., & Fine, S. E. (2010). Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*, *19*(1), 1–29. https://doi.org/10.1111/j.1467-9507.2009.00543.x

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., Halpern-Meekin, S., Fox, N. A., & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. *Proceedings of the National Academy of Sciences*, *119*(5), e2115649119. https://doi.org/10.1073/pnas.2115649119

UCAS. (2016). *Factors associated with predicted and achieved A level attainment*. https://www.ucas.com/file/71796/download?token=D4uuSzur

UCAS. (2021). *NVQS*. https://www.ucas.com/further-education/post-16-qualifications/qualifications-you-can-take/nvqs

Ulker, A., & Guven, C. (2011). Determinants of maternity leave duration in Australia: Evidence from the HILDA survey. *Economic Record*, *87*(278), 399–413. https://doi.org/10.1111/j.1475-4932.2011.00729.x

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

Van Smeden, M., Lash, T. L., & Groenwold, R. H. H. (2020). Reflection on modern methods: Five myths about measurement error in epidemiological research. *International Journal of Epidemiology*, *49*(1), 338–347. https://doi.org/10.1093/ije/dyz251

Velthuis, S., Lupton, R., Thomson, S., & Unwin, L. (2018). *The characteristics and post-16 transitions of GCSE "lower attainers."*

Vernon-Feagans, L., & Bratsch-Hines, M. E. (2013). Caregiver–child verbal interactions in child care: A buffer against poor language outcomes when maternal language input is less. *Early Childhood Research Quarterly*, *28*(4), 858–873. https://doi.org/10.1016/j.ecresq.2013.08.002

Vernon-Feagans, L., Garrett-Peters, P., Willoughby, M., Mills-Koonce, R., Cox, M., Blair, C., Burchinal, P., Burton, L., Crnic, K., Crouter, N., Garrett-Peters, P., Granger, D., Greenberg, M., Lanza, S., Miccio, A., Mills-Koonce, R., Skinner, D., Stifter, C., Werner, E., & Willoughby, M. (2012). Chaos, poverty, and parenting: Predictors of early language development. *Early Childhood Research Quarterly*, *27*(3), 339–351. https://doi.org/10.1016/j.ecresq.2011.11.001

Vihman, M., & Croft, W. (2007). Phonological development: toward a "radical" templatic phonology. *Linguistics*, *45*(4). https://doi.org/10.1515/LING.2007.021

Viner, R. M., Ross, D., Hardy, R., Kuh, D., Power, C., Johnson, A., Wellings, K., McCambridge, J., Cole, T. J., Kelly, Y., & Batty, G. D. (2015). Life course epidemiology: Recognising the importance of adolescence. *Journal of Epidemiology and Community Health*, *69*(8), 719–720. https://doi.org/10.1136/jech-2014-205300

von Stumm, S. (2017). Socioeconomic status amplifies the achievement gap throughout compulsory education independent of intelligence. *Intelligence*, *60*, 57–62. https://doi.org/10.1016/j.intell.2016.11.006

von Stumm, S., & Plomin, R. (2021). Using DNA to predict intelligence. *Intelligence*, *86*, 101530. https://doi.org/10.1016/j.intell.2021.101530

von Stumm, S., Rimfeld, K., Dale, P. S., & Plomin, R. (2020). Preschool Verbal and

Nonverbal Ability Mediate the Association Between Socioeconomic Status and School Performance. *Child Development*, *91*(3), 705–714. https://doi.org/10.1111/cdev.13364

Wadman, R., Botting, N., Durkin, K., & Conti-Ramsden, G. (2011). Changes in emotional health symptoms in adolescents with specific language impairment. *International Journal of Language and Communication Disorders*, *46*(6), 641–656. https://doi.org/10.1111/j.1460-6984.2011.00033.x

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

Wake, M., Tobin, S., Girolametto, L., Ukoumunne, O. C., Gold, L., Levickis, P., Sheehan, J., Goldfeld, S., & Reilly, S. (2011). Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's learn language cluster randomised controlled trial. *BMJ (Online)*, *343*(7821), 1–10. https://doi.org/10.1136/bmj.d4741

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

Wellcome Trust. (2016). *Purpose and quality of education in England inquiry*. https://wellcome.org/sites/default/files/wtp060177.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

Westrupp, E. M., Reilly, S., McKean, C., Law, J., Mensah, F., & Nicholson, J. M. (2019). Vocabulary Development and Trajectories of Behavioral and Emotional Difficulties Via Academic Ability and Peer Problems. *Child Development*, *91*(2), e365–e382. https://doi.org/10.1111/cdev.13219

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

White, M., Adams, J., & Heywood, P. (2009). How and why do interventions that increase health overall widen inequalities within populations? In S. Babones (Ed.), *Health, inequality and society* (pp. 65–82). Policy Press.

https://doi.org/10.46692/9781847423221.006

Wigelsworth, M., Humphrey, N., Kalambouka, A., & Kalambouka, A. (2010). A review of key issues in the measurement of children's social and emotional skills. *Educational Psychology in Practice*, *26*(2), 173–186. https://doi.org/10.1080/02667361003768526

Wiliam, D., & Bartholomew, H. (2004). It's not which school but which set you're in that matters: the influence of ability grouping practices on student progress in mathematics1. *British Educational Research Journal*, *30*(2), 279–293. https://doi.org/10.1080/0141192042000195245

Willner, C. J., Gatzke-Kopp, L., & Bray, B. C. (2017). The Dynamics of Internalizing and Externalizing Comorbidity Across the Early School Years. *Developmental Psychopathology*, *28*(4), 1033–1052. https://doi.org/10.1017/S0954579416000687

Wilson, A. C., & Bishop, D. V. M. (2019). "If you catch my drift...": ability to infer implied meaning is distinct from vocabulary and grammar skills. *Wellcome Open Research*, *4*(68), 1–30. https://doi.org/10.12688/wellcomeopenres.15210.3

Woodward, P. (2019). Higher Education and Social Inclusion: Continuing Inequalities in Access to Higher Education in England. *Handbook on Promoting Social Justice in Education*, 1–23. https://doi.org/10.1007/978-3-319-74078-2_9-1

World Health Organization. (2018). *Mental health: strengthening our response*. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

World Health Organization. (2019). The WHO Special Initiative for Mental Health (2019-2023): Universal Health Coverage for Mental Health. *The WHO Special Initiative for Mental Health (2019-2023)*, 1–4. https://apps.who.int/iris/bitstream/handle/10665/310981/WHO-MSD-19.1-eng.pdf?ua=1

World Health Organization. (2021). *Depression*. https://www.who.int/news-room/fact-sheets/detail/depression

Yeung, W. J., & Conley, D. (2008). Black-white achievement gap and family wealth. *Child Development*, *79*(2), 303–324. https://doi.org/10.1111/j.1467-8624.2007.01127.x

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(5), 516–524. https://doi.org/10.1111/jcpp.12009