The forgotten fifth:

Examining the early education trajectories of teenagers who fall below the expected standards in GCSE English language and maths examinations at age 16

Lee Elliot Major and Sam Parsons

CENTRE FOR
LONGITUDINAL
STUDIES

UKRI

Economic
and Social
Research Council

Contact the authors

Lee Elliot Major

University of Exeter

L.Elliot-Major@exeter.ac.uk

Sam Parsons

UCL Centre for Longitudinal Studies

sam.parsons@ucl.ac.uk

Disclaimer

This working paper has not been subject to peer review.
CLS working papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.
Any opinions expressed here are those of the author(s) and not those of the UCL Centre for Longitudinal Studies (CLS), the UCL Social Research Institute, University College London, or the Economic and Social Research Council.

How to cite this paper

Elliot Major, L., Parsons, S. (2022) The forgotten fifth: examining the early education trajectories of teenagers who fall below the expected standards in GCSE English language and maths examinations at age 16. CLS Working Paper 2022/6. London: UCL Centre for Longitudinal Studies.

Abstract

Successive Governments have failed to address an issue that continues to plague the British education system: many teenagers leave secondary school without the 'expected standard' of a grade 4 pass in GCSE English language and maths. We use the UK Millennium Cohort Study (MCS) to assess the antecedents and academic trajectories of the 1 in 5 (18%) of teenagers in England who did not achieve these two thresholds, despite benefitting from many education reforms. We find that identification of falling below expected standards in pre-school assessment of 'school readiness' (age 3) and in teacher assessment of literacy and number skills at school entrance (age 5) are both highly predictive of failure to attain a grade 4 or higher in GCSE English language and maths at age 16, even after controlling for family background and individual characteristics. One in 4 children below expected levels at age 5 fail to achieve a grade 4 pass or above in their English language and maths GCSE at age 16 compared with 1 in 10 children above expected levels at age 5. Half of pupils who fail at age 16 were judged to be behind at age 5. We conclude that future attempts to improve standards in English and maths will likely only succeed if high quality support is provided during the pre-school years, support is provided to improve the home learning environment, and teachers are able to identify, diagnose and respond appropriately to children falling behind at early education stages.

Acknowledgements

Introduction

This research aims to shed light on one of education's seemingly most intractable problems: how can it be that so many pupils in England leave school lacking basic levels of literacy and number skills? To those outside education it is unfathomable that after 12 years of schooling, hundreds of thousands of pupils are assessed in public examinations at age 16 to have not met the required standards in English language and maths.

In practical terms this means many pupils will leave school unable to digest a train timetable, compare the cost of products and services, spot fake news or bias in the media, or understand a medical prescription according to the National Literacy Trust and the National Numeracy organisations which campaign to improve literacy and numeracy (National Literacy Trust, 2017; National Numeracy, 2022). Despite countless policy initiatives, successive Governments have failed to address an issue that has plagued the education system for at least several decades (Rashid & Brooks, 2010). In 2017 the then education secretary, Justine Greening announced that a GCSE grade 4 in England was considered by the Government as the 'standard' pass rate for pupils (Greening, 2017). Achieving a grade 4 in a maths GCSE means "you have the essentials of numeracy, the core skills needed for daily life" according to National Numeracy (2022). Grade 4 is widely interpreted as a 'good pass', being equivalent to the former (low) grade C pass, and acts as an effective passport to A Levels and equivalent Level 3 courses and apprenticeships. It also determines whether pupils need to continue to study English and maths and re-sit GCSE exams as a condition of the funding of their post-16 study (Gov.UK, 2022a).

International comparisons suggest that the UK, at least historically, has had a particular problem in producing school leavers with these fundamental life skills compared with many other nations (Elliot Major & Machin, 2018). As a result, a high proportion of the population are unable to live full and functional lives and contribute fully to society and the economy. The skills these examination passes represent are critical to our capacity to communicate and live and work together, and their contribution to workforce skills has increasingly been recognised as critical to economic success (Kuczera et al., 2016). Furthermore, the intergenerational transmission of this educational disadvantage is highly likely to persist as low skilled parents find it harder to support their own child's reading or learning (Sammons et al., 2014) and inequalities in education achievement at age 16 are an important factor in determining the persistence of family income disadvantage across generations (Blanden, Gregg & Macmillan, 2010; Blanden & Macmillan, 2016).

There is a high correlation between literacy and number skills; students who do well in English language examinations also perform well in maths (Rashid & Brooks,

2010). Our focus here is on the pupils who fail to pass a grade 4 in *both* English language and maths GCSEs, a group of pupils who are particularly characterized by low achievement more generally. Using data presented by Lupton et al. (2021) on the 24% of pupils in state funded schools in England who did not achieve an A*-C (grade 4+) GCSE in English and maths in 2015, it is apparent that just one in 50 of these pupils achieved 5 or more grade 4+ GCSEs in other subjects, with 5+ good grade GCSEs being another standard school achievement level.[1] Therefore, whilst acknowledging pupils without a grade 4 in both English language and maths are a heterogeneous group who will possess other skills and achievements, this group of students do overlap considerably with what Kuczera at al. (2016) define as the 'low skilled' – those who were assessed as below level 2 in either literacy or numeracy in the OECD's Survey of Adult Skills.

Members of the UK Millennium Cohort Study (MCS) in England are an ideal population to study students' GCSE performance. The children were born in 2000/1 and have experienced successive Government policy interventions aimed at improving pupil outcomes, many focused on improving literacy and number skills (see Appendix Table A1). Given the scale of major education reforms experienced, this cohort might be considered a golden generation with more support being offered during their school career than any generation before or since. Since the first survey when they were nine months of age, the MCS children have been followed up a further six times between age three and 17, providing an ideal resource to examine educational trajectories from the pre-school years to performance in public examinations at age 16. The aim of our research is to further our understanding of why so many young people continue to fail, or more specifically not meet expected standards, in key public examinations at age 16 – and to what extent direct assessment of 'school readiness' at age 3 and teacher assessments of children's skills and aptitudes in the earliest year of formal schooling (Reception year) predict these future education outcomes. Previous evidence shows that pupils among the lowest fifth of performers in England in learning and development at the end of the Early Years Foundation Stage (EYFS) – at age 5 – are six times more likely to be in the lowest fifth at Key Stage 1 at age 7 (DCSF, 2008) and performance in Key Stage 2 (age 11) is highly correlated with GCSE attainment (Sammons et al., 2014). Early identification of need followed by appropriate support should clearly be part of any effective approach to tackling education disadvantage and helping children overcome specific obstacles to learning.

Understanding more about the antecedents and determinants of low literacy and inadequate number skills in the earliest years of schooling and the association between early and later assessment performance, may offer valuable insight into how skill levels in the future can be improved and this seemingly ever repeating story

---

[1] See Table 3.1 and Table 4.1 in Lupton et al., (2021). 576,132 pupils took GCSEs in 2015, of which 140,884 did not gain an A*-C in English and maths. Among this group, just 3,129 (2.2%) achieved 5+ A*-C GCSEs.

can be avoided – or at least minimised. The richness of data gathered in the MCS allows us to incorporate a wide range of individual and family background characteristics including children's home learning environments into our analysis to help identify the key risk factors associated with poor literacy and number skills in the early years and with not achieving the required standard pass in English language and maths GCSE qualifications, which increasingly supports successful post-16 transitions in the 21st century.

Later life outcomes are profoundly shaped by the cognitive and non-cognitive skills exhibited during the pre-school years (Heckman & Mosso, 2014), with low levels of education and poor adult basic skills impacting negatively across all domains of adult life (e.g. see Bynner & Parsons, 1997, 2006). Here we seek to highlight the trajectory of disadvantage that follows a group of children from pre-school to school entrance and achievement 12 years later, and to highlight areas of learning which might help teachers (and parents) to do more to help this often neglected group of children, and to also aid policy makers to deliver effective strategies to back up their lofty rhetoric about banishing low literacy and number skills. We hope this work will provide some practical ways forward.

The paper proceeds as follows. We first review the relevant literature in this area; we then set out our assumptions, data and methods; we then present our results; finally, we discuss our findings and the implications for Government policy.

Literature review

Prevalence of low attainment at age 16

For at least 50 years, and probably more, a substantial proportion of school leavers in England have left education with poor literacy and number skills, increasingly defined by their failure to attain a 'good grade' in public examinations in English language and maths at the age of 15 and 16. In recent years England's exams regulator, Ofqual, has undertaken an annual review to assess evidence from examiners to determine whether to alter the proportion of passes in any given year (Newton, 2021). In the decade preceding the Covid pandemic (during which alternative and teacher assessments replaced examinations) the proportion of pupils passing the expected benchmarks in these core subjects remained largely unchanged: GCSE examinations have effectively been 'norm-referenced' with the same proportions of pupils fixed at successive grade boundaries. The expected national benchmark for English language and maths GCSEs changed in 2017 from a grade C to a grade 4 (equivalent to a low C grade). All else being equal, one might have predicted that pass rates would have increased slightly in light of this change.

National Reference Tests, undertaken since 2017, meanwhile have shown little change in standards (Ofqual, 2021). In essence, whether the expected national benchmark is a grade C or a grade 4, a substantial proportion of pupils are destined to not get over the line. In 2019, when standard examinations in England were last sat before the Covid-19 pandemic, a third of year 11 students did not attain a grade 4 or higher in either English Language or Maths once again (ASCL, 2019; Ofqual, 2019).

Since September 2015 students have been required to remain in education or learning until age 18 (Gov.UK, 2022b) and many re-sit their English language or maths examination if they do not gain a grade 4 or higher the first-time round. Attaining a 'good grade' pass in these subjects is increasingly fundamental for accessing the widest range of possible post-16 transitions and prospering in life after school, with this link appearing to be particularly strong in England (Kuczera at al., 2016). Concentrating specifically on students not passing expected standards in GCSE English or maths, Lupton et al (2021) found that post-16 transitions for lower attainers tended to be more complex and difficult when compared with their higher attaining peers. Whether pupils fall far below or just miss the 'expected standard' (grade C or 4) pass line may have equal long-standing consequences. For example, Machin et al (2020) show that pupils who narrowly miss out on a grade C/4 in their English GCSE by just a few marks can pay a heavy price for this failure: they are for example less likely to study for A-levels and to attend university. Attainment in English and maths at 16 can also influence later labour market outcomes (see Dickerson et al., 2020), with stronger literacy and numeracy skills in adulthood being associated with higher employment rates (Vignoles, 2016), and in spending increased amounts of time in employment over the lifecourse (Bynner & Parsons 1997, 2002). In general, pupils doing poorly in GCSEs can be scarred for many years, finding it hard to recover in the workplace (Bell & Blanchflower, 2010; Crawford et al., 2011; Ralston et al., 2016; Thompson, 2017).

Looking wider than GCSE attainment, Rashid and Brooks (2010) summarised an extensive range of national and international surveys and test and examination results over the period 1948–2009, and concluded that the proportion of 16-19 year-olds with poor reading and poor numeracy had remained stubbornly high – at about 17% and 22% respectively – meaning a fifth of the population could not understand basic texts and simple information described by numbers – skills needed for full participation in society. Elliot Major and Machin (2018) estimated that 25% of 16-19-year-olds in England left school without basic literacy and number skills, after analysing 2012 data from the OECD's Programme for the International Assessment of Adult Competencies (PIAAC). This was a higher proportion than in most developed countries. A separate OECD review labelled British youngsters the "most illiterate in the developed world" (OECD, 2016).

The cost of this 'failure' is not just incurred in terms of the poorer life prospects of individuals, but collectively, with the annual cost to the nation of low numeracy alone having been estimated to be about 1.3 per cent of the country's Gross Domestic Product (Martin et al., 2014), with far more claimants of Jobseekers Allowance in England in 2011 having poorer literacy and numeracy skills (BIS, 2012) than the general working age adult population. It has never been more important to increase our understanding of the barriers to learning that prevent this group of students from achieving the expected standards after 12 years of compulsory schooling.

Education policies

As noted, there has been a plethora of government initiatives aimed at improving the literacy and number skills of children and adults in England (and the UK more broadly), particularly since the late 1990s and the time of New Labour and their '*education, education, education*' mantra[2]. As such, the 2000/1 Millennium Cohort is a particularly pertinent population to study given the major Government policy initiatives aimed at improving pupil outcomes have potentially 'benefitted' this cohort during their school years more than previous generations. Many were implemented when New Labour were in power (1997-2010). These have included: Sure Start centres providing support during the pre-school years; Early Years Foundation Stage (EYFS) profile assessments to diagnose pupil needs; nationally prescribed literacy hours and numeracy hours in primary schools; significant increases in school funding; academies, schools independent of local authority control, to take over struggling schools; 'pupil premium' funds for children qualifying for free school meals, coupled with evidence-informed guidance on how to spend funds effectively; and reforms to the curriculum and grading of English language and maths GCSEs. (See Appendix Table A1 for a comprehensive list of policy reforms from 1996-2013). As evidenced by the stubborn proportion of students who continue to fall below the expected standard in their public examinations at age 16 (ASCL, 2019), it appears however that these extensive policy efforts to improve outcomes at age 16 for all pupils have fallen short. One in three students in England do not achieve a grade 4 or higher in English language or maths in their GCSE examinations at the end of year 11.

---

[2] Part of Tony Blair's 2001 speech launching Labour's education manifesto: 'Our top priority was, is and always will be education, education, education. To overcome decades of neglect and make Britain a learning society, developing the talents and raising the ambitions of all our young people.' Full speech available at: https://www.theguardian.com/politics/2001/may/23/labour.tonyblair

Family and individual characteristics associated with low attainment

One possible explanation for this failure is that many factors driving outcomes occur early in children's lives and emanate from the home and family environment – limiting what can be achieved solely in the classroom. An extensive literature has documented the many early years factors associated with children who go on to suffer low educational and cognitive achievement in general, as well as in English and maths GCSEs specifically. The social-economic gradient in cognitive and academic achievements over the life-course is well established – whether classified in terms of parental income, social class or education levels (see for example: Halsey et al., 1980; Feinstein, 2003; Blanden, Gregg, & Macmillan, 2007; Parsons et al., 2011; Sullivan et al., 2013; Stopforth et al., 2020; Stopforth & Gayle, 2022), with the gap in GCSE attainment in England by eligibility for free school meals being an omnipresent statistic (Sutherland et al., 2015; DfE, 2020, 2022a). The intergenerational transmission of low education attainment or skills is particularly pertinent here, with children whose parents have no or few formal qualifications being far less likely to attain good grade GCSEs at age 16 in England (Sammons et al., 2014), and children of parents with the poorest numeracy skills being twice as likely to perform poorly in number skills assessments (Bynner & Parsons, 2006). Poor education begets poor education.

Disadvantaged neighbourhoods (Sammons et al., 2014), poor housing and overcrowding in the home are also related to lower academic attainment (Goux & Maurin, 2003; Office of the Deputy Prime Minister, 2004; Schoon, 2020) and children growing up in a workless household have poorer early academic outcomes and make less progress between age three and five than those living in working households (Parsons, Schoon & Vignoles, 2014).

The early family socio-economic environment that children are exposed to has long-lasting consequences, with the early years having long been identified as the key time for interventions to address skills inequalities among children (Coghlan, 2010; Heckman & Mosso, 2014). The school readiness gap between the poorest and richest children is already 19 months before they have stepped into a classroom (Bradbury et al., 2012), with the other firmly established fact from the literature being the strong link between early performance in cognitive or academic tests and later academic attainment. Around 60 per cent of achievement gaps in England at the age of 14 (KS3) are already present at the start at school (Bradbury et al., 2015), and Passaretta et al (2022) find that 50–80 per cent of language gaps observed at end of primary school are explained by gaps settled before formal schooling in the United Kingdom (as well as in Germany and the Netherlands). Therefore, whilst only 20–50 per cent of school-age gaps were generated during schooling, the failures of the education system to better address the poor skills of students at the start of their education career just store up a bigger problem for future generations.

Bradbury et al (2012) found that around half of the school readiness gap between the poorest and richest children could be linked with the quality of the early home learning environment (HLE). Others have also found a positive HLE to be strongly associated with children's school readiness and early educational and cognitive development (Melhuish et al., 2008), with longer-term effects also observed (Sammons et al., 2014, 2007; Pungello, Kainz & Burchinal, 2010). Key indicators of early home-learning include parents reading to their children, teaching literacy and numeracy and encouraging literacy activities and library visits, all of which are associated with improved school achievement (see also de la Rochebrochard 2012; Kiernan & Huerta, 2008). Importantly, recent research by Stopforth and Gayle (2022) found that engagement in reading related activities was 'mildly influential' in reducing the GCSE attainment gap by family social class.

Focusing specifically on students in England who did not attain a good grade in English language and maths at age 16, a rather neglected group of pupils in the literature, Cassen and Kingdon (2007) found the key characteristics that predicted being a low achiever at age 16 was to be a British white male, to be eligible for free school meals, diagnosed with special educational needs and to have performed poorly in earlier literacy assessments (KS1 and KS2) in primary school. They also drew attention to the importance of pre-school education, helping families to improve the early home learning environment, offering parenting help and support and for policies to address poverty and income inequalities alongside improving schooling environments. All these factors had essential parts to play individually, but collectively they could significantly work to reduce low achievement levels. In contrast, the more disadvantages a child experiences can exacerbate low achievement levels. As noted by Lupton et al (2021) combinations of characteristics are particularly associated with low attainment in GCSE English and maths. For example, the attainment gap between pupils on Free School Meals (FSM) and those not on FSM is bigger for White pupils than for other ethnic groups (Kingdon & Cassen, 2010; Strand, 2014).

Other individual characteristics associated with poorer cognitive and academic attainment include being male, being born with low birthweight, not being breastfed, being younger in the school year and exhibiting health and behaviour problems. Boys do less well in primary school assessments (DfE, 2019; Parsons & Hallam, 2014) and girls have consistently outperformed boys in GCSE attainment over the last three plus decades since 1988 (Smithers, 2014; DfE, 2020). Low birthweight children go on to experience more cognitive deficits on average (Currie & Hyson, 1999), including poorer vocabulary (Taylor et al., 2013), to perform less well in primary school (KS1) (Parsons & Hallam, 2014) and to pass fewer public examinations at age 16 (Case et al., 2006), whereas having been breastfed for longer is associated with improved cognitive development (Quigley et al., 2012). Children born between June and August have been shown to do less well at school on average than their autumn-born classmates in primary school (Sykes et al., 2009;

Parsons & Hallam, 2014) and in public examinations at age 16 (Crawford et al., 2013) and health and behaviour problems in early childhood have been shown to stifle cognitive and academic progress through the school years (Barbaresi et al., 2007; Gutman & Vorhaus 2012; Parsons et al., 2021; Washbrook et al., 2013; Whiteside et al., 2017).

Research aims and questions

Our research adds to this literature in several distinctive ways. First, we assess how many members of the UK Millennium Cohort Study at school in England were not 'school ready' at age 3; who were not at the expected standard in the EYFS Communication, Literacy and Language (CLL) and Maths Development (MD) scales at age 5; the proportion who do not attain a grade 4 or higher in English language and maths GCSEs at age 16; and importantly, show the association between performance in the early years and at the end of secondary schooling. This generation of children, born 2000/1, is of particular interest as it benefited from an unprecedented range of Government reforms aimed at improving literacy and numeracy. By using a longitudinal cohort study, we can trace the early life education trajectories of children who end up without English language and math GCSEs at age 16 and help identify the key family background and individual characteristics that are associated with low literacy and number skills and provide possible levers for policy intervention.

Our specific research questions are:
- What family and individual characteristics are associated with the children who are not 'school ready' at age 3; who are not judged by teachers to be at the expected standards in the Early Years Foundation Stage (EYFS) at age 5; and who do not achieve a grade 4 or higher in GCSE English language and maths at age 16?
- To what extent does not being school ready at age 3 predict not being assessed to be at the expected standard in the EYFS at age 5 – specifically in Communication, Language and Literacy (CLL) and Maths Development (MD)?
- To what extent does not being school ready at age 3 or at the expected standard in the EYFS at age 5 predict not attaining a grade 4 or higher in GCSE English language and maths at age 16?
- What family environment, educational and individual characteristics attenuate the association between performance in assessments at age 3, 5 and 16?
- What are the implications of these findings for policies and strategies aimed at improving basic skills in literacy and numeracy for future generations of pupils?

Data and methods

Millennium Cohort Study

The Millennium Cohort Study (MCS) is a multi-purpose ongoing longitudinal study of approximately 19,000 babies born to families living in the UK between September 2000 and January 2002 (Plewis, 2007; Connelly & Platt, 2014; Joshi & Fitzsimons, 2016). Data has been collected when the children were aged around 9 months, 3, 5, 7, 11, 14 and 17 when approximately 10,700 study members participated. Here we draw on information collected from personal interviews and self-completion questionnaires administered to parents, teachers, and the cohort children from 9 months, 3 years, 5 years and 17 years (University of London, 2022a 2022b 2022c 2021). Information collected includes a wide range of robust family socio-economic, employment, qualification, health, wellbeing, and parenting behaviour, together with child characteristics, cognitive ability, and education attainment.

Analytic sample

Of the 18,552 families who first took part in sweep 1, we restrict our sample to the 11,532 families who lived in England who had provided information on the sex and ethnicity of the cohort child, giving a final analytic sample of 11,524. As in all longitudinal studies, MCS suffers from attrition over time, and at age 17 the response rate for the whole UK sample was 57%.

Multiple Imputation

We used Multiple Imputation (MI) to deal with attrition and item non-response to restore sample representativeness, adopting a chained equations approach (White, Royston & Wood, 2011) under the assumption of 'missing at random' (MAR), which assumes that the most important predictors of missing data are included in our models. To maximise the plausibility of the MAR assumption the most important predictors of missing data are included in our models to further reduce bias and retain power (see Mostafa & Wiggins, 2015; Mostafa et al., 2020; Silverwood et al., 2020). All reported analyses are averaged across 25 replicated data sets based upon Rubin's Rule for the efficiency of estimation under a reported degree of missingness across the whole data of around 0.25 (Little & Rubin, 2014). Missingness in the variables ranged from less than 1% in many of the sweep 1 measures to a high of 51% for information provided by a cohort child teacher at age 11 (See Appendix Table A2 for the level of missingness in all variables included in the imputation.)

The analyses were additionally weighted to adjust for the survey's stratified clustered sampling design (Plewis, 2007).

Key Measures

We have three key outcomes measures: being 'delayed' in the Bracken (1998) direct assessment of school readiness at age 3; being assessed at 'below the expected level' in the literacy and number skills components of the EYFS by teachers at age 5; and not reaching the 'expected standard' of a grade 4 or higher in GCSE English language and maths examinations at age 16.

Bracken School Readiness

At age three MCS cohort members completed the Bracken School Readiness Assessment-Revised (BSRA-R), which is one component of the Bracken Basic Concept Scale-Revised (Bracken, 1998). The BSRA-R is used to assess the 'readiness' of a child for formal education by testing their knowledge and understanding of basic concepts (Bracken, 1998). Basic concepts are defined as aspects of children's knowledge that are taught by parents and pre-school teachers to prepare a child for formal education. The assessment consists of 85 items across five basic concept sub-tests: Colours (10); Letters (15); Numbers/Counting (18); Size/Comparisons (22); and Shapes (20). The acquisition of basic concept knowledge and skills is important for a child's future educational attainment (Breen, 1985; Duncan et al., 2007; Bracken & Crawford, 2010). All items are summed to produce a total score which is age standardised. The age standardised score is used to place cohort members into a five-category 'Normative Classification' variable which ranges from 'very advanced' 'advanced' 'average' 'delayed' and 'very delayed'. Here we compare children who are 'delayed' (combining delayed or very delayed) against those who are 'school ready' (combining average, advanced or very advanced). For further details see Connelly (2013).

Early Years Foundation Stage (EYFS) profile

The Early Years Foundation Stage (EYFS) is the standard set for education, teaching, learning and care of 0 to 5-year-olds. It was first introduced as the Foundation Stage Profile (FSP) in 2000 (Qualifications and Curriculum Authority, 2003), and later became part of the 2006 Childcare Act and must be followed by all Ofsted registered settings and childminders. All teachers of children at school in England complete an Early Years Foundation Stage (EYFS) profile in the final term of Reception Year in Primary school.

The main purpose is to support a successful transition for pupils in their first years of schooling during Key Stage 1 (KS1). The idea is to inform Year 1 teachers about each child's stage of development and learning needs to help them to plan the curriculum to meet the needs of all their pupils. Parents are also provided with the EYFS profile of their child's development.

The EYFS profile is intended to provide a reliable, valid and accurate assessment of each child's development. In its current form, it is made up of an assessment of the child's outcomes relating to 17 early learning goal (ELG) descriptors. Children are defined as having reached a Good Level of Development (GLD) at the end of the EYFS if they have achieved at least the expected level for the ELGs in the prime areas of learning and the specific areas of mathematics and literacy. This helps teachers and parents to understand what a child can do in relation to national expectations. Each ELG has a score range of 0-9, with a score of 6 or higher indicating a child has reached the expected level in a specific ELG (DfE 2022b).

EYFS profile scores: focus on Communication, Language and Literacy and Maths Development

In this research we concentrate specifically on assessments in two areas: Communication, Language and Literacy, and Maths Development. (See Appendix 1 for further details.)

Communication, Language and Literacy (CLL), combines assessment in four separate early learning goal (ELG) descriptors:

- Language for communication and thinking
- Linking sounds and letters
- Reading
- Writing

The maximum score available is 36, with 24 or higher indicating the child has reached the expected level overall.

Maths Development (MD) combines assessment in three separate early learning goal (ELG) descriptors:

- Numbers as labels and for counting
- Calculating
- Shape, space and measures

The maximum score available is 27, with 18 or higher indicating the child has reached the expected level overall.

GCSE examinations in English language and maths

General Certificate in Secondary Education (GCSE) examinations were introduced in the 1980s and remain the standard qualifications that are under-taken by pupils in England and Wales at the end of Year 11 (aged 15-16) (Department for Education 1985, Mobley et al. 1986, North 1987). The students will usually study for nine GCSE subjects (Carroll & Gill 2017), which will (usually) include the core subjects of English, maths and science (Jin, Muriel, & Sibieta 2011). Reforms to GCSEs were introduced in 2015, with the first cohorts taking the new exams in 2017 and 2018. GCSEs grades now range from 1 to 9, with a 'good grade' being a grade 4 or higher. Prior to this, GCSE grades ranged from A*-G, with an A*-C grade representing the expected national standard, with a grade C and grade 4 or 5 being broadly equivalent. However, a grade 4 is viewed as a 'standard pass' whereas a grade 5 is a 'strong pass' (Greening, 2017), and the Government report on both. Here we focus on the substantial minority of students who did not secure a grade 4 or above in both English language and maths. Mastering English and maths is the most basic requirement for prospering in life after school – one of the reasons why these core subjects have been increasingly prioritised in school accountability measures. Since 2014 students who did not gain at least a grade C or grade 4 in English language or maths have needed to continue studying the subjects and to re-sit the examination (Lupton et al., 2021).

The MCS teenagers sat their GCSEs in 2016/2017 and reported their grades when interviewed in 2018.

For all three of our binary outcome measures, 0 indicates having reached the expected level of achievement; 1 indicates not having reached the expected threshold at age 3, 5 or 16.

Covariates

We examine school readiness, EYFS assessment and GCSE attainment across a wide range of family environment and individual characteristic measures identified to be associated with academic and cognitive performance from the literature.

Socio-economic characteristics (SES) (9 months)

These measures are taken from the first MCS survey when cohort members were age 9 months. The specific SES measures included in the models are parent highest qualification level (four categories ranging across NVQ levels from None/NVQ1 [below GCSE grade 4 or equivalent]; NVQ2 [GCSE grade 4+ or equivalent]; NVQ3 [A Level or equivalent]; NVQ4/5 [degree+ or equivalent]); whether someone in the household is working (0) or it is a workless household (1); if only English is spoken in the home (0) or English and/or only another language is spoken (1). In terms of

housing, we include the Index of Multiple Deprivation (IMD) which classifies the area where a cohort member lives into 10 deciles and compare the eight more affluent decile areas (0) against the bottom two deciles (1); if the housing is owner occupied (0) or rented (1); whether the home is overcrowded, comparing homes with <1 person per room (0) against those with 1+ person per room (1); and whether the home suffers from dampness, classified as no damp in the home (0) or there is dampness in the home (1). For family status we include if the family has two-parents (0) or a single parent (1) in the home and whether the mother was an older (0) or teenage mother (1). For health, we include a measure of parent general health, comparing those who self-report good, very good or excellent (0) against those who report having poor or fair health (1); whether the mother exhibits a high number of depressive symptoms, as assessed by the shortened 9 question version of the Malaise Inventory which is an established scale to measure signs of psychological distress or depression in teenagers and adults (Rutter, 1970). In the shortened version scores range between 0-9, with a score of 0-3 indicating no/low signs of depression (0), 4+ indicating the mother is experiencing signs of depression (1). For the majority binary measures, 0 indicates the reference category. For parental highest qualification level, NVQ4 or 5 (degree level or higher) is the reference group.

Early education support measures (3 and 5)

We include two early education support measures. At age 3 an overall assessment of the home learning environment, which includes such activities as how often a parent read to the child, sang songs, played with them, or took them to a park (for further details see de la Rochebrochard, 2012). The scores ranged from 0-42 and were split into quintiles. We compared the top four quintiles (0) to the bottom quintile (1). At age 5 we included how often the child was read to. As nearly half of parents reported their child was read to everyday, we compared 0-6 days per week (0), everyday (1).

Individual characteristics

From the first survey we include a child's sex (male = 0; female = 1), their ethnicity (BAME = 0; white = 1), birthweight (normal = 0; low birthweight < 2515 grams = 1), birth order (2nd or later in birth order = 0; 1st born = 1), whether they were breastfed (0=yes; 1=never) and the season they were born in (Autumn = 0; Winter = 1; Spring = 2; Summer = 3) and from the second survey at child age 3 whether they had a longstanding illness (no = 0; yes = 1).

Early behaviour problems (age 3)

Behaviour problems were assessed from parent reports on the Strengths and Difficulties Questionnaire [SDQ]. The SDQ is widely validated cross-nationally and

cross-culturally for use in non-clinical settings. For further details see Goodman (1997, 2001). The SDQ includes 25 measures comprising five scales of five items each. For each negative attribute, the parent is asked to say whether it is 'not true' (0), 'somewhat true' (1) or 'certainly true' (2) about their child's behaviour, with scores reversed for positive attributes. We use the four problem behaviour scales, setting aside the non-problems scale of pro-social behaviour.

- SDQ Conduct problems
- SDQ Hyperactivity problems
- SDQ Peer problems
- SDQ Emotional problems

Each problem behaviour scale ranges from 0-10 but can be dichotomised to indicate 'abnormal' behaviour. A score of 3+ indicates conduct and peer problems; 4+ emotional problems; and 6+ hyperactivity problems (Youth In Mind, 2016). In each binary variable no problems are coded as 0, behaviour problems as 1.

Results

Proportions of children not being 'school ready' pre-school and lacking basic literacy and number skills at age 5 and 16

We first detail the school readiness of our sample at age 3 and then their literacy and number skills levels as assessed by teachers in the EYFS at age 5 and GCSE attainment at age 16.

At age 3, 86% of children were directly assessed to be 'school ready' with the remaining 14% being 'delayed'. At age 5, we find that 76% of children had reached expected levels in maths development (MD); 62% of children meanwhile had reached expected levels in communication, language and literacy (CLL). When teacher assessments in both EYFS areas are combined, we find:

- 60% had reached the 'expected level' in CLL and MD
- 18% had reached the 'expected level' in CLL *or* MD
- 22% had not reached the 'expected level' in CLL *and* MD

Turning to GCSE attainment at age 16, 72% of our sample achieved a good pass in English language, and 71% in maths. This compares favourably with statistics from the Department for Education (DfE) which show that around 7 in 10 students taking GCSEs in England at the end of Year 11 achieved a grade 4 or higher in English

language or maths in 2017 and 2018 (Ofqual, 2018). Considering the benchmark of a grade 4 or higher in both GCSE English language and maths at age 16, we find:

- 61% had a 4-9 (or A*-C) grade in English language and maths
- 21% had a 4-9 (or A*-C) grade in English language *or* maths
- 18% did not have a 4-9 (or A*-C) grade in English language *and* maths

The similar proportions of children at age 5 and at age 16 assessed below expected standards in both subject areas provide a plausible estimate of around 20% of all pupils lacking basic literacy and number skills. This equates to approximately 1.8 million children among the 8.9 million pupils attending schools in England in 2020/21.

Association between early performance and later examination success

Among the vast majority (86%) of children who were 'school ready' at age 3, 83% went on to assessed at the expected standard in both the CLL and MD parts of the EYFS profile at age 5, compared to just 47% of those who were 'delayed'. Looking further ahead, school readiness and teacher assessments in the early years are extremely strong predictors of success in examinations at age 16. For children assessed as 'school ready' at age 3, 85% went on to gain a grade 4 of higher in English Language and Maths at age 16, falling to 65% for those who were 'delayed'. Although around half of children assessed below the expected level in EYFS CLL at age 5 went on to attain a grade 4 or higher GCSE in English language (56%), this increased to 82% for children assessed at the expected level in CLL. There was a similarly strong association between EYFS MD and GCSE maths performance, with 47% of those assessed below the expected level attaining a grade 4 or higher compared to 78% of children at the expected level. For children assessed below the expected level in both CLL and MD, just a third (34%) achieved a grade 4 or higher in both GCSE English language and maths, compared to 74% assessed at the expected level.

Individual and family background characteristics of Millennium Cohort Study children

We now build a richer picture of the early life of children in the MCS. First we examine the association between the Bracken 'school readiness', EYFS teacher assessments and GCSE attainment, and then detail the family background and individual characteristics of a) the children identified as falling behind in the pre-school years and in the expected levels in literacy and numeracy at the start of their school career and b) the teenagers who do not attain a grade 4 or higher pass in GCSE English language and maths (there is a strong overlap amongst these populations).

Tables 1 and 2 provide a descriptive profile of the early family environment and individual characteristics of children by whether they were school ready at age 3, had reached the expected levels in the EYFS communications, language, literacy and maths development at age 5; or had achieved a grade 4 or higher in GCSE English language and maths at age 16.

Early family environment

Table 1 shows that children who were not school ready at age 3 and who were below expected standard levels at age 5 and age 16 share several distinguishing family background characteristics compared to their higher performing peers. We use proportions for children by GCSE results to highlight the differences below.

At each age, children identified as struggling were twice as likely to be born to a teenage mother (13% to 5%) and be living with a single parent (24% to 10%), and three times more likely to be living in a workless household (33% to 11%). Their parents are three times more likely to have no or poor education qualifications (equivalent to low GCSE grades) (31% to 10%) and were less likely to have attained a degree or higher qualifications (22% to 52%). Their home is more likely to be rented (58% to 28%), overcrowded (34% to 21%) or damp (19% to 12%) and situated in poorer areas (classified in the bottom two deciles of the Index for Multiple Deprivation (IMD)) (36% to 18%).

Young children who were not school ready at age 3 were three times more likely to be part of a household where an additional language other than English is spoken (26% to 9%), with differences between groups reducing at age 5 (17% to 10%) and reversing at age 16 (9% to 12%).

Considering early educational support at home, children not school ready or below expected levels at age 5 and 16 were more likely to have a poor home learning environment at age 3 (including being read to and being taught letters and numbers) and were less likely to be read to everyday at age 5.

Individual characteristics

Turning to table 2, we consider children's individual characteristics. Here we see young children and teenagers below expected standard levels are less likely to be female (39% to 53%) and to be the first-born child (37% to 44%), to have never been breastfed (43% to 21%) and they were twice as likely to have had a low birthweight (10% to 6%). Children who were not school ready at age 3 or were below expected standards at age 5 were more likely to be non-white (BAME) (32% to 12%), but differences between achievement groups disappeared – or in fact reversed – by age 16 (13% to 16%). Under achievers at each age point were also more likely than their more capable peers to have identified behaviour problems on the parent completed

SDQ scales at age 3, and under achievers at age 5 were less than half as likely to be autumn born children (14% to 31%) but twice as likely to be younger summer born children in their school year (40% to 19%), but it was only being summer born that was associated with poor GCSE attainment (29% to 24%).

Association between School Readiness, EYFS assessments and GCSEs

As discussed earlier, earlier performance is a strong predictor or later performance. More children below the expected standard at age 5 were classified as 'delayed' from their performance in the Bracken School Readiness assessment (Bracken, 1998). Similarly, being delayed at age 3 and below the expected standard in teacher assessments at the EYFS were both extremely strong predictors of success in GCSE examinations. Teenagers who did not gain a GCSE grade 4 or higher in English Language and maths were three times more likely than their peers who did to be identified as 'delayed' at age 3 (28% to 9%); and almost half (48%) of those who failed to gain a grade 4 or higher in English Language and maths had been assessed below the expected level in EYFS CLL and MD at age 5, compared to 1 in 8 (12%) of those who did get a grade 4 or higher in both. Conversely, whereas three-quarters (72%) of those who did attain a GCSE grade 4 or higher in English Language and maths had been at the expected level in EYFS CLL and MD, this fell to a third (31%) amongst those who had not.

Table 1: Early Family Environment characteristics of cohort children by assessment at age 3, 5 and 16 (column proportions)

	Bracken School Readiness		EYFS (CLL and MD)[1]		English Language and Maths[2]	
	School Ready	Delayed	>= Expected Level	< Expected Level	Grade 4+ both	< Grade 4 both
Family environment (9 months)						
Single Parent	.12	.24	.10	.22	.10	.24
Teenage mother	.07	.12	.05	.12	.05	.13
Workless Household	.14	.36	.10	.32	.11	.33
Parent No quals/NVQ1 quals	.13	.35	.09	.32	.10	.31
Parent Degree+	.47	.21	.53	.23	.52	.22
Rented Home	.34	.61	.28	.59	.28	.58
Overcrowded Home	.22	.42	.19	.38	.21	.34
Damp Home	.13	.20	.13	.18	.12	.19
Bottom two deciles IMD	.20	.45	.17	.37	.18	.36
English +/or other language spoken at home	.09	.26	.10	.17	.12	.09
Parent poor/fair general health	.16	.24	.13	.24	.14	.24
Mother depressive symptoms	.13	.18	.12	.18	.12	.18
Education support (age 3, 5)						
HLE (bottom quintile) (3)	.13	.25	.10	.23	.12	.20
HLE (top quintile) (3)	.25	.12	.28	.15	.25	.19
Read to child everyday (5)	.52	.44	.55	.43	.53	.45
N(100%)=	9911	1613	6926	2535	7064	2051

Note: [1] Cohort Members assessed at the 'expected level' in CLL or MD are excluded from this table; [2] Cohort Members with a Grade 4+ qualification in English language *or* Maths are excluded from this table; proportions in bold indicate difference statistically significant at p<.05

Table 2: Individual characteristics of cohort children by assessment at age 3, 5 and 16 (column proportions)

	Bracken School Readiness		EYFS (CLL and MD)[1]		English Language and Maths[1]	
	School Ready	Delayed	>= Expected Level	< Expected Level	Grade 4+ both	< Grade 4 both
Birth characteristics (9 mths)						
Female	.51	.38	.54	.42	.53	.39
Non-white (BAME)	.12	.32	.12	.22	.16	.13
Low birthweight	.06	.12	.06	.11	.06	.10
1st born child	.44	.34	.46	.37	.44	.37
Never breastfed	.25	.39	.22	.39	.21	.43.
Longstanding illness	.16	.19	.16	.19	.15	.21
Autumn born[3]	na	na	.31	.14	.26	.23
Winter born	na	na	.26	.20	.24	.23
Spring born	na	na	.24	.26	.25	.25
Summer born	na	na	.19	.40	.24	.29
Early behaviour (age 3)						
SDQ Emotional problems	.07	.17	.07	.15	.07	.13
SDQ Conduct problems	.49	.62	.47	.61	.46	.63
SDQ Hyperactivity problems	.23	.41	.20	.39	.20	.39
SDQ Peer problems	.24	.42	.22	.36	.24	.33
Early education (age 3, 5)						
Bracken School Readiness (delayed)	na	na	.06	.35	.09	.28
< expected level CLL and MD (5)	na	na	na	na	.12	.48
>= expected level CLL and MD (5)	na	na	na	na	.72	.31
N(100%)=	9911	1613	6926	2535	7064	2051

Note: [1] Cohort Members assessed at the 'expected level' in CLL or MD are excluded from this table; [2] Cohort Members with a Grade 4+ qualification in English language *or* Maths are excluded from this table; [3] not relevant for school readiness categorisation which is based on an age standardised score; proportions in bold indicate difference statistically significant at p<.05

Regression analyses

To isolate the specific correlations linking different characteristics to later outcomes, we ran a series of logistic regression analyses predicting a) being 'delayed' in the Bracken assessment of school readiness; b) being assessed as having below expected literacy and number skills levels at age 5 (EYFS), and c) failing to gain a grade 4 or higher in GCSE English language and maths. For the EYFS outcome Model 1(a) included performance in the Bracken school readiness assessment at age 3; for the GCSE outcome Model 1(b) included performance in the Bracken school readiness assessment at age 3 and the EYFS CLL and MD. Models 2 to 6 each include the relevant earlier skills assessment measure plus the sets of characteristics as detailed below. For the pre-school Bracken assessment there was no Model 1 as no prior skills assessment was included in the modelling.

- Model 1a: Bracken school readiness: average/advanced v delayed
- Model 1b: Bracken school readiness: average/advanced v delayed and EYFS CLL and/or MD: above/at v below expected level
- Model 2: + early family socio economic environment
- Model 3: + early family education support
- Model 4: + child demographic characteristics
- Model 5: + child behaviour problems
- Model 6: + all measures

Our analysis allows us to identify the predictive power of earlier skills and learning on later learning outcomes. Model 1 shows the raw relationship between earlier and later skills, Model 2 to 5 show the key characteristics associated with low achievement within each set of measures considered, while Model 6 includes all measures together. The full sets of regression results for each of the three outcome measures are included in the appendix (Appendix Tables A3-A5). For ease of interpretation, we have calculated the predicted probabilities of earlier assessment performance being associated with later assessment performance and present these graphically.

Predicting delayed school readiness at age 3

Concentrating on the final model (Model 6), being part of a workless household, having parents with low level qualifications who spoke a language other than English, living in crowded, rented housing in a deprived area and a poor home learning environment all significantly predicted delayed school readiness. In terms of a child's own characteristics, having low birthweight, not being breastfed and having hyperactivity and peer problems all significantly increased the likelihood of having delayed school readiness; being female, white and the first-born child significantly decreased the likelihood of being delayed. (See Appendix Table A3 for full results.)

Predicting early failure at school entrance

We found being 'delayed' in the Bracken school readiness assessment (at age 3) is a very strong predictor of not reaching the expected levels in EYFS CLL and MD. Importantly, this association is barely attenuated by the wide range of family background and individual characteristics included in the different models. The predicted probability of not reaching the expected levels in EYFS CLL and MD by Bracken school readiness assessment in each of the 6 models is detailed in Figure 1.

Figure 1 shows that in Model 1, 53% of children assessed with delayed school readiness at age 3 go on to be below the expected level in CLL and MD at age 5 compared to 17% of children with average/advanced school readiness. Even when all other family background, home learning and individual characteristics are taken into account (Model 6), the 'gap' in proportions assessed below the expected level remains large: 38% to 14% respectively. (See Appendix Table A4 for full results.)

Figure 1: % not at expected level in CLL and MD assessment: predicted probabilities by Bracken School Readiness at age 3 after adjustment family background, early education support and individual characteristics: Model 1 to Model 6
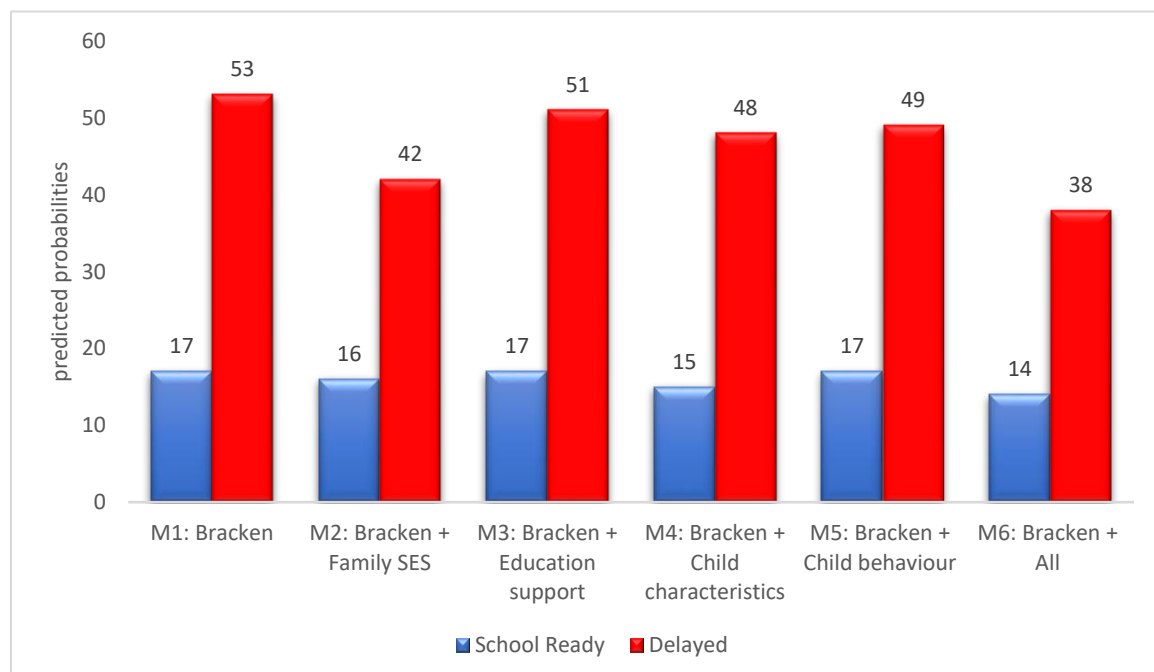
What other measures are important?

Concentrating on results from Model 6, after adjusting for all family environment and individual characteristics, we found that children with increased odds of being

assessed below expected levels at age 5 (ordered broadly from largest to smallest within each set of characteristics) were more likely to have less educated parents, to live in a rented, overcrowded home, to be part of a workless household, to have a parent with poor or fair general health and to have a poor home learning environment. In terms of individual characteristics, children assessed below expected levels at age 5 were more likely to be younger in the school year (particularly summer born), to have hyperactivity behaviour problems (at age 3), to have been born with a low birthweight and to not have been breastfed. They were also more likely to be male.

Predicting outcomes in GCSE English language and maths

Similarly, our regression analysis here allows us to identify the predictive power of early assessments on GCSE examinations and how this association is attenuated by the wide range of family background and individual characteristics included in the modelling.

The Bracken assessment of school readiness at age 3 and the EYFS teacher assessment at age 5 both remain as a significant predictor of whether children will go on to attain a 'standard pass' or higher grade in their English language and maths GCSE examinations 11 years later after adjusting for a host of family and individual characteristics. As displayed in Figure 2, Model 1 shows that 25% of those with delayed school readiness and 35% of those below the expected level in EYFS CLL and MD did not gain a grade 4 or higher pass in GCSE English language and maths, compared to 15% and 12% respectively for those who were 'school ready' at 3 and assessed at the expected level at age 5. Even when all other covariates were included in the modelling (Model 6), 19% of those who were delayed at 3 and 27% assessed as below the expected level in CLL and MD at age 5 were predicted not to achieve a grade 4 pass or above in their English language and maths GCSE at age 16. This compares with 13% who were school ready and 11% of children who were assessed as at or above the expected levels in CLL and MD at age 5. (See Appendix Table A5 for full results.)

What other measures are important?

In addition to the Bracken school readiness and the EYFS measures, we found that children with increased odds of not achieving a grade 4 or higher in English language and maths were more likely (again ordered broadly from largest to smallest within each set of characteristics) to have less educated parents – specifically, compared to teenagers who have a degree educated parent, teenagers with parents who have no or very low levels qualifications have more than twice the odds of not gaining a grade 4+ in both maths and English – and to be in a rented home where

only English was spoken. As with earlier assessments, the teenagers were also more likely to be male, to have had hyperactivity behaviour problems in early childhood, to have been born with low birthweight, to have not been breastfed and to have had a longstanding illness in early childhood.

Whereas earlier education disadvantage was associated with being non-white (BAME) and living in a home where an additional language other than English was spoken, not attaining a grade 4 or higher GCSE in English language and maths was associated with being White and only English being spoken in the home.

Figure 2 % not attaining a grade 4+ in English language and maths GCSE: predicted probabilities by Bracken School Readiness at age 3 and EYFS CLL and MD assessment at age 5 after adjustment for individual characteristics and family background: Model 1 to Model 6
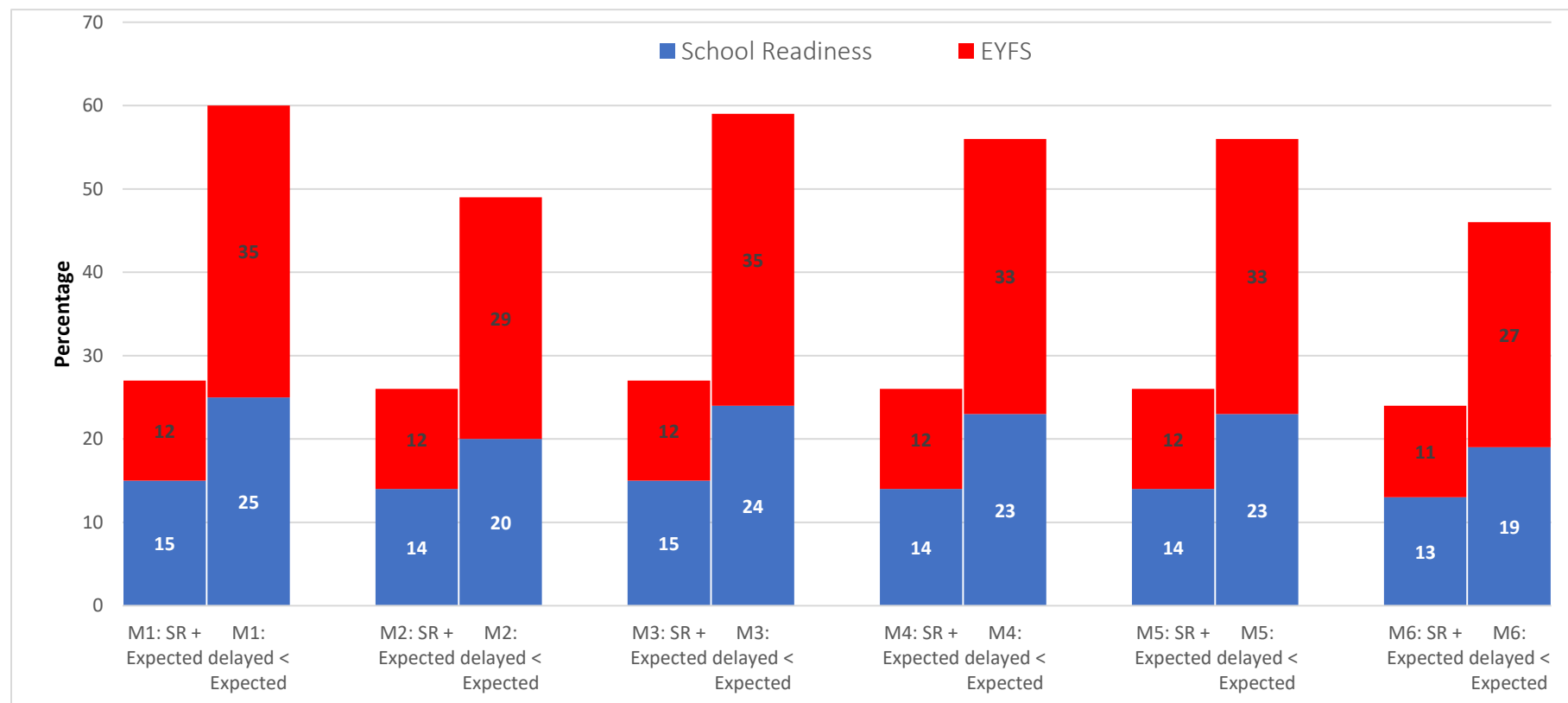
Note: Model 1: earlier education performance; Model 2: + early family socio economic environment; Model 3: + early family education support; Model 4: + child demographic characteristics; Model 5: + child behaviour problems; Model 6: + all measures. The predicted probabilities are obtained from a single logistic regression, thus it is important to acknowledge that the estimates (and therefore probabilities) for the school readiness are conditional on the EYFS which in terms of temporal ordering are on the causal path to GCSE results, being completed two years after the Bracken school readiness assessment.

Discussion

In this analysis we have focused on the 1 in 5 teenagers in England who fail to achieve the standard pass rates in both English language and maths GCSEs. This 'forgotten fifth' of pupils are associated with low achievement more generally and poor prospects in later life. We have shown that the warning signals of this educational failure emerge in the early years: these teenagers are very likely to have entered school not being 'school ready' at age 3 and to then be assessed by their teacher at age 5 as being below the expected standard in the EYFS in both Communication, Language and Literacy and Maths Development. A half of pupils who 'fail' in these two crucial examinations at age 16 were judged by teachers to be below the expected standards at age 5.

After controlling for a plethora of other family and individual characteristics, 38% of children assessed as delayed in the Bracken school readiness assessment at age 3 are predicted to be below the expected EYFS CLL and MD levels at age 5. This compares to 14% who were 'school ready' at age 3. Moving forward to GCSE attainment, school readiness at age 3 and teacher assessment in the EYFS CLL and MD remain significant predictors of failure to attain the expected standard of a grade 4 in GCSE English language and maths at the end of secondary schooling after controlling for other characteristics. Our analysis adds to a large literature pointing to the need for educational support before schooling begins in the critical early formative years (Heckman, 2006; Heckman & Masterov,2007; Heckman & Mosso, 2014), whilst also profiling the established 'tell-tale' characteristics of children who are more likely to fall behind during the earliest stages of schooling and eventually leave school without the basic literacy and number skills needed to function and prosper in life after school.

Highlighting risk factors is only useful however if they elicit a substantive response to alter the educational trajectories of pupils so that they are helped to get back on course. Diagnosing and responding to children's needs in the first years of schooling was one of the explicit aims of the EYFS when it was introduced for this cohort of children. Yet we find that all else being equal, children assessed by teachers as below expected levels at this stage are still far more likely to fail to achieve a grade 4 pass or above in their English language and maths GCSEs at age 16. This raises questions about the effectiveness of this early identification and diagnosis of pupils falling behind in the classroom, and the strategies put in place to enable pupils to develop their early literacy and number skills.

At the Conservative party conference in October 2021, the then education secretary Nadhim Zahawi vowed to "tackle innumeracy and illiteracy" as part of the

Government's schools white paper in 2022.[3] In the March 2022 schools white paper (England) (Roberts, 2022), the Government unveiled a new goal of ensuring that 90% of children leaving primary school would reach expected standards in reading, writing and maths by 2030. These would be a significant increase on the 65% of children in England meeting all three of these benchmarks in 2019. For GCSEs, the Government also outlined a target for the average grade in English language and maths GCSE to rise to 5, a 'strong pass', also by 2030, from the 4.5 in 2019, the last year in which exams were held. The problem with using averages in this way may mean that it may not explicitly address the needs of the long tail of students who currently fail to reach the grade 4 benchmark. In our view this neglected group should be the policy priority.

The Minister's remarks echoed similar promises uttered by politicians in the past. Yet despite countless policy initiatives over the years, successive Governments have failed to fulfil these lofty ambitions. It is difficult to assess the impact of a set of policy reforms on one cohort of children, but our analysis concurs with previous assessments and suggests that little changed in the proportions of children leaving school without basic literacy and number skills in England for a generation of children who arguably benefitted from more education reforms than any cohort in recent history. Reviewing Labour's record on education from 1997 to 2010, Lupton and Obolenskaya (2013) for example conclude that while attainment levels overall increased and socioeconomic gaps were reduced over the period, a persistent minority of young people remained disaffected with school, achieving little and facing very poor post-school prospects. Similarly, an Ofsted review in 2010 found that "the rising tide of educational change was still not lifting all the boats" (Ofsted, 2013).

Profiling the early life circumstances of teenagers who leave school without good grade GCSE passes indicating a lack of basic literacy and number skills proficiency may provide some indications of why policy efforts have failed to have the impact policy makers would have envisaged. Our analysis of the Millennium cohort reveals a series of consistent family and individual characteristics from children's early lives that are highly predictive of not being 'school ready' in the pre-school years and falling below expected literacy and number skill levels at the start of primary and end of secondary education stages. These can be thought of as risk factors that may place children on a trajectory of educational failure. In line with the vast literature on children's educational outcomes, we find growing up in more difficult socio-economic circumstances is related to not reaching educational standards at age 3, 5 and 16. In particular we find that having parents with low levels of parental education (those with no qualifications or 'good grade' GCSE passes or equivalent) is a strong measure predicting failure at every age. Living in an overcrowded and workless household with a parent in poor health is also related to not reaching expected

---

[3] See: https://schoolsweek.co.uk/zahawi-to-tackle-innumeracy-and-illiteracy-with-white-paper-next-year/

standards in the early years. Importantly, so is the quality of the area that families live in, with more failing children living in areas of higher deprivation which links directly to the quality of the local schools – the higher the property prices are in an area, the better the local primary and secondary school is (DfE, 2017).

In terms of individual characteristics, children born with low birthweight who are not breastfed are more likely to have delayed school readiness at age 3 and to fail at the first educational hurdle at age 5 when teachers assess pupils at the beginning of their schooling to be below the expected standards. Being male and having hyperactivity problems at age 3 is also associated with education disadvantage at all three age points. Summer born children are almost five times as likely as otherwise similar autumn born children to be assessed below expected literacy and number skills levels at age 5, with being summer born remaining associated with poor GCSE performance although not when other family and individual characteristics are taken into account. This (somewhat) contradicts earlier studies that have found a small summer born effect remaining at age 16 for earlier cohorts of pupils (see Crawford et al, 2013).

Our results confirm the need for more effective education policies to help younger and poorer behaved children in the early years – and point to the need for combined health and education initiatives that could for example draw attention to the importance of breastfeeding for the long-term education success of children which would be particularly important for low-birth babies.

The importance of parental education and the home learning environment

Parental education is an indicator of what parents do, as well as who they are (Sylva et al., 2004). While one response to our findings might be to commit to long term efforts to improve educational levels of future generations of parents, they also point to the possibility of changing the behaviours of current parents to improve the home learning environment. A large literature shows that, on average, parents with higher education levels engage in more cognitive stimulation, interact with more warmth and consistency, and use harsh discipline less often with young children than do parents with lower education levels (see for example Rowe, 2018; Kalil, Ryan, & Corey, 2012; Linver, Brooks-Gunn & Kohen, 2002).

A positive early home-learning environment is derived in the MCS by finding out whether children experience a range of educational activities including being read to, playing with numbers, and being taught letters and numbers, and hearing songs/poems/rhymes. In line with prior research, we find that children in homes lacking these activities are significantly more likely to not be school ready at age 3 and to be assessed by teachers at age 5 as being below expected levels in the EYFS, which are in turn strongly related to GCSE performance.

Our findings resonate with previous studies that have investigated past generations of lower achievers in English language and maths. Rashid and Brooks (2010) for example conclude "that children who fall behind in the early stages must be identified and given targeted catch-up programmes immediately" and that "family literacy and numeracy programmes can make a contribution to preventing early failure". Carroll and Gill (2018) meanwhile conclude that "a better job could be done in identifying and supporting children who are behind in reading and writing" while "early-years provision has to do better in reaching the most disadvantaged, particularly to help improve parenting and early learning".

In our future work we plan to review policy efforts in this area more extensively, but our analysis already points to several policy implications likely to be important in any policy efforts to improve literacy and numeracy levels. These include:

Identifying, diagnosing and responding to low literacy and number skills in children at every stage of their education

We now have clear evidence-informed approaches (see Education Endowment Foundation reviews, 2018, 2020) to improve literacy and numeracy in classrooms. Yet these will continue to be ineffective for many children without a comprehensive and effective strategy to identify, diagnose and respond appropriately to children falling behind at every life and education stage from birth onwards throughout their schooling journeys. Our results suggest that the Early Years Foundation Stage (EYFS) profile assessments intended to diagnose pupil needs at age 5 were either ineffective, or that teachers lack the time and resources to respond to the needs identified.

Investing in high quality early years education

Our analysis confirms the need to prioritise education investment to create a comprehensive pre-school education and support service for disadvantaged and at-risk pupils. Crucially this must be of consistent high quality, overseen by well qualified staff (See for example Sylva et al.,2004).

Improving parental engagement and the home learning environment

In our view the Government could do more to work with schools to ensure they develop clear school parent engagement plans. These would encourage schools to better understand their communities, provide information and advice on improving the home learning environment, and make themselves inclusive to all parents. Recent trials outside the UK have produced promising results using low-cost behavioural tools to help motivate parents to read with younger children (see for example Barone et al., 2020; Mayer et al., 2019). A dedicated programme could be

initiated to develop evidence-informed approaches to help parents change habits in the home environment to support learning.

Recognising summer born disadvantage

The systematic differences by relative age of pupils in early years assessments highlights the issue of summer born disadvantage (see Elliot Major & Higgins, 2019). Summer-born children are far less mature than their older peers during the early stages of schooling, but this should not be mistaken for being lower achievers. They should be treated like any other vulnerable groups of pupils – with progress monitored in relation to what is expected for their age.

Considering reforms to examinations

The continuing underperformance of so many pupils also calls into question school examinations themselves. An alternative approach would be to assess children against a basic threshold of key literacy and number skills – alongside higher academic grades. Maths and English language GCSEs could each be split into two separate qualifications: a compulsory test examining basic number and literacy skills, and a separate exam for pupils pursuing more academic study. Currently, pupils achieving a grade 3 or lower in English and/or mathematics at GCSE are automatically enrolled to retake the exams again. Success rates however remain predictably low (Velthuis et al., 2018).

Strengths and Limitations

Our conclusions and analyses are based on correlations of data, suggestive of relationships rather than offering strict causal evidence. Alternative explanations could be offered for the unchanging patterns in pupils' achievement over successive years. For example, it could be that the school system simply worked harder to counter-balance wider inequalities in society during the period we consider (although there was no significant change in income inequality for the MCS generation). It is difficult to assess the specific causal impact of one policy initiative, let alone a range of policy reforms for one generation.

Our data meanwhile is derived from an observational longitudinal study, so bias due to unmeasured confounding cannot be ruled out. As in any longitudinal survey, missing data due to attrition is unavoidable. We employed multiple imputation, augmenting our models with auxiliary variables in the imputation phase to maximise the plausibility of the MAR assumption and restore sample representativeness, but bias due to a non-ignorable missing data-generating mechanism cannot be ruled out.

A further limitation is that our findings strictly can only be generalised to those born in Britain in 2000/2, or close to these years.

References

ASCL (2019). The Forgotten Third. Final report of the Commission of Inquiry. Available at: https://www.ascl.org.uk/ASCL/media/ASCL/Our view/Campaigns/The-Forgotten-Third_full-report.pdf

Barbaresi, W.J., Katusic, S.K., Colligan, R.C., Weaver, A.L. & Jacobsen, S.J. (2007) Modifiers of long-term school outcomes for children with attention-deficit/hyperactivity disorder: does treatment with stimulant medication make a difference? Results from a population-based study, *Journal of Developmental and Behavioral Pediatrics*, 28(4), 274–287. DOI: 10.1097/DBP.0b013e3180cabc28

Barone, C., Fougère, D., Martel, K. (2020) Reading Aloud to Children, Social Inequalities, and Vocabulary Development: Evidence from a Randomized Controlled Trial IZA DP No. 13458

Bell, D. & Blanchflower, D. (2010). UK Unemployment in the Great Recession. *National Institute Economic Review*, 214 (1): pp. R3–R25. doi: 10.1177/0027950110389755.

BIS (2012). The 2011 Skills for Life Survey: A Survey of Literacy, Numeracy and ICT. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/36000/12-p168-2011-skills-for-life-survey.pdf - page=294

Bracken, B. (1998). Bracken Basic Concept Scale Revised: Examiner's Manual. London: The Psychological Corporation.

Bracken, B. & Crawford, E. Basic Concepts in Early Childhood Educational Standards: A 50-State Review. *Early Childhood Educ J* **37,** 421–430 (2010). https://doi.org/10.1007/s10643-009-0363-7

Blanden, J. & Macmillan, L. (2016). Educational Inequality, Educational Expansion and Intergenerational Mobility. *Journal of Social Policy*, 45(4), 589–614. doi:10.1017/S004727941600026X

Blanden, J., Gregg, P. & Macmillan, L. (2007). Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education, *Economic Journal*, 117: C43-C60. https://doi.org/10.1111/j.1468-0297.2007.02034.x

Bradbury B., Corak, M., Waldfogel, J. & Washbrook, E. (2012) Inequality During the Early Years: Child Outcomes and Readiness to Learn in Australia, Canada, United Kingdom and United States, in Ermisch, Jantti, M. and Smeeting, T. (eds.) From

Parents to Children: The Intergenerational Transmission of Advantage, New York: Russell Sage Foundation.

Bradbury B., M. Corak, J. Waldfogel & E. Washbrook (2015). Too Many Children Left Behind: The U.S. Achievement Gap in Comparative Perspective, Russell Sage Foundation.

Breen, M. (1985). Authenticity in the Language Classroom. *Applied Linguistics*, 6 (1), 60–70. https://doi.org/10.1093/applin/6.1.60
Bynner, J. & Parsons, S. (1997): It doesn't get any better. The Basic Skills Agency: London.

Bynner, J. & Parsons, S. (2002). Social Exclusion and the Transition from School to Work: The Case of Young People Not in Education, Employment or Training. *Journal of Vocational Behavior*, 60, p289-309

Bynner, J. & Parsons, S. (2006) New Light on Literacy and Numeracy: Results of the literacy and numeracy assessment in the age 34 follow-up of the 1970 British Cohort Study (BCS70). London: National Research and Development Centre for adult literacy and numeracy. http://dera.ioe.ac.uk/22309/1/doc_3186.pdf

Carroll, M. & Gill, T. (2018). Uptake of GCSE subjects 2017. Statistics Report Series No. 120. Cambridge, UK: Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/518813-uptake-of-gcse-subjects-2017.pdf
Case, A., Fertig, A. & Paxson, C. (2006) The lasting impact of childhood health and circumstance, *Journal of Health Economics*, 24, 365–389.
DOI: 10.1016/j.jhealeco.2004.09.008

Cassen, R., & Kingdon, G. (2007). Tackling low educational achievement. Joseph Rowntree Foundation. https://www.jrf.org.uk/sites/default/files/jrf/migrated/files/2063-education-schools-achievement.pdf

Connelly, R. (2013). Millennium Cohort Study Data Note: Interpreting test scores. CLS Data Note 2013/1. London, Centre for Longitudinal Studies. https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-data-note-20131-Test-Scores-Roxanne-Connelly.pdf

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
Coghlan, M., Bergeron, C., White, K., Sharp, C., Morris, M. & Wilson, R. (2010). Narrowing the gap in outcomes for young children through effective practices in the early years. London: Centre for Excellence and Outcomes in Children and Young

People's Services. Available at:
https://www.basw.co.uk/system/files/resources/basw_123326-4_0.pdf

Crawford, C., Greaves, E., Jin, W., Swaffield, J., and Vignoles, A. (2011). The impact of the minimum wage regime on the education and labour market choices of young people. Report to the Low Pay Commission. London: Low Pay Commission. https://ifs.org.uk/publications/6127

Crawford, C., Dearden, L. and Greaves, E. (2013), 'When you are born matters: Evidence for England', IFS Reports (No. R80). London: Institute for Fiscal Studies. https://ifs.org.uk/comms/r80.pdf

Currie, J. & Hyson, R. (1999). Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight, *American Economic Review*, 89 (2): 245-250. DOI: 10.1257/aer.89.2.245

de la Rochebrochard, E. (2012). The home learning environment as measured at age 3. Millennium Cohort Study Data Note 1. https://cls.ucl.ac.uk/wp-content/uploads/2017/07/FINAL-The-home-learning-environment-as-measured-in-the-MCS-at-age-3-MCS-data-note-1.pdf

Department of Education (1985) *General Certificate of Secondary Education: A general introduction*, London: Her Majesty's Stationery Office.

DfE. (2017). House prices and schools: do houses close to the best- performing schools cost more? Ad hoc research note March 2017. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/600623/House_prices_and_schools.pdf

DfE (2019). Early years foundation stage profile results in England, 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/839934/EYFSP_2019_Main_Text_Oct.pdf

DfE (2020). Academic Year 2019/20: Key Stage 4 Performance (revised). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/863815/2019_KS4_revised_text.pdf

DfE. (2022a). Academic Year 2020/21: Key stage 4 performance. https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-4-performance-revised/2020-21

DfE. (2022b). Early years foundation stage profile, 2022 handbook. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1024319/Early_years_foundation_stage_profile_handbook_2022.pdf

Dickerson, A., McDool, E. & Morris, D. (2022) Post-compulsory education pathways and labour market outcomes, *Education Economics*. https://doi.org/10.1080/09645292.2022.2068137

Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H., . . . Duckworth, K. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446. doi: https://doi.org/10.1037/0012-1649.43.6.1428

Education Endowment Foundation. (2018). 'Preparing for literacy: Improving communication, language and literacy in the early years: Guidance report', https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/literacy-early-years

Education Endowment Foundation. (2020). Improving Mathematics in the Early Years and Key Stage 1, https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/early-maths

Elliot Major, L, and Machin S. (2018). Social Mobility and Its Enemies. London: Pelican Books.

Elliot Major, L. and Higgins, S. (2019) What Works? Research and Evidence for Successful Teaching, London: Bloomsbury Education.

Feinstein, L. (2003) Inequality in the early cognitive development of British children in the 1970 cohort, Economica, 70, 73–97. https://doi.org/10.1111/1468-0335.t01-1-00272

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38(5), 581-586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 1337- 1345. https://doi.org/10.1097/00004583-200111000-00015

Goux, D. & Maurin, E. (2003) The effect of overcrowded housing on children's performance at school (Paris, INSEE).

Gov.UK (2022a). 16 to 19 funding: maths and English condition of funding. Available at: https://www.gov.uk/guidance/16-to-19-funding-maths-and-english-condition-of-funding

Gov.UK (2022b). School leaving age. https://www.gov.uk/know-when-you-can-leave-school

Greening, J. (2017) Letter to Neil Carmichael MP, 28th March. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/603594/ESC_letter.pdf.

Gregg, P. and Macmillan, L. (2010). Family income, education and cognitive ability in the next generation: exploring income gradients in education and test scores for current cohorts of youth. Longitudinal and Life Course Studies, 1(3), 259-280. http://dx.doi.org/10.14301/llcs.v1i3.95

Gutman, L. & Vorhaus, J. (2012). The impact of pupil behaviour and wellbeing on educational outcomes. (DfE Research Report DFE-RR253). London: Department for Education. https://dera.ioe.ac.uk//16093/

Halsey, A.H., Heath, A.F. & Ridge, J.M. (1980) Origins and destinations: Family, class, and education in modern Britain (Oxford, Oxford University Press).

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902. doi:doi:10.1126/science.1128898

Heckman, J. & Masterov, D. (2007). The Productivity Argument for Investing in Young Children. *Review of Agricultural Economics*, 29 (3), 446-493. DOI:10.1111/j.1467-9353.2007.00359.x

Heckman, J. & Mosso, S. (2014) The Economics of Human Development and Social Mobility, Annual Review of Economics, 6(1), 689-733. https://doi.org/10.1146/annurev-economics-080213-040753

Jin, W., Muriel, A., & Sibieta, L. (2011). *Subject and course choices at ages 14 and 16 amongst young people in England: Insights from behavioural economics* (DfE Research Report DFE-RR160). London: Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/182677/DFE-RR160.pdf

Joshi, H. & Fitzsimons, E. (2016). The Millennium Cohort Study: the making of a multi-purpose resource for social science and policy. Longitudinal and Life Course Studies, 7(4), 409-430. http://dx.doi.org/10.14301/llcs.v7i4.410
Kalil, A., Ryan, R. & Corey, M. (2012). Diverging destinies: maternal education and the developmental gradient in time with children. Demography, 49(4), 1361-83. DOI: 10.1007/s13524-012-0129-5.

Kiernan, K. E. & Huerta, M. C. (2008). Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood.

The British Journal of Sociology, 59 (4), 783- 806. https://doi.org/10.1111/j.1468-4446.2008.00219.x

Kingdon, G. & Cassen, R. (2010). Ethnicity and low achievement in English schools. British Educational Research Journal, 36: 403-431. https://doi.org/10.1080/01411920902989185

Kuczera, M., Field, S. & Windisch, H. (2016) 'Building Skills for All: A Review of England', OECD Report, https://www.oecd.org/unitedkingdom/building-skills-for-all-review-of-england.pdf

Linver, M., Brooks-Gunn, J., Kohen, D. (2002). Family processes as pathways from income to young children's development. *Developmental Psychology*, 38 (5): 719-34. PMID: 12220050.

Little, R. & Rubin, D. (2014). Statistical Analysis with Missing Data. 2nd Edition, John Wiley & Sons, Hoboken.

Lupton, R. and Obolenskaya, P. (2013) Labour's record on education: policy, spending and outcomes 1997-2010. Social policy in a cold climate working paper (SPCCWP03). Centre for Analysis of Social Exclusion (CASE), London, UK. http://eprints.lse.ac.uk/51214/1/__libfile_REPOSITORY_Content_CASE_CASE%20Social%20policy%20in%20a%20cold%20climate%20series_Working%20papers_spccwp03_full%20text.pdf

Lupton, R., Thomson, S., Velthuis, S. and Unwin, L. (2021). Moving on from initial GCSE 'failure': Post-16 transitions for 'lower attainers' and why the English education system must do better. Available at: https://www.nuffieldfoundation.org/wp-content/uploads/2021/02/Post16-transitions-for-lower-attainers-Final-report.pdf

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. https://doi.org/10.1016/j.jpubeco.2020.104224

Martin, R. and H. Hodgson, with A. Maloney and I. Rayner (2014) 'Cost of outcomes associated with low levels of adult numeracy in the UK', Pro Bono Economics Report for National Numeracy, http://www.probonoeconomics.com/sites/probonoeconomics.com/files/files/reports/PBE National Numeracy costs report 11Mar.pdf

Mayer, S., Kalil, A., Oreopoulos, & Gallegos, S. (2019) Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together Intervention, *Journal of Human Resources*, 54(4), 900-925. DOI: 10.3368/jhr.54.4.0617.8835R

Mobley M., Emerson C., Goddard Y., Goodwin S. and Letch R. (1986) *All about GCSE*, London: Heinemann.

Mostafa, T. & Wiggins, R. (2015). The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies. Longitudinal and Life Course Studies, 6(2), pp.131-146. http://dx.doi.org/10.14301/llcs.v6i2.312

Mostafa, T., Narayanan, M., Pongiglione, B., Dodgeon, B., Goodman, A., Silverwood, R., & Ploubidis, G. (2020). Improving the plausibility of the missing at random assumption in the 1958 British birth cohort: A pragmatic data driven approach. London: UCL Centre for Longitudinal Studies Working Paper Series. https://cls.ucl.ac.uk/wp-content/uploads/2017/02/CLS-working-paper-2020-6-Improving-the-plausibility-of-the-missing-at-random-assumption-in-the-1958-British-birth-cohort-1.pdf

Melhuish, E., Phan, M., Sylva, K., Sammons, P., Siraj, I. & Taggart, B. (2008). Effects of the Home Learning Environment and Preschool Center Experience upon Literacy and Numeracy Development in Early Primary School. *Journal of Social Issues* 64(1):95 – 114. https://doi.org/10.1111/j.1540-4560.2008.00550.x

National Literacy Trust. (2017). What do adult literacy levels mean? Available at: https://literacytrust.org.uk/parents-and-families/adult-literacy/what-do-adult-literacy-levels-mean/

National Numeracy. (2022). What do adult numeracy 'levels' mean? Available at: https://www.nationalnumeracy.org.uk/what-numeracy/what-do-adult-numeracy-levels-mean

Newton, P. (2021). Demythologising A level Exam Standards, *Research Papers in Education*, DOI: 10.1080/02671522.2020.1870543

North J. (1987) *The GCSE: An Examination*, London: The Clarridge Press.
OECD (2016). British youngsters 'most illiterate' in developed world, says OECD. Available at: https://www.trainingjournal.com/articles/news/british-youngsters-most-illiterate-developed-world-says-oecd.
Office of the Deputy Prime Minister. (2004). The impact of overcrowding on health and education: a review of evidence and literature. London: Office of the Deputy Prime Minister. https://dera.ioe.ac.uk/5073/1/138631.pdf

Ofqual (2018). Guide to GCSE results for England, 2018. https://www.gov.uk/government/news/guide-to-gcse-results-for-england-2018

Ofqual (2019). Guide to GCSE results for England, 2019. Available at:
https://www.gov.uk/government/news/guide-to-gcse-results-for-england-2019

Ofqual (2021) National Reference Test Results Digest
https://www.gov.uk/government/publications/the-national-reference-test-in-2021/national-reference-test-results-digest-2021

Ofsted (2013) Unseen children: access and achievement 20 years on, Evidence report
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/379157/Unseen_20children_20-_20access_20and_20achievement_2020_20years_20on.pdf

Parsons, S. & Hallam, S. (2014). The Impact of Streaming on Attainment at Age Seven: Evidence from the Millennium Cohort Study. *Oxford Review of Education*, 40 (5), 567-589. DOI:10.1080/03054985.2014.959911

Parsons, S., Schoon, I., Rush, R. & Law, J. (2011) Long-term outcomes for children with early language problems: Beating the odds, *Children & Society*, 25(3), 202–214.
https://doi.org/10.1111/j.1099-0860.2009.00274.x

Parsons, S., Schoon, I. & Vignoles, A. (2014). Parental worklessness and children's early school achievement and progress. Longitudinal and Life Course Studies, 5(1), 19-41. http://www.llcsjournal.org/index.php/llcs/article/view/230

Parsons S., Sullivan, A., Moulton, V., Fitzsimons, E. & Ploubidis, G. (2021) The relationship between child behaviour problems at school entrance and teenage vocabulary acquisition: a comparison of two generations of British children born 30 years apart. *British Educational Research Journal*, 47 (5), 1361-1387.
http://doi.org/10.1002/berj.3732

Passaretta, Giampiero & Skopek, Jan & Huizen, Thomas. (2022). Is Social Inequality in School-Age Achievement Generated before or during Schooling? A European Perspective. European Sociological , jcac005, https://doi.org/10.1093/esr/jcac005

Plewis, I. (Ed.) (2007). Millennium Cohort Study first survey: Technical report on sampling (4th edn) (London, Centre for Longitudinal Studies).
https://cls.ucl.ac.uk/wp- content/uploads/2017/07/Technical-Report-on-Sampling-4th-Edition-August-2007.pdf

Pungello, E. P., Kainz, K., Burchinal, M., Wasik, B. H., Sparling, J. J., Ramey, C. T., & Campbell, F. A. (2010). Early educational intervention, early cumulative risk, and the early home environment as predictors of young adult outcomes within a high-risk

sample. Child Development, 81(1), 410–426. https://doi.org/10.1111/j.1467-8624.2009.01403.x

Qualifications and Curriculum Authority (2003). Foundation Stage Profile: Handbook (2003). Available at: http://doc.ukdataservice.ac.uk/doc/6847/mrdoc/pdf/foundation_stage_profile_handbook.pdf

Quigley, M., Hockley, C., Carson, C., Kelly, Y., Renfrew, M. and Sacker, A. (2012). Breastfeeding is Associated with Improved Child Cognitive Development: A Population-Based Cohort Study. *Journal of Pediatrics*, 160, 25-32. https://doi.org/10.1016/j.jpeds.2011.06.035

Ralston, K., Feng, Z., Everington, D. & Dibben, C. (2016). Do young people not in education, employment or training experience long-term occupational scarring? A longitudinal analysis over 20 years of follow-up, *Contemporary Social Science*, 11:2-3, 203-221, DOI: 10.1080/21582041.2016.1194452

Rashid, S. and Brooks, G. (2010). The levels of attainment in literacy and numeracy of 13- to 19-year-olds in England, 1948–2009. Research report. Available at: https://dera.ioe.ac.uk/21953/1/doc_4690.pdf

Roberts, N. (2022). March 2022 schools white paper (England). House of Commons Library Briefing Paper CBP 09511. https://researchbriefings.files.parliament.uk/documents/CBP-9511/CBP-9511.pdf

Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives*, *12*(2), 122-127.

Rutter, M., Tizard, J., & Whitmore, K. (1970). Education, health, and behaviour. London: Longman.

Sammons, P., Sylva, K., Melhuish,E., Siraj, I., Taggart, B., Toth, K. and Smees, R. (2014). Influences on students' GCSE attainment and progress at age 16: Effective Pre-School, Primary & Secondary Education Project (EPPSE). Research Report. Available at: https://dera.ioe.ac.uk/20875/1/RR352_-_Influences_on_Students_GCSE_Attainment_and_Progress_at_Age_16.pdf
Schoon, I. (2020). The Wellbeing of Children in the Face of Socio-Economic Deprivation and Family Instability. Revue des politiques sociales et familiales, 131-2(4), 51-65. https://www.persee.fr/doc/caf_2431-4501_2019_num_2131_2431_3359.

Silverwood, R., Narayanan, M., Dodgeon, B., & Ploubidis, G. B. (2020). Handling missing data in the National Child Development Study: User guide. London: UCL Centre for Longitudinal Studies. https://cls.ucl.ac.uk/wp-

content/uploads/2020/04/Handling-missing-data-in-the-National-Child-Development-Study-User-Guide.pdf

Smithers, A. (2014) GCSE Trends: 1988-2014. Centre for Education and Employment Research. University of Buckingham. https://www.buckingham.ac.uk/wp-content/uploads/2019/02/GCSE14AGS.pdf

Stopforth, S., Gayle, V. and Boeren, E. (2020): Parental social class and school GCSE outcomes: two decades of evidence from UK household panel surveys, *Contemporary Social Science*, 16 (3), 309-324. Available at: https://doi.org/10.1080/21582041.2020.1792967

Stopforth, S. & Gayle, V. (2022) Parental social class and GCSE attainment: Re-reading the role of 'cultural capital', *British Journal of Sociology of Education*, 43 (5), 680-699, DOI: 10.1080/01425692.2022.2045185

Strand, S. (2014) School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11, Oxford Review of Education, 40:2, 223-245. DOI: 10.1080/03054985.2014.891980

Sullivan, A., Ketende, S., & Joshi, H. (2013). Social class and inequalities in early cognitive scores. Sociology, 47,1187-1206. http://dx.doi.org/10.1177/0038038512461861

Sutherland, A., Ilie, S. & Vignoles, A. (2015) Factors associated with achievement: key stage 4. DfE Research report. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/473673/RR407_-_Factors_associated_with_achievement_-_key_stage_4.pdf

Sykes, E. D. A., Bell, J. F. and Rodeiro, C. V. (2009), 'Birthdate effects: A review of the literature from 1990-on'. Cambridge: Cambridge Assessment. http://www.cambridgeassessment.org.uk/Images/109784-birthdate-effects-a-review-of-the- literature-from-1990-on.pdf

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004) The Effective Provision of Pre-School Education (EPPE) Project: Findings from Pre-school to end of Key Stage 1. https://dera.ioe.ac.uk/18189/2/SSU-SF-2004-01.pdf

Taylor, C., Christensen, D., Lawrence, D., Mitrou, F. & Zubrick, S. (2013) Risk factors for children's receptive vocabulary development from four to eight years in the longitudinal study of Australian children, *PlosOne*, 8(11). https://doi.org/10.1371/annotation/a730446c-0150-4079-98ea-95b12e1e3c28

Thompson, M. N., Dahling, J. J., Chin, M. Y., & Melloy, R. C. (2017). Integrating Job Loss, Unemployment, and Reemployment With Social Cognitive Career Theory. *Journal of Career Assessment, 25*(1), 40-57. https://doi.org/10.1177/1069072716657534

University of London, Institute of Education, Centre for Longitudinal Studies. (2021). Millennium Cohort Study: Seventh Survey, 2018. [data collection]. 2nd Edition. UK Data Service. SN: 8682, http://doi.org/10.5255/UKDA-SN-8682-2

University of London, Institute of Education, Centre for Longitudinal Studies. (2022a). Millennium Cohort Study: First Survey, 2001-2003. [data collection]. 14th Edition. UK Data Service. SN: 4683, http://doi.org/10.5255/UKDA-SN-4683-6

University of London, Institute of Education, Centre for Longitudinal Studies. (2022b). Millennium Cohort Study: Second Survey, 2003-2005. [data collection]. 11th Edition. UK Data Service. SN: 5350, http://doi.org/10.5255/UKDA-SN-5350-6

University of London, Institute of Education, Centre for Longitudinal Studies. (2022c). Millennium Cohort Study: Third Survey, 2006. [data collection]. 9th Edition. UK Data Service. SN: 5795, http://doi.org/10.5255/UKDA-SN-5795-6

Velthuis, S., Lupton, R., Thomson, S. and Unwin, L. (2018). The characteristics and post-16 transitions of GCSE 'lower attainers. Available at: http://hummedia.manchester.ac.uk/institutes/mui/igau/lower_attainers_working_paper_oct2018.pdf

Vignoles, A. (2016) 'What is the Economic Value of Literacy and Numeracy? Basic Skills in Literacy and Numeracy are Essential For Success in the Labour Market', IZA World of Labour 2016: 229, http://wol.iza.org/articles/what-is-economic-value-of-literacy-and-numeracy-1.pdf

Washbrook, E., Propper, C. & Sayal, K. (2013) Pre-school hyperactivity/attention problems and educational outcomes in adolescence: Prospective longitudinal study, The British Journal of Psychiatry, 203, 265–271. DOI: 10.1192/bjp.bp.112.123562
White, I.R., Royston, P., & Wood, A.M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. Stat Med, 30(4): 377-399. https://doi.org/10.1002/sim.4067

Whiteside, K., Gooch, D. & Norbury, C. (2017). English Language Proficiency and Early School Attainment Among Children Learning English as an Additional Language. Child Development, 88 (3), 812–827. https://doi.org/10.1111/cdev.12615
Youth In Mind (2016). Scoring the SDQ. https://www.sdqinfo.org/py/sdqinfo/c0.py

Appendices

Appendix 1

Early Year Foundation Stage (EYFS)

The Curriculum guidance for the foundation stage sets out six areas of learning covering children's physical, intellectual, emotional and social development. All six of these areas of learning are included in the Profile:
- personal, social and emotional development
- communication, language and literacy
- mathematical development
- knowledge and understanding of the world
- physical development
- creative development

Communication, Language and Literacy (CLL): combines assessment in four separate early learning goal (ELG) descriptors:

Language for communication and thinking

- Listens and responds
- Initiates communication with others, displaying greater confidence in more informal contexts
- Talks activities through, reflecting on and modifying actions
- Listens with enjoyment to stories, songs, rhymes and poems, sustains attentive listening and responds with relevant comments, questions or actions
- Uses language to imagine and recreate roles and experiences
- Interacts with others in a variety of contexts, negotiating plans and activities and taking turns in conversation
- Uses talk to organise, sequence and clarify thinking, ideas, feelings and events, exploring the meanings and sounds of new words
- Speaks clearly with confidence and control, showing awareness of the listener
- Talks and listens confidently and with control, consistently showing awareness of the listener by including relevant detail. Uses language to work out and clarify ideas, showing control of a range of appropriate vocabulary
- Linking sounds and letters
- Joins in with rhyming and rhythmic activities
- Shows an awareness of rhyme and alliteration
- Links some sounds to letters
- Links sounds to letters, naming and sounding letters of the alphabet
- Hears and says sounds in words

- o Blends sounds in words
- o Uses phonic knowledge to read simple regular words
- o Attempts to read more complex words, using phonic knowledge
- o Uses knowledge of letters, sounds and words when reading and writing independently

Reading

- o Is developing an interest in books.
- o Knows that print conveys meaning.
- o Recognises a few familiar words.
- o Knows that, in English, print is read from left to right and top to bottom.
- o Shows an understanding of the elements of stories, such as main character, sequence of events and openings.
- o Reads a range of familiar and common words and simple sentences independently.
- o Retells narratives in the correct sequence, drawing on language patterns of stories.
- o Shows an understanding of how information can be found in non-fiction texts to answer questions about where, who, why and how.
- o Reads books of own choice with some fluency and accuracy.

Writing

- o Experiments with mark-making, sometimes ascribing meaning to the marks.
- o Uses some clearly identifiable letters to communicate meaning.
- o Represents some sounds correctly in writing.
- o Writes own name and other words from memory.
- o Holds a pencil and uses it effectively to form recognisable letters, most of which are correctly formed.
- o Attempts writing for a variety of purposes, using features of different forms.
- o Uses phonic knowledge to write simple regular words and make phonetically plausible attempts at more complex words.
- o Begins to form captions and simple sentences, sometimes using punctuation.
- o Communicates meaning through phrases and simple sentences with some consistency in punctuating sentences.

Maths Development (MD): combines assessment in three separate early learning goal (ELG) descriptors:

Numbers as labels and for counting

- o Says some number names in familiar contexts, such as nursery rhymes.
- o Counts reliably up to three everyday objects.

- o Counts reliably up to six everyday objects.
- o Says number names in order
- o Recognises numerals 1 to 9.
- o Counts reliably up to 10 everyday objects.
- o Orders numbers, up to 10.
- o Uses developing mathematical ideas and methods to solve practical problems.
- o Recognises, counts, orders, writes and uses numbers up to 20.

Calculating

- o Responds to the vocabulary involved in addition and subtraction in rhymes and games.
- o Recognises differences in quantity when comparing sets of objects.
- o Finds one more or one less from a group of up to five objects.
- o Relates addition to combining two groups.
- o Relates subtraction to taking away.
- o In practical activities and discussion, begins to use the vocabulary involved in adding and subtracting.
- o Finds one more or one less than a number from 1 to 10.
- o Uses developing mathematical ideas and methods to solve practical problems.
- o Uses a range of strategies for addition and subtraction, including some mental recall of number bonds.

Shape, space and measures

- o Experiments with a range of objects and materials showing some mathematical awareness.
- o Sorts or matches objects and talks about sorting.
- o Describes shapes in simple models, pictures and patterns.
- o Talks about, recognises and recreates simple patterns.
- o Uses everyday words to describe position.
- o Uses language such as 'circle' or 'bigger' to describe the shape and size of solids and flat shapes
- o Uses language such as 'greater', 'smaller', 'heavier' or 'lighter' to compare quantities.
- o Uses developing mathematical ideas and methods to solve practical problems.
- o Uses mathematical language to describe solid (3D) objects and flat (2D) shapes.

Appendix tables

Table A1: Policy Reforms that the MCS cohort will have been exposed to

Year	Policy reform
1996	National Literacy Task Force: Set-up by Labour to develop a strategy to raise standards of literacy in primary schools over a five- and ten-year period. Built on work of the National Literacy Project. http://www.educationengland.org.uk/documents/literacytaskforce/implementation.html
1997	Numeracy Task Force set up in a similar vein to literacy – built on work of the National Numeracy project (Sept 1996). http://www.educationengland.org.uk/documents/literacytaskforce/implementation.html
1997	Excellence in Schools White Paper: Set out an agenda to raise standards and highlighted underachievement in Maths and English. Targets literacy and numeracy priorities in primary education. Sets targets for 75% and 80% of 11-year olds to reach expected standards for their age in maths and English respectively by 2002. In 1996 fewer than 6 in 10 achieved these levels. http://www.educationengland.org.uk/documents/wp1997/excellence-in-schools.html
1997	Guidelines for teachers developed to spend at least an hour a day on English and an hour on maths in primary schools as part of the national strategies for improving standards of literacy and numeracy. http://www.educationengland.org.uk/documents/wp1997/excellence-in-schools.html
1999	Sure Start programme launched aimed at improving the health, well-being and educational attainment of 0- to 3-year-olds in disadvantaged areas through a wide range of health, education and social services. https://researchbriefings.files.parliament.uk/documents/CBP-7257/CBP-7257.pdf
2000	Network of 'city academies' independent of local authority, effectively private schools paid for by the state. https://dera.ioe.ac.uk/3000/1/City_academies_-_schools_to_make_a_difference_(July_2000).pdf
2001	Building on Our Success Green paper. Aim to build on success at primary level in secondary schools with targets for 14-year-olds in English and Maths. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/250873/5050.pdf

Year	Policy reform
2001	Schools Achieving Success White Paper: Detailed post-election plans. By 2007 85% of 14-year-olds were to achieve Level 5 or above in English, Maths and ICT at the end of Key Stage 3. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/355105/Schools_Achieving_Success.pdf
2001	Secondary National Strategy set-up to improve English and Maths at key stage 3 https://www.nfer.ac.uk/media/1347/91018.pdf
2002	Education Act: Establishes into law targets at Key Stages 1,2, 3 & 4 http://www.educationengland.org.uk/documents/acts/2002-education-act.html#06
2003	Primary National Strategy combines the literacy and numeracy strategies. https://dera.ioe.ac.uk/4817/7/pri_excel_enjoy_strat_Redacted.pdf
2004	Making Mathematics Count: Critical Review of arrangements http://www.mathsinquiry.org.uk/report/MathsInquiryFinalReport.pdf
2005	White Paper 14-19 Education and Skills: Details plans on extending functional skills in English and Maths to GCSE level. http://www.educationengland.org.uk/documents/pdfs/2005-white-paper-14-19-education-and-skills.pdf
2010	Pupil premium introduced to improve outcomes for children on Free School meals, alongside Education Endowment Foundation evidence informed guidance to spend the pupil premium effectively. https://researchbriefings.files.parliament.uk/documents/SN06700/SN06700.pdf
2013	New GCSE syllabuses for English language and maths announced, to be taught in schools from September 2015. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/254513/GCSE_consultation_-_government_s_response.pdf

Table A2: Level of Missingness in each variable included in the data imputation

Variable Name	Valid N	% missing	Variable Name	Valid N	% missing	Variable Name	Valid N	% missing
workless1	11,524	0.00	parent2	9,356	0.19	cdnvabil	8,993	0.22
sex	11,524	0.00	workless2	9,310	0.19	cdpcabil	8,966	0.22
eth1	11,524	0.00	hle2	9,276	0.20	CDCONDA0g	8,757	0.24
born1	11,524	0.00	lsi2	9,276	0.20	BDCONDA0g	8,750	0.24
aimdscoe	11,524	0.00	disorg2	9,275	0.20	CDEMOTA0g	8,743	0.24
aimdeste	11,524	0.00	canthear2	9,275	0.20	CDPEERA0g	8,738	0.24
parent1	11,524	0.00	tenure2	9,275	0.20	BDEMOTA0g	8,731	0.24
lang1g	11,524	0.00	poverty2	9,251	0.20	CDHYPEA0g	8,689	0.25
season	11,524	0.00	parent3	9,147	0.21	BDPEERA0g	8,681	0.25
ageCMbirth	11,521	0.00	fsm5	9,147	0.21	bdbasa00	8,679	0.25
nparcareg	11,514	0.00	workless3	9,146	0.21	BDHYPEA0g	8,661	0.25
ghlth1	11,513	0.00	ccyage00	9,143	0.21	BDPEERA0g	8,681	0.25
lsli1	11,510	0.00	regbed5	9,099	0.21	bdbasa00	8,679	0.25
parnvq	11,508	0.00	readto5a	9,097	0.21	BDHYPEA0g	8,661	0.25
room1g	11,501	0.00	regmeal5	9,097	0.21	workless4	8,344	0.28
nbfed	11,494	0.00	cmhealth5	9,096	0.21	parent4	8,344	0.28
dishome1	11,492	0.00	lsi5	9,092	0.21	DCAGEY00	8,337	0.28
tenure1	11,492	0.00	tenure3	9,086	0.21	poverty4	8,332	0.28
damp1	11,488	0.00	poverty3	9,075	0.21	lsi7	8,303	0.28
disarea1	11,487	0.00	numsch11	9,019	0.22	stayp7	8,293	0.28
lbw	11,485	0.00	afterclub5	9,017	0.22	tenure4	8,265	0.28
poverty1	11,481	0.00	reluct5a	9,016	0.22	MATHS7SA	8,211	0.29
malg1	10,935	0.05	cdpsabil	9,005	0.22	DCWRAB00	8,210	0.29
formatc	9,357	0.19	bbasage	9,001	0.22	Bracken	8,196	0.29
DCPCAB00	8,177	0.29	cmteachget11	7,681	0.33	FCWRDSC	6,671	0.42
poverty5	8,143	0.29	EPEXTU0A	7,673	0.33	eng	6,228	0.46

Variable Name	Valid N	% missing	Variable Name	Valid N	% missing	Variable Name	Valid N	% missing
parent5	8,143	0.29	EPEXTU0B	7,673	0.33	maths	6,228	0.46
DDHYPEA0g	8,027	0.30	BDMPCLA0	7,556	0.34	prepsec11	5,778	0.50
tenure5	8,001	0.31	ssch_lkfor11	7,369	0.36	EQ33	5,769	0.50
CC_FSPTOTALg	7,993	0.31	poverty6	7,249	0.37	parin11g	5,762	0.50
CC_CLLg	7,992	0.31	parent6	7,249	0.37	stay16	5,761	0.50
CC_MDg	7,992	0.31	workless6	7,249	0.37	behav11	5,742	0.50
AGE	7,988	0.31	fsm14	7,132	0.38	univ18	5,715	0.50
stayp11	7,971	0.31	lsi14	7,132	0.38	sen11b	5,699	0.51
EVSABIL	7,965	0.31	secsch14	7,132	0.38	eng11b	5,690	0.51
univp11	7,964	0.31	susp14	7,131	0.38	occasp14g	4,826	0.58
BDMPIAA0	7,950	0.31	stayp14	7,110	0.38			
kessler2b	7,936	0.31	tenure6	7,096	0.38			
schhappy11	7,898	0.31	FPEXTU0A	7,096	0.38			
cmparint11	7,854	0.32	FPEXTU0B	7,096	0.38			
EDPEER00g	7,793	0.32	occasp11g	7,032	0.39			
EDCOND00g	7,790	0.32	quiet14	7,019	0.39			
EDEMOT00g	7,785	0.32	FPEERg	7,015	0.39			
EDHYPE00g	7,768	0.33	FCONDUCTg	7,013	0.39			
cmstaysch11	7,730	0.33	FEMOTIONg	7,012	0.39			
regbed11	7,718	0.33	truant14	7,007	0.39			
lsi11	7,713	0.33	FHYPERg	7,004	0.39			
afterclub11	7,707	0.33	FCSTYY00	6,948	0.40			
fsm11	7,703	0.33	FCSTYU00	6,717	0.42	AVERAGE MISSING		0.25

Table A3: predicting delayed school readiness at age 3

	Family SES	Family Education Support	Child Characteristics	Child Behaviour	All
Family SES (9 months)					
Single parent	0.93				0.89
	(0.10)				(0.10)
Teenage Mother	1.00				1.15
	(0.11)				(0.14)
Workless household	1.68***				1.60***
	(0.18)				(0.18)
Highest Qualification (ref: NVQ4+)					
None/NVQ1	2.76***				2.22***
	(0.32)				(0.29)
NVQ2	2.15***				1.91***
	(0.24)				(0.22)
NVQ3	1.62***				1.53***
	(0.20)				(0.19)
Rent home	1.36***				1.30**
	(0.11)				(0.11)
Overcrowded home	1.41***				1.30**
	(0.12)				(0.12)
Damp home	1.14				1.12
	(0.12)				(0.12)
IMD [bottom 2 quintiles]	1.54***				1.38***
	(0.13)				(0.12)
Poor/Fair general health	1.19				1.07
	(0.12)				(0.11)
Depressive symptoms	1.02				0.90
	(0.10)				(0.10)

	Family SES	Family Education Support	Child Characteristics	Child Behaviour	All
English + Other Language	2.48***				1.63***
	(0.23)				(0.22)
Education Support					
Home Learning Environment: bottom 20%		2.26***			1.33**
		(0.18)			(0.12)
Child Characteristics					
Female			0.58***		0.61***
			(0.04)		(0.04)
White			0.26***		0.55***
			(0.03)		(0.07)
Low birthweight			1.78***		1.57***
			(0.19)		(0.19)
1st born			0.70***		0.74***
			(0.05)		(0.06)
Longstanding illness (3)			1.23*		1.20
			(0.11)		(0.11)
Never breastfed			2.21***		1.31**
			(0.17)		(0.11)
Behaviour (3)					
SDQ Emotional problems				1.69***	1.19
				(0.16)	(0.13)
SDQ Conduct problems				1.26**	1.07
				(0.09)	(0.09)
SDQ Hyperactivity problems				1.82***	1.50***
				(0.13)	(0.12)
SDQ Peer problems				1.88***	1.38***
				(0.14)	(0.11)

	Family SES	Family Education Support	Child Characteristics	Child Behaviour	All
R^2	.12	.02	.08	.05	.15
N	11524	11524	11524	11524	11524

Exponentiated coefficients; Standard errors in parentheses; [*] $p < 0.05$, [**] $p < 0.01$, [***] $p < 0.001$

Table A4: predicting being below expected level in EYFS CLL and MD

	Bracken	Family SES	Family Education Support	Child Characteristics	Child Behaviour	All
Bracken School Readiness: Delayed	5.61*** (0.39)	3.74*** (0.30)	5.27*** (0.37)	5.14*** (0.42)	4.81*** (0.35)	3.64*** (0.32)
Family SES (9 months)						
Single parent		0.81* (0.09)				0.81 (0.09)
Teenage Mother		1.17 (0.12)				1.22 (0.14)
Workless household		1.42** (0.15)				1.40** (0.16)
Highest Qualification (ref: NVQ4+)						
None/NVQ1		2.60*** (0.25)				2.30*** (0.24)
NVQ2		1.83*** (0.17)				1.73*** (0.16)
NVQ3		1.43*** (0.14)				1.35** (0.14)
Rent home		1.49*** (0.12)				1.46*** (0.12)
Overcrowded home		1.35*** (0.10)				1.34*** (0.11)
Damp home		0.96 (0.07)				0.97 (0.08)
IMD [bottom 2 quintiles]		1.11 (0.09)				1.07 (0.09)
Poor/Fair general health		1.31***				1.21*

	Bracken	Family SES	Family Education Support	Child Characteristics	Child Behaviour	All
		(0.10)				(0.10)
Depressive symptoms		1.10				1.01
		(0.09)				(0.09)
English + Other Language		1.09				1.05
		(0.12)				(0.15)
Education Support						
Not read to every day (5)			1.32***			1.09
			(0.08)			(0.07)
Home Learning Environment: bottom 20%			1.72***			1.35***
			(0.13)			(0.11)
Child Characteristics						
Female				0.74***		0.75***
				(0.04)		(0.05)
White				0.62***		0.89
				(0.06)		(0.09)
Low birthweight				1.57***		1.39**
				(0.17)		(0.16)
1st born				0.83**		0.88
				(0.05)		(0.07)
Longstanding illness (3)				1.19		1.15
				(0.10)		(0.11)
Never breastfed				1.92***		1.22**
				(0.13)		(0.09)
Season Born (ref: Autumn)						
Winter				1.63***		1.69***
				(0.15)		(0.16)

55

	Bracken	Family SES	Family Education Support	Child Characteristics	Child Behaviour	All
Spring				2.42***		2.66***
				(0.21)		(0.25)
Summer				4.42***		4.87***
				(0.41)		(0.48)
Behaviour (3)						
SDQ Emotional problems					1.47***	1.16
					(0.14)	(0.12)
SDQ Conduct problems					1.26**	1.10
					(0.08)	(0.08)
SDQ Hyperactivity problems					1.68***	1.48***
					(0.13)	(0.12)
SDQ Peer problems					1.30***	1.10
					(0.09)	(0.08)
R^2	.08	.14	.09	.14	.10	.20
N	11524	11524	11524	11524	11524	11524

Exponentiated coefficients; Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5: predicting not gaining GCSE English and Maths Grade 4 or higher

	EYFS CLL+MD	Family SES	Family Education Support	Child Characteristics	Child Behaviour + Cognition	All
School readiness: delayed	1.90***	1.62***	1.85***	1.81***	1.75***	1.49***
	(0.18)	(0.17)	(0.18)	(0.19)	(0.17)	(0.16)
CLL+MD: Below expected level	3.90***	3.08***	3.81***	3.69***	3.60***	2.88***
	(0.30)	(0.24)	(0.29)	(0.30)	(0.28)	(0.24)
Family SES (9 months)						
Single parent		0.97				0.98
		(0.13)				(0.14)
Teenage Mother		1.03				1.06
		(0.14)				(0.16)
Workless household		1.25				1.22
		(0.16)				(0.16)
Highest Qualification (ref: NVQ4+)						
None/NVQ1		2.48***				2.12***
		(0.28)				(0.25)
NVQ2		1.88***				1.69***
		(0.17)				(0.17)
NVQ3		1.47**				1.37**
		(0.17)				(0.16)
Rent home		1.32**				1.29*
		(0.14)				(0.14)
Overcrowded home		1.09				1.08
		(0.10)				(0.11)
Damp home		1.14				1.14
		(0.11)				(0.11)
IMD [bottom 2 quintiles]		1.18				1.18

	EYFS CLL+MD	Family SES	Family Education Support	Child Characteristics	Child Behaviour + Cognition	All
		(0.11)				(0.11)
Poor/Fair general health		1.27**				1.24*
		(0.12)				(0.12)
Depressive symptoms		1.06				1.01
		(0.10)				(0.10)
English + Other Language		0.41***				0.58***
		(0.05)				(0.09)
Education Support						
Not read to every day (5)			1.11			0.97
			(0.07)			(0.07)
Home Learning Environment: bottom 20%			1.24*			1.12
			(0.12)			(0.11)
Child Characteristics						
Female				0.69***		0.68***
				(0.05)		(0.05)
White				1.50***		1.34*
				(0.16)		(0.20)
Low birthweight				1.37*		1.29*
				(0.17)		(0.17)
1st born				0.89		0.89
				(0.07)		(0.08)
Longstanding illness (3)				1.28**		1.24*
				(0.12)		(0.11)
Never breastfed				1.95***		1.38***
				(0.14)		(0.10)
Season Born (ref: Autumn)						

58

	EYFS CLL+MD	Family SES	Family Education Support	Child Characteristics	Child Behaviour + Cognition	All
Winter				0.93		0.94
				(0.09)		(0.09)
Spring				0.89		0.94
				(0.09)		(0.09)
Summer				0.90		0.95
				(0.09)		(0.10)
Behaviour (3)						
SDQ Emotional problems (3)					1.18	1.09
					(0.13)	(0.13)
SDQ Conduct problems (3)					1.38***	1.16
					(0.11)	(0.09)
SDQ Hyperactivity problems (3)					1.52***	1.34***
					(0.11)	(0.10)
SDQ Peer problems (3)					1.00	0.90
					(0.08)	(0.08)
R^2	.09	.13	.09	.11	.10	.15
N	11524	11524	11524	11524	11524	11524

Exponentiated coefficients; Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$