

Reading for pleasure and attainment in vocabulary and mathematics

Dr. Alice Sullivan (corresponding author)*
Director, 1970 British Cohort Study
Centre for Longitudinal Studies

Mr. Matt Brown
Survey Manager, 1970 British Cohort Study
Centre for Longitudinal Studies

*Contact details:
Centre for Longitudinal Studies
Institute of Education, University of London
55-59 Gordon Square
LONDON
WH1H 0NU
Tel: 020 76126661
Email: a.sullivan@ioe.ac.uk

Acknowledgements: We are grateful to the BCS70 cohort members for their contribution over many years. Thanks to Tarek Mostafa for providing the analysis in Appendix A1. Thanks to Dick Wiggins for comments on an earlier draft, and to two anonymous reviewers.

Reading for pleasure and progress in vocabulary and mathematics

Abstract

This paper examines inequalities in attainment in vocabulary and mathematics at age 16 for a nationally representative cohort of people born in Britain in 1970 (the 1970 British Cohort Study). Our analytical sample is $n=3,583$ cohort members who completed vocabulary and mathematics tests at age 16. We explore whether inequalities due to childhood social background are similar across the linguistic and mathematical domains, or whether they differ, and to what extent these inequalities are driven by families' social class position, parents' education and home reading resources. We examine the role of children's own reading for pleasure controlling for all these background factors. As reading can be seen as an indicator of 'cultural capital', we also test the influence of an alternative indicator of cultural capital, playing a musical instrument. Our longitudinal analysis addresses the question of the extent to which differences in attainment are determined by age 10; and which factors are linked to a growth in differentials during adolescence. We show that childhood reading is linked to substantial cognitive progress between the ages of 10 and 16, whereas playing an instrument is not. Reading is most strongly linked to progress in vocabulary, with a weaker, but still substantial link to progress in maths. Strikingly, reading for pleasure is more strongly linked than parental education to cognitive progress in adolescence.

Keywords: Reading; longitudinal; BCS70; vocabulary; mathematics.

Introduction

Persistent socio-economic inequalities in educational attainment and cognitive scores have been documented by many studies over the years, and the explanation of these social inequalities is one of the central problems within the sociology of education (Halsey et al. 1980). Debate continues regarding the mechanisms via which privileged families ensure the educational and subsequent occupational success of their children, and the relative importance of economic and cultural resources in determining class differentials in educational outcomes.

Sociologists have put forward both cultural and economic explanations for educational inequalities. The most prominent theory emphasising the importance of cultural resources, Bourdieu's theory of cultural reproduction (Bourdieu & Passeron [1977] 1990) has generated much debate about the nature of cultural capital. Interpretations can be divided into the 'status seeking' and 'information processing' views (Ganzeboom 1982). According to the status seeking view, children who display cultural attributes that are valued by the school are rewarded by teachers with high attainment. In contrast, the 'information processing' view suggests that particular cultural activities actually foster intellectual development. We have argued that reading is distinctive because it develops linguistic ability and wider knowledge (Sullivan 2002; 2007). Reading differs in this way from other cultural activities such as music or going to galleries and museums (often termed 'beaux arts' participation). We know that the home reading culture is important for children's early cognitive scores (Byford et al. 2012). Past studies have also found that books in the home and children's reading behaviour, but not 'beaux arts' participation, help to explain social differentials in children's outcomes (De Graaf et al. 2000; Sullivan 2001) (Cheung & Andersen 2003; Georg 2004; Jaeger 2011; Sullivan 2001).

The role of language has been neglected in most empirical studies of cultural reproduction, yet Bourdieu's discussion of cultural capital emphasises the importance of language as the key to success in school. "*Obvious in the literary disciplines but more subtle in the sciences, the ability to manipulate academic language remains the principal factor in success in examinations*" (Bourdieu et al. 1994) (p.21). Hirsch (1983) also emphasises vocabulary, though from a different perspective, on the grounds that knowledge of words is both an adjunct to knowledge of concepts and assists further learning. It has been argued that reading must be a particularly important driver of vocabulary development, given the paucity of vocabulary used in speech compared to books, even comparing children's books to adult speech (Cunningham & Stanovich 1998).

Children pick up styles of speech, vocabulary and forms of reasoning simply by hearing their parents talk, and also pick up reading habits through seeing their parents read, and having reading materials readily available in the home (Sullivan 2007). We may expect this passive cultural transmission to be most important in the case of linguistic skill, since studies have found huge differences in the number of different words that children are exposed to in middle and working class homes (Hart & Rinsley 1995). Vocabulary is transmitted within the home almost constantly, without any conscious effort. In contrast, while parents may well

seek to promote their children's success in subjects such as mathematics, this must typically be done consciously, with discrete time set aside for the task. It is arguable therefore that we should expect wider disparities according to parental cultural resources in children's linguistic ability than in their abilities in mathematics. On the other hand if reading ability and vocabulary are central to further learning in all disciplines, we may expect the influences on vocabulary development to also drive attainment in other subjects including mathematics. Evidence for this view is provided by a recent study of monozygotic twins (Ritchie et al. 2014 in press) showing that learning to read well leads to improved general intelligence. In this paper, we are able to address whether the socio-economic and cultural factors driving vocabulary are essentially the same as those driving mathematics attainment, or whether they differ.

Studies focussing on children's reading have faced challenges in unpacking the reciprocal relationship between ability and participation in reading (Cunningham & Stanovich 1998). We know that children who read a lot perform well in school, but are they bright because they read, or is it simply that they read because they are bright? Reviews of the literature find extensive evidence for an association between reading frequency and reading attainment (Twist et al. 2007), but note the difficulty in establishing whether reading frequency actually leads to improved attainment in the absence of compelling longitudinal evidence (Clark & De Zoysa 2012; Clark & Rumbold 2006; Department for Education 2012; Department for Education Education Standards Research Team 2012). While some longitudinal studies on reading exist, they have typically been small scale, covered relatively short periods, and lacked controls for socio-economic background (Taylor et al. 1990).

The growth in cognitive inequalities according to socioeconomic status during childhood has been established by analyses of the British cohort studies of 1946, 1958, 1970 and 2000 (Douglas 1964; Feinstein 2003; 2004; Fogelman & Goldstein 1976; Sullivan et al. 2013), but for the 1970 cohort this has only so far been examined up to age 10. Here we analyse the extent to which social inequalities in the 1970 cohort continued to grow during adolescence, and which childhood resources drive this growth. Specifically, what is the role of reading for pleasure during childhood and adolescence? As far as we are aware, ours is the first large study to take a longitudinal approach to reading and cognitive development during adolescence.

Research questions

In this paper, we focus on the potential role of both parents' and children's reading in explaining differentials in cognitive test scores at age 16 for a cohort of children born in 1970. The BCS70 is a large, nationally representative, longitudinal birth cohort study, containing rich measures of both cognition and home background, which provides some strong advantages in tackling these questions.

We address the following questions.

1. Are inequalities due to parental social background similar across the domains of vocabulary and mathematics, or do they differ? We hypothesize that parental education (which is typically seen as reflecting cultural and cognitive resources) but not occupational social class (which is an indicator of economic position) may be more strongly linked to vocabulary than to maths attainment.
2. Are inequalities due to parental social background explained by parental reading environment, behaviour and ability? We hypothesise that the link between parental education and children's test scores is more likely to be accounted for in this way than the link between parental social class and children's test scores.
3. Is the influence of parental reading environment, behaviour and ability explained by children's own reading behaviour?
4. Which factors are linked to changing test scores between the ages of 10 and 16? In particular, is the child's own reading linked to cognitive progress? We hypothesise that children's reading should be more strongly linked than children's music participation to cognitive progress. We also hypothesise that childhood reading should be more strongly linked to progress in vocabulary than to progress in maths, since reading directly exposes children to new vocabulary. Nevertheless, we expect that reading will develop mathematical abilities to some degree, since reading ability facilitates learning in all subjects.

Data and variables

The 1970 British Cohort Study (BCS70) follows the lives of more than 17,000 people born in England, Scotland and Wales in a single week of 1970 (Elliott & Shepherd 2006). Over the course of cohort members' lives, the BCS70 has collected information on health, physical, educational and social development, and economic circumstances among other factors. Since the birth survey in 1970, there have been eight surveys (or 'waves') at ages 5, 10, 16, 26, 30, 34, 38 and 42. An understanding of the educational progress of this cohort during their childhood is vital to understanding their later life course trajectories.

The 1970 cohort study is rich in cognitive test scores throughout the early years, and the early test scores (up to age ten) have been analysed extensively, including influential work by Feinstein (2003; 2004). The cognitive scores at age ten have also been shown to be important for adult outcomes, including in employment (Breen & Goldthorpe 2001) and health (Batty et al. 2007). To date, almost no research has been carried out using the age 16 test scores (though see (Duncan et al. 2012)), due to a lack of awareness of the existence of these scores among the research community, and because the arithmetic dataset was not deposited until 2008. We aim to encourage wider use of these variables by researchers.

Because we exploit data from all of the childhood waves of the study, including the age 16 wave, the problem of missing data must be addressed. The age 16 survey employed sixteen separate survey instruments, and unfortunately coincided with a teachers' strike which affected the completion of those instruments, including cognitive tests, that were administered via schools (Dodgeon 2008). This led to substantial instrument non-response. Nevertheless, the 1986 sample is more representative in terms of the birth characteristics of the sample in 1970 than any other wave of the study excluding the birth wave (Mostafa & Wiggins 2014). Appendix A1 shows logit response models for response to both the arithmetic and vocabulary tests in terms of the birth characteristics of the 1970 sample. Respondents who took both tests were no more highly selected than the 1986 sample as a whole. Levels of missing data for the variables used in our analysis are provided in table 1. As list-wise deletion was not a practical option, we use multiple imputation to 'fill-in' values of any missing items in the variables selected for our analysis adopting Schafer's algorithm under the assumption of 'missing at random' (MAR). All reported analyses are averaged across twenty replicates based upon Rubin's Rule for the efficiency of estimation under a reported degree of missingness across the whole data of just under 20% (Little & Rubin 1987). The analytical sample consists of the 3, 583 cohort members who completed the age 16 vocabulary and maths tests.

It is important to acknowledge that people's levels of motivation and compliance, as well as potential stereotype-threat (Croizet & Claire 1998; Spencer et al. 1999) will affect their scores in cognitive tests. We also acknowledge that multiple-choice tests do not capture the full range of academic skills, and girls tend to fare worse in multiple-choice tests than in other forms of assessment (Gipps & Murphy 1994). We do not interpret the tests used here as providing an estimation of innate intelligence. They are simply tests of attainment based on the capability and motivation to complete a particular task under given conditions.

Analytical Strategy

We will first present a descriptive analysis of children's attainment trajectories. Subsequently we will model mathematics and vocabulary scores at age 16 as a function of family characteristics, childhood reading behaviour and children's prior attainment scores. This will be presented as a series of multivariate linear regression models. The outcome and explanatory variables to be used in these analyses are described below.

Outcome variables: Mathematics and vocabulary at age 16

Influences on cognitive scores and changes in these scores may be expected to differ according to the nature of the assessment, and the demands the assessment makes on processing capacity or problem solving as opposed to knowledge (Richards & Sacker 2003). Here we examine attainment scores in vocabulary and maths. Vocabulary reflects linguistic competence, which we expect to be developed within the home and through reading rather than primarily through schoolwork. An advantage for our purposes of the vocabulary test

used here is that it is purely a test of linguistic competence, with no verbal reasoning element. The vocabulary test is purely dependent on recall, while the maths test can be seen as a test of problem solving, although of course background knowledge is important here too.

Maths attainment was assessed using the Applied Psychology Unit (UPU) Arithmetic test - a 30 minute assessment comprising 60 multiple choice items covering arithmetic, probabilities and area (Closs & Hutchings 1976; Dodgeon 2008). Vocabulary was assessed using a 75 item multiple choice test. Table 1 shows descriptive statistics for the mathematics and vocabulary tests.

TABLE 1

Explanatory variables

This section outlines the predictors to be used in the regression models presented in the results section of this paper. The descriptives for the variables in our regression analyses are shown in tables 2 and 3.

TABLE 2

TABLE 3

Model 1: Socio-economic background, sex and siblings

Social class is coded according to the NS-SEC schema (Goldthorpe 1997) at age ten (in 1980). NS-SEC is an occupational schema, and determines class position in terms of employment relations. It reflects not just income, but longer term economic security, stability and prospects, as reflected in a person's labour market position. It also reflects power in terms of relationships of authority, control and autonomy within the workplace (Goldthorpe & McKnight 2006). We adopt a 'dominance' approach to household social class, taking the mother or father's occupation whichever is higher, using the three category classification¹. Parents' education is coded as the highest qualification of the mother or father (whichever of the two is higher).

Position in the birth order is a well established predictor of child outcomes (Nisbet 1953).

Sex is included to test whether development in maths and vocabulary differed between boys and girls.

Model 2: Model 1 + Home reading culture

We are able to provide a relatively thorough operationalisation of the home reading culture compared to many previous studies. We include not just reading to the child, but also parental reading behaviour, newspapers (broadsheet or tabloid) in the home and parental reading ability.

When the cohort members were aged 5, mothers were asked on how many days of the last 7 the child had been read to. In 1986, the mother was asked whether she and her husband read books. Seeing parents reading may affect children's attitudes to reading, and parents' reading habits are also likely to be positively linked to parents' reading ability. Mothers were also asked which newspapers were in the home and were thus available for the teenager to read. We are able to differentiate between broadsheet and tabloid households and those who did not have newspapers at home. The prose style of tabloid (then as now) was simpler and geared towards a lower reading age and smaller vocabulary than the broadsheets. During the 1980s, newspaper readership was high, and the type of newspaper read was (as it remains) a strong cultural identifier (Chan & Goldthorpe 2007).

In the absence of a reading assessment for parents, we rely on a self-reported measure of reading difficulties. Mothers were asked whether they or their husband had reading difficulties, either when learning to read or currently (maternal self-completion 1986). Positive responses to these items are low, with 5% of mothers admitting to any difficulties for themselves, and 4% for their husbands. Item non-response was high on these questions (6% for mothers and 9% for fathers) and, in preliminary analyses, was found to be significantly associated with poor performance in both assessments suggesting that those with reading difficulties may have been embarrassed or reluctant to report this. We acknowledge also that subjective reporting of difficulties tends to be much lower than actual tested difficulties (Bynner & Parsons 2006). Here we use a binary variable which indicates whether the mother reported that either she herself or the father had difficulties with reading. Although parental reading ability and habits and reading materials available in the home were captured when the cohort member was age 16, we consider these to be variables which would be unlikely to be subject to significant change during the preceding years of the cohort member's life and therefore do not see it as problematic to treat these variables as predictors of outcomes at age 16. We acknowledge the drawback that we have no measure of books in the home, a variable which has been shown to be a powerful predictor of children's educational attainment (Chiu & Chow 2010; Evans et al. 2010).

Model 3: Model 2 + Cohort member's own reading and playing a musical instrument

The 1980 self-completion pupil questionnaire includes items on reading books and going to the library. The 1986 cohort member self-completion questionnaires contained items on reading books and newspapers. Book reading declined between the ages of ten and 16. Some difference may be due to the earlier variable being reported by mothers while the later variable is self-reported, but it is also likely that there was a genuine decline in reading among teenagers, perhaps partly due to a lack of availability and promotion of suitable books for this age group. For example, libraries during the 1980s typically devoted very little space to books aimed at adolescents. This decline in reading for pleasure as children get older is in line with previous research (Clark & Rumbold 2006).

We also include variables indicating whether the cohort member played a musical instrument at age 5 and age 10. Playing an instrument would not be expected to directly develop skills in maths or vocabulary, but can be seen as a more general indicator of the cultural climate of the

home, and arguably represents ‘status seeking’ behaviour, in the sense that parents who encourage their children to play an instrument are complying with a strong social norm within the educated middle-classes.

Model 4: Model 3 + Cognitive tests at five and ten

The cohort members took age-appropriate tests at age five and ten. These are included in our final model, to assess cognitive progress between the ages of 10 and 16. Descriptive statistics for these tests are shown in table 4. We use Principal Components Analysis (PCA) to extract a single scale for cognition at each age (5 and 10). Full details of the tests used and the PCA analysis are provided in Parsons (2014).

TABLE 4

Age five tests

The cohort members took five tests designed to capture verbal and non-verbal skills: *Copying designs* (Rutter et al. 1970); *English picture vocabulary* (Brimer & Dunn 1962); *Human figure drawing* (draw-a-man) (Goodenough 1926; Harris 1963); *Complete a profile*; and *Schonell graded reading* (Golding 1975).

Age ten tests

The eight tests taken at age ten were: *Shortened Edinburgh reading test* (Godfrey Thompson Unit 1978); *Pictorial language comprehension test*; *Friendly maths test*; *Spelling* (dictation task); *British Ability Scales* (BAS) (Elliott et al. 1979; Hill 2005): Two verbal subscales (word definitions and word similarities) and two non-verbal subscales (digit recall and matrices) (Butler et al. 1980).

Results

We begin with a simple graphical presentation of children’s trajectories over time, before presenting the results of our regression analyses.

Attainment trajectories

First we look at children’s progress in vocabulary and mathematics in terms of their percentile rankings. Figure 1 shows vocabulary trajectories between the ages of five and 16 comparing frequent readers to infrequent readers and the children of at least one graduate to those whose parents had no qualifications. Young people who read books most frequently at 10 and 16 (20%) are classified as high readers. Those who were in the lowest book reading category at both ages or the lowest at one and second lowest at the other (21%) were categorised as low readers.

FIGURE 1

High readers started out in the 57th percentile at age 5, and increased their position by 5 percentiles by age 10 and a further 7 percentiles by age 16, meaning that by age 16, the high readers had caught up with the children of graduates. The least frequent readers saw a decline of 8 percentiles over the same period.

The children of both graduates and parents with no qualifications maintained a roughly constant position between the ages of 5 and 10. Interestingly, the children of parents with no qualifications maintained a roughly constant position between 5 and 16, with a slight increase in their mean rank at 10 and a slight decline by 16. However, the gap between the children of graduates and the children of the unqualified grew during the secondary school years, as the children of graduates improved their position.

FIGURE 2

Turning to children's maths trajectories, we must first note that there was no maths test at age 5, so we use a non-verbal score (Copying Designs) as a baseline. The children of graduates improve their position between 5 and 16 by 3 percentiles (as opposed to 5 for vocabulary). The position of the children of parents with no qualifications barely changed between 5 and 16. The gap between frequent and infrequent readers remained fairly constant between 5 and 16.

In summary, these charts suggest that parents' education and children's reading make a difference to children's progress in both vocabulary and maths during the secondary school years. However, the role of children's reading appears to be particularly important for vocabulary development. In the analyses to follow, we test whether these descriptive results hold given an extensive set of control variables.

Multivariate Regression Analysis

We present a multivariate regression analysis (also known as multivariate general linear model or multivariate response model), which treats the outcome variables jointly, and hence uses only those cases with data for both test scores at age 16. This allows us to compare model fit across the two outcomes, and to examine differences in the predictors across the two outcomes. The dependent variables are treated as percentage test scores, in order to make the coefficients interpretable as percentage point differences. We have run parallel analyses (available on request) treating the dependent variables as standardised z scores, which did not change the results in any substantial way.

TABLE 5

Table 5 shows the multivariate regression analysis results. In Model 1 we control only for sex and family background. Model 1 shows no link between gender and either maths or

vocabulary scores. Parents' qualifications are significantly linked to both test scores, with the children of more highly educated parents achieving higher scores. Children who had a parent with a degree scored around 13 percentage points higher than those whose parents had no qualifications in both vocabulary and maths. The advantage associated with managerial and professional occupations ('the salariat') compared to the manual/routine class was considerably smaller, 4 percentage points for vocabulary and 5 for maths. Older siblings are negative for both vocabulary and maths, with a 2 percentage point disadvantage for each additional older sibling.

In model 1, we can also note that model fit is weak for both maths and vocabulary, but fit is stronger in the case of vocabulary ($R^2 = 0.15$) than for maths (0.10), suggesting that ascribed social characteristics are more important predictors of variability in vocabulary than in maths.

In model 2, we introduce variables related to the home reading climate. How often the child was read to at age five is significantly positive for both vocabulary and maths, with an advantage of 1.2 and 0.9 percentage points respectively for each additional day the child was read to in the last week. Parents' reading books in their spare time is also significantly positive for vocabulary (an advantage of 1.6 percentage points), but not for maths. Parents' self-reported reading problems are significantly negative for both scores (-2.9 percentage points for vocabulary, -4.9 for maths). Having broadsheet newspapers in the home is linked to higher scores on both vocabulary (4.9) and maths (4.4), while tabloids in the home were not significant compared to having no national newspapers in the home.

In this model, the coefficients for both parental education and social class are somewhat reduced, but remain consistently statistically significant. In model 2, the improvement in model fit is greatest for vocabulary (R^2 increases from 0.15 to 0.20, compared to 0.10 to 0.12 for maths) suggesting that the home reading climate accounts for more of the variability in vocabulary than for maths.

Model 3 introduces the child's own reading behaviour. Book reading at 10 and 16 and newspaper reading at 16 are powerfully linked to both vocabulary and maths. Reading often at age 10 is linked to an advantage of 13 percentage points in vocabulary and 14 in maths. Reading more than weekly at 16 is linked to a further advantage of 7 percentage points in vocabulary and 3 in maths, while reading a newspaper more than weekly at age 16 is associated with an additional 5 percentage points in vocabulary and 7 in maths. This model also includes playing a musical instrument at 10 and 16. Playing an instrument is linked to test scores in this model, but much less strongly than reading is. Playing an instrument at 10 is linked to an advantage of only 1 percentage point in vocabulary and 2 percentage points in maths. Playing an instrument at 16 is not linked to a statistically significant advantage in maths, but is linked to a 1.5 percentage point advantage in vocabulary.

Interestingly, gender becomes significant for both subjects in this model, with positive coefficients for boys. This suggests that, while boys' absolute performance was not different from girls' in either maths or vocabulary, boys performed at higher levels than girls for any

given level of recreational reading – in other words, boys performed as well as girls, despite not reading as much as girls. The influence of variables reflecting the parents' reading culture is reduced in this model, suggesting that the influence of parents' reading culture is somewhat explained by the child's own reading. However, the negative link between tabloid newspapers and vocabulary (a 1 percentage point reduction in scores) becomes statistically significant in this model.

Model 3 also shows a substantially improved model fit for vocabulary (from $R^2 = 0.2$ to 0.32), with a smaller improvement for maths (0.12 to 0.18), suggesting that the child's own reading is most important for vocabulary development. In particular, the cohort member's reading at age 16 was more strongly linked to vocabulary than to mathematics attainment.

Model four introduces the cohort member's test scores at the ages of five and ten. Principal Components Analysis (PCA) is used to extract a single standardised Z score for cognition at each age (5 and 10).. The inclusion of the age five and ten test scores in the model means that it becomes a model of how far the predictions of model 3 had already been established by age ten, and how far they continued to be reflected in changes in the child's test scores between age ten and age 16. Essentially it is a model of progress, with the proviso that the tests taken at ages five and ten were not the same as those taken at 16, although vocabulary was measured at both 5 and 10, and maths was also measured at age 10. Coefficients in this model could be biased if measurement error in the cognitive tests at ages 5 and 10 are linked to other variables of interest (Jerrim & Vignoles 2013). Therefore, we have attempted to minimise the risk of spurious results due to measurement error in any given test by including the full set of test scores at both ages in our analysis.

An increase of 1 standard deviation in cognitive scores at age 5 was associated with an increase of 1 percentage point in vocabulary, but no significant gain for maths. The age 10 score was more highly predictive of attainment at 16, as would be expected. A one standard deviation increase in age 10 scores was linked to gains of 9 percentage points in vocabulary and 13 percentage points in maths. Model fit improves substantially, to $R^2 = 0.61$ for vocabulary and 0.49 for maths.

Many variables that were significant in model 3 become non-significant or marginally significant in model 4 because they are linked to absolute attainment in the test scores at age 16, but not to progress between ten and 16. Parents' education remains significant, but the coefficient is much reduced in size (the advantage due to a parental degree falls from 8 to 2.5 percentage points for vocabulary and from 9 to 2 percentage points for maths). Social class is non-significant in this model. The negative influence of elder siblings remains significant for both maths and vocabulary, with a one percentage point disadvantage in progress in maths and vocabulary for each additional elder sibling.

The influence of the home reading environment is much reduced, although broadsheet newspapers in the home and having been read to at age 5 remain significantly positive for vocabulary, though not for maths. Overall though, the influence of the home reading

environment has been largely mediated by the combination of the children's own reading behaviour and by their early attainment scores.

Importantly, the coefficients for cohort member's own reading behaviour, including reading books and newspapers, remain large and highly statistically significant in model 4. This suggests that it is not just the case that academically able children read more, but that leisure reading is linked to greater cognitive progress during the teenage years. This is in marked contrast to playing a musical instrument, which becomes non-significant in this model, apart from a borderline significant *negative* relationship between playing an instrument at age 10 and vocabulary development.

Taking the three variables reflecting childhood reading together, being in the top categories for all three adds up to a gain of 10 percentage points in vocabulary and 5 percentage points in maths. This compares to a difference associated with a parental degree of 2.5 percentage points for vocabulary and 2 percentage points for maths. In other words, the influence of reading on growth in academic attainment scores is substantial. Given that parental education has typically been found to be the strongest determinant of children's educational and cognitive outcomes, the fact that childhood reading matters more than parental education for cognitive progress during the secondary school years is striking.

Finally, in additional analysis (available on request) we tested whether there was an interaction between childhood reading and parents' education. This was to assess whether the cognitive return to reading differed according to parents' educational level. We found that there was a positive interaction for reading at age 16 (but not age 10) only for vocabulary. This differential may suggest that 16 year olds with more educated parents had access to higher quality reading materials, but unfortunately we lack the data to test this hypothesis directly.

Conclusions

Our initial descriptive analyses showed that differentials in vocabulary scores between the children of graduates and the children of parents with no qualifications increased sharply between the ages of ten and 16, but differentials for maths were stable. However, our subsequent regression analyses showed that differentials in test scores at age 16 due to parental social background did not vary greatly across the domains of mathematics and vocabulary. We hypothesised that parental education would be more strongly linked to vocabulary than to maths scores, but this was not the case. Rather, parental education was much more strongly linked than parental social class to both vocabulary and maths scores. This finding broadly supports Bourdieu's emphasis on cultural resources, suggesting that they matter more than economic resources, at least for cognitive outcomes.

The home reading culture, including reading to the child, parents' reading books and newspapers, and parental reading problems, was linked to children's test scores, and this had

a role in mediating the influence of parents' education, and also to some extent in mediating parents' social class. In order to interpret the influence of the home reading culture on children's outcomes, it is important to acknowledge that elements of this culture are likely to be strongly related to one another. For example, a mother who struggles with reading is likely to struggle to read to her child, and unlikely to read in her leisure time. Given the prevalence of adult literacy problems in Britain, (National Audit Office 2008) this is likely to be an important aspect of the educational disadvantage suffered by many children. Eight per cent of the BCS70 cohort were assessed as having poor basic skills in literacy at age 34, with strong evidence of intergenerational transmission of poor literacy and numeracy to their own children (Bynner & Parsons 2006)

Children's own reading behaviour was strongly linked to test scores in maths and vocabulary, and this accounted for some of the influence of parents' reading. Our findings support other work suggesting that children's leisure reading is important for educational attainment and social mobility (Taylor 2011), and suggest that the mechanism for this is increased development of cognitive skills. Once we controlled for the child's test scores at age five and ten, the influence of the child's own reading remained large and statistically significant, suggesting that the positive link between leisure reading and cognitive outcomes is not purely due to more able children being more likely to read a lot, but that reading for pleasure is actually linked to increased progress over time. We found a stronger link between reading and progress in vocabulary development than in mathematics. This is in line with our hypothesis, based on the grounds that books directly expose readers to new words, and therefore reading should influence vocabulary directly, whereas the influence on attainment in other areas would be indirect, via the fact that improved reading ability improves an individual's ability to learn across the whole curriculum. We expected this indirect influence of reading on progress in mathematics to be weaker than the hypothesised direct effect on vocabulary. The differential influence of reading on the two cognitive scores can therefore be seen as tentatively supporting a causal interpretation of the role of reading in vocabulary development. As a further check on the interpretation of the effect of reading, we included playing a musical instrument in our models. While it could be argued that reading simply proxies a particular cultural milieu within the family, or child characteristics such as diligence or concentration, we would argue that the lack of a significant association between playing a musical instrument (which should also proxy such characteristics) and cognitive progress does not accord with this view. In summary, although statistical regression analyses can never prove causality, we suggest that the combined evidence of temporal ordering and differential effects in line with theory makes a causal interpretation of the role of reading more plausible than the alternative explanations in this case.

From a policy perspective, our findings strongly back the need to support and encourage children's reading in their leisure time, especially given that the available evidence on trends over time suggests that children's reading for pleasure has declined in recent years (Clark & Rumbold 2006). Reasons for this decline in reading for pleasure may include increases in competing demands on young people's time, including homework, organised activities and the internet. However, there is scope for new technologies to be exploited to provide greater

access to books. Supporting young children in becoming confident readers is clearly necessary but not sufficient to achieve the goal of encouraging reading for pleasure throughout childhood and adolescence. Schools need to foster a love of reading as well as teaching reading as a skill. In light of the decline in leisure reading between the ages of ten and 16, our findings suggest the particular need to support teenagers' reading. We would also argue that supporting reading for pleasure among disadvantaged children could potentially provide a powerful tool in closing education attainment gaps (Connelly et al. 2014). Further research is needed on effective approaches to promoting reading for pleasure, but it seems clear that library services both within and outside schools are necessary to promoting reading, and sharing knowledge with individual young people about books they may enjoy, especially for those children who do not live in homes that are lined with books. The lack of library provision in many British schools is a cause for concern (APPG 2014). By definition, no one can be forced to read for pleasure, but practices such as silent reading periods in school may help to establish reading as a habit to be enjoyed.

In this paper, we have attempted to shed new light on educational inequalities and cultural reproduction processes by examining the role of reading in attainment trajectories over time, and the distinctive link between cultural resources and vocabulary. In future work, we intend to assess the role of cognition in determining educational attainment and life chances, including the question of whether vocabulary plays a distinctive role in determining educational attainment and social reproduction and mobility.

Bibliography

- APPG (2014) *The beating heart of the school: Improving educational attainment through school libraries and librarians* (London, Libraries All Party Parliamentary Group).
- Batty, G. D., Deary, I. J., Schoon, I. & Gale, C. R. (2007) Mental ability across childhood in relation to risk factors for premature mortality in early life: the 1970 British Cohort Study, *Journal of Epidemiology and Community Health*, 61, 977-1003.
- Bourdieu, P., Passeron, J.-C. & Saint-Martin, M. (1994) *Academic Discourse* (Cambridge, Polity Press).
- Bourdieu, P. & Passeron, J. C. ([1977] 1990) *Reproduction in education, society and culture* (London ; Beverly Hills, Sage Publications).
- Breen, R. & Goldthorpe, J. H. (2001) Class, mobility and merit - The experience of two British birth cohorts, *European Sociological Review*, 17(2), 81-101.
- Brimer, M. A. & Dunn, L. M. (1962) *English Picture Vocabulary Test* (Educational Evaluation Enterprises).
- Butler, B., Despotidou, S. & Shepherd, P. (1980) *1970 British Cohort Study (BCS70): Ten Year Follow-up* (London, Social Statistics Research Unit, City University).
- Byford, M., Kuh, D. & Richards, M. (2012) Parenting practices and intergenerational associations in cognitive ability, *International Journal of Epidemiology*, 41, 263-272.
- Bynner, J. & Parsons, S. (2006) *New light on literacy and numeracy* (London, NRDC Research Report).
- Chan, T. W. & Goldthorpe, J. H. (2007) Social Status and Newspaper Readership, *American Journal of Sociology*, 112(4), 1095-1134.
- Cheung, S.-Y. & Andersen, R. (2003) Time to read: Family resources and educational outcomes in Britain, *Journal of Comparative Family Studies*, 34(3), 413-433.
- Chiu, M. M. & Chow, B. W. Y. (2010) Culture, motivation and reading achievement: High school students in 41 countries, *Learning and Individual Differences*, 20, 579-592.
- Clark, C. & De Zoysa, S. (2012) *Mapping the interrelationships of reading enjoyment, attitudes, behaviour and attainment* (London, National Literacy Trust).
- Clark, C. & Rumbold, K. (2006) *Reading for pleasure: A research overview* (London, National Literacy Trust).
- Closs, S. J. & Hutchings, M. J. (1976) *APU arithmetic test* (London, Hodder and Stoughton).
- Connelly, R., Sullivan, A. & Jerrim, J. (2014) *Primary and secondary education and poverty: Anti poverty strategy review* (London, Joseph Rowntree Foundation).
- Croizet, J.-C. & Claire, T. (1998) Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds, *Personality and Social Psychology Bulletin*, 24(6), 588-594.
- Cunningham, A. E. & Stanovich, K. E. (1998) What reading does for the mind, *American Educator*, 22(1-2), 8-15.
- De Graaf, N. D., De Graaf, P. M. & Kraaykamp, G. (2000) Parental Cultural Capital and Educational Attainment in the Netherlands: A Refinement of the Cultural Capital Perspective, *Sociology of Education*, 73, 92-111.
- Department for Education (2012) *Research Evidence on Reading for Pleasure* (London, DfE <http://www.education.gov.uk/schools/teachingandlearning/pedagogy/b00192950/encouraging-reading-for-pleasure/what-the-research-says-on-reading-for-pleasure>).
- Department for Education Education Standards Research Team (2012) *Research Evidence on Reading for Pleasure* (London, DfE <http://www.education.gov.uk/schools/teachingandlearning/pedagogy/b00192950/encouraging-reading-for-pleasure/what-the-research-says-on-reading-for-pleasure>).
- Dodgeon, B. (2008) *Guide to the Dataset: BCS70 16 year follow up: APU arithmetic test* (London, Centre for Longitudinal Studies).
- Douglas, J. W. B. (1964) *The Home and the School* (London, MacGibbon and Kee).
- Duncan, G. J., Bergman, L., Duckworth, K., Kokko, K., Lyyra, A.-L., Metzger, M., Pulkkinen, L. & Simonton, S. (2012) The Role of Child Skills and Behaviors in the Intergenerational

Transmission of Inequality: A Cross-National Study, in: J. Ermisch, M. Jäntti & T. Smeeding (Eds) *From Parents to Children: The Intergenerational Transmission of Advantage* (New York, Russell Sage Foundation).

Elliott, C. D., Murray, D. J. & Pearson, L. S. (1979) *British Ability Scales* (Slough, NFER).

Elliott, J. & Shepherd, P. (2006) Cohort Profile: 1970 British birth cohort (BCS70), *International Journal of Epidemiology*, 35(4), 836-843.

Evans, M. D. R., Kelley, J., Sikora, J. & Treiman, D. J. (2010) Family scholarly culture and educational success: Books and schooling in 27 nations, *Research in social stratification and mobility*, 28, 171-197.

Feinstein, L. (2003) Inequality in the Early Cognitive Development of British Children in the 1970 Cohort, *Economica*, 70(1), 73-97.

Feinstein, L. (2004) Mobility in pupils' cognitive attainment during school life, *Oxford Review of Economic Policy*, 20(2), 213-229.

Fogelman, K. R. & Goldstein, H. (1976) Social Factors Associated with Changes in Educational Attainment between 7 and 11 Years of Age, *Educational Studies*, 2, 95-109.

Ganzeboom, H. (1982) Explaining Differential Participation in High-Cultural Activities - A Confrontation of Information-Processing and Status-Seeking Theories, in: W. Raub (Ed) *Theoretical Models and Empirical Analyses: Contributions to the Explanation of Individual Actions and Collective Phenomena* (Utrecht, E.S. Publications), 186-205.

Georg, W. (2004) Cultural capital and social inequality in the life course, *European Sociological Review*, 20(4), 333-344.

Gipps, C. & Murphy, P. (1994) *A fair test? : assessment, achievement and equity* (Milton Keynes, Open University Press).

Godfrey Thompson Unit (1978) *Edinburgh Reading Test* (Sevenoaks, Hodder and Stoughton).

Golding, J. (1975) *The 1970 Birth Cohort 5-Year Follow-up: Guide to the dataset* (University of Bristol, Institute of Child Health).

Goldthorpe, J. & McKnight, A. (2006) The economic basis of social class, in: S. Morgan, D. B. Grusky & G. S. Fields (Eds) *Mobility and Inequality: Frontiers of research from sociology and economics* (Stanford, Stanford University Press).

Goldthorpe, J. H. (1997) The 'Goldthorpe' class schema: some observations on conceptual and operational issues in relation to the ESRC review of government social classifications, in: D. Rose & K. O'Reilly (Eds) *Constructing Classes: Towards a New Social Classification for the UK* (Swindon, ESRC/ONS).

Goodenough, F. L. (1926) *The measurement of intelligence by drawings* (New York, World Book Company).

Halsey, A. H., Heath, A. & Ridge, J. (1980) *Origins and Destinations: family, class, and education in modern Britain* (Oxford, OUP).

Harris, D. B. (1963) *Children's drawings as measures of intellectual maturity* (New York, Harcourt, Brace and World).

Hart, B. & Rinsley, T. R. (1995) *Meaningful differences in the everyday experiences of young American children* (Baltimore, M.D., Paul H. Brookes).

Hill, V. (2005) Through the past darkly: A review of the British Ability Scales Second Edition, *Child and Adolescent Mental Health*, 10, 87-98.

Hirsch, E. D. (1983) Cultural Literacy, *The American Scholar*, 52(2), 159-169.

Jaeger, M. M. (2011) Does Cultural Capital Really Affect Academic Achievement? New evidence from combined sibling and panel data, *Sociology of Education*, 84(4), 281-298.

Jerrim, J. & Vignoles, V. (2013) Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes, *Journal of the Royal Statistical Society (Series A)*, 176(4), 887-906.

Lareau, A. (2003) *Unequal childhoods : class, race, and family life* (Berkeley, University of California Press).

Little, R. J. A. & Rubin, D. B. (1987) *Statistical analysis with incomplete data*, (New York, Wiley).

- Mostafa, T. & Wiggins, D. (2014) Handling attrition and non-response in the 1970 British Cohort Study, *CLS Working Paper*, 2014/2.
- National Audit Office (2008) *Skills for Life: Progress in improving adult literacy and numeracy* (London, House of Commons Public Accounts Committee).
- Nisbet, J. (1953) Family Environment and Intelligence, *Eugenics Review*, XLV, 31-42.
- Parsons, S. (2014) Childhood cognition in the 1970 British Cohort Study, *CLS Data Note*.
- Richards, M. & Sacker, A. (2003) Lifetime antecedents of cognitive reserve, *Journal of clinical and experimental neuropsychology*, 25(5), 614-624.
- Ritchie, S. J., Bates, T. C. & Plomin, R. (2014 in press) Does Learning to Read Improve Intelligence? A Longitudinal Multivariate Analysis in Identical Twins From Age 7 to 16, *Child development*.
- Rutter, M., Tizard, J. & Whitmore, K. (1970) *Education, Health and Behaviour* (London, Longmans).
- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999) Stereotype threat and women's math performance, *Journal of experimental social psychology*, 35(1), 4-28.
- Sullivan, A. (2001) Cultural Capital and Educational Attainment, *Sociology*, 35(4), 893-912.
- Sullivan, A. (2002) Bourdieu and Education: How Useful is Bourdieu's Theory for Researchers?, *Netherlands' Journal of Social Sciences*, 38(2), 144-166.
- Sullivan, A. (2007) Cultural Capital, Cultural Knowledge, and Ability, *Sociological Research Online*, 12(6).
- Sullivan, A., Ketende, S. & Joshi, H. (2013) Social class and inequalities in early cognitive scores, *Sociology*, 47(6), 1187-1206.
- Taylor, B. M., Frye, B. J. & Maruyama, G. M. (1990) Time spent reading and reading growth, *American Educational Research Journal*, 27(2), 351-362.
- Taylor, M. (2011) Life course outcomes of cultural practices - instrumental benefits?, *paper presented at British Sociological Association annual conference* (London School of Economics, London, England).
- Twist, L., Schagan, I. & Hodgson, C. (2007) *Progress in International Reading Literacy Study (PIRLS): Reader and Reading National Report for England 2006* (Slough, NFER and DCSF).

Tables and Figures

Table 1: Age 16 Arithmetic and Vocabulary Scores – Percentage points

	N	Min.	Max.	Mean	Std. Deviation
Arithmetic	3583	0	100	61.6	19.4
Vocabulary	3583	0	96	53.5	14.9

Table 2: Descriptive statistics for categorical regressors

	Imputed %	Original N	% missing
Child sex			
Male	46.1	1653	
Female	53.9	1930	
Missing	-	0	0
Parental social class (NS-SEC)			
Managerial / professional	29.9	1044	
Intermediate	28.1	835	
Routine and manual	38.5	1322	
Not working	3.6	128	
Missing	-	254	7.1
Parental highest qualification (Age 10)			
No qualifications	27.1	972	
Vocational/Other	19.1	684	
O-Levels/A-levels	34.4	1113	
Degree+	19.3	692	
Missing	-	122	3.4
Parental reading (Age 16 survey)			
At least one parent reads books	88.7	2127	
Missing	-	1040	29.0
Parental reading problems (Maternal)			
At least one parent has reading problem	9.6	345	
Missing	-	970	27.1
Reading materials available in home			
Tabloid paper	43.6	1562	
Broadsheet paper	11.8	423	
Missing	-	970	27.1
How often reads books (Age 10)			
Often	54.9	1962	
Sometimes	39.8	1034	
Never or hardly ever	5.3	161	
Missing	-	426	11.9
How often reads newspapers (Age 16)			
Rarely or never	11.6	414	
Less than once a week	7.1	256	

Once a week	20.6	562	
More than once a week	60.7	2130	
Missing	-	221	6.2
How often reads books (Age 16)			
Rarely or never	33.3	1192	
Less than once a week	20.6	736	
Once a week	16.9	472	
More than once a week	29.2	967	
Missing	-	216	6.0
Plays musical instrument			
Age 10	46.5	1662	
Missing	-	436	
Age 16	24.0	861	
Missing	-	249	6.9

Table 3: Descriptive statistics for continuous regressors

	Imputed mean	Original N	% missing
Birth order	1.8	3163	11.7
Number of days child read to in past week (age 5)	4.5	2872	19.8

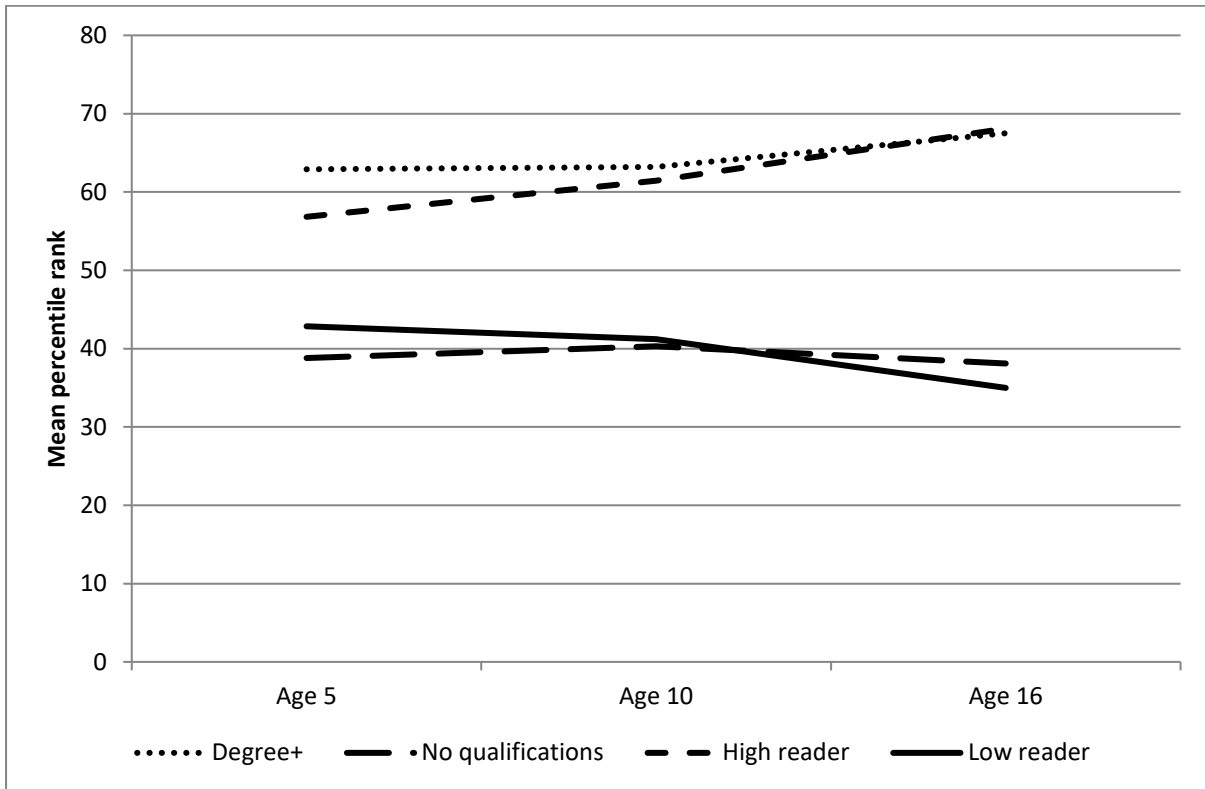
Table 4: Age 5 and 10 test scores – Imputed where missing

	Original N	Min.	Max.	Mean.	Std. Deviation	% missing
<i>Age 5</i>						
Copying designs	2968	0	8	4.9	1.8	17.2%
English picture vocabulary	2775	0	62	38.1	12.0	22.6%
Human figure drawing	2931	1	22	10.5	2.8	18.2%
Complete-a-profile	2876	0	16	6.9	3.6	19.7%
Schonell graded reading	2962	0	49	1.7	4.0	17.3%
<i>Age 10</i>						
Edinburgh Reading Test	2690	0	64	39.7	13.0	24.9%
Pictorial Language Score	2904	24	100	63.0	9.6	19.0%
Friendly Maths Test	2690	3	72	45.9	11.2	24.9%
Spelling score	2880	0	56	36.3	9.7	19.6%
BAS word definitions	2675	0	29	10.8	4.7	25.3%
BAS word similarities	2662	0	20	12.3	2.3	25.7%
BAS digit recall score	2671	7	34	22.5	3.8	25.5%
BAS Matrices	2666	0	28	16.2	4.9	25.6%

Table 5: Multivariate regression analysis of vocabulary and mathematics % scores (n for all models =3, 583)

	Model 1		Model 2		Model 3		Model 4	
	Maths	Vocab	Maths	Vocab	Maths	Vocab	Maths	Vocab
Intercept	57.7**	50.9**	54.6**	46.3**	35.6**	29.6**	57.6**	46.3**
Sex (Male)	0.6	-0.1	0.8	0.2	3.4**	2.9**	0.4	0.6
Social class (Ref = Manual/Routine)								
Managerial / professional	4.8**	3.9**	3.6**	2.6**	3.2**	2.1**	0.4	-0.1
Intermediate	2.2**	2.0**	1.8*	1.5**	1.7*	1.4**	0.2	0.3
Not working	0.6	0.8	1.3	1.4	1.7	1.2	0.5	0.5
Highest parental qual (Ref=None)								
Vocational/Other	3.9**	2.8**	3.4**	2.2**	3.1**	2.0**	0.9	0.3
O-Levels/A-levels	5.7**	5.6**	4.7**	4.5**	4.2**	3.9**	0.6	1.2**
Degree+	13.2**	12.6**	10.9**	9.7**	9.4**	8.2**	2.0*	2.5**
Birthorder	-2.1**	-2.2**	-1.8**	-1.9**	-1.6**	-1.6**	-0.8**	-1.0**
Age 5 - number of days child read to in past week			0.9**	1.2**	0.5**	0.9**	-0.2	0.3**
Age 16 - Parent(s) read books			1.3	1.6*	0.5	0.7	-0.9	-0.4
Age 16 - Parental reading problems			-4.9**	-2.9**	-4.1**	-2.2**	-1.2	0.0
Age 16 - Tabloid newspaper in home			0.2	-0.1	-0.9	-0.9*	-0.3	-0.4
Age 16 - Broadsheet newspaper in home			4.4**	4.9**	3.5**	3.7**	1.0	1.8**
Age 10 - How often reads books (Ref = Never or hardly ever)								
Often					13.9**	12.7**	2.6*	4.2**
Sometimes					10.3**	7.8**	3.4**	2.6**
Age 16 - How often reads newspapers (Ref = Rarely or never)								
Less than once a week					3.3*	2.0*	1.5	0.5
Once a week					3.2**	1.7*	2.1*	0.8
More than once a week					6.9**	4.9**	2.8**	1.8**
Age 16 - How often reads books (Ref = Rarely or never)								
Less than once a week					2.9**	2.7**	0.7	1.0*
Once a week					1.6	3.1**	0.4	2.2**
More than once a week					3.0**	6.5**	-0.4	4.0**
Age 10 - Plays musical instrument					2.3**	1.1*	0.1	-0.7
Age 16 - Plays musical instrument					1.1	1.5**	-0.3	0.4
Age 5 ability score							0.5	1.0**
Age 10 ability score							12.6**	9.2**
Adjusted R²	0.10	0.15	0.12	0.20	0.18	0.31	0.48	0.61

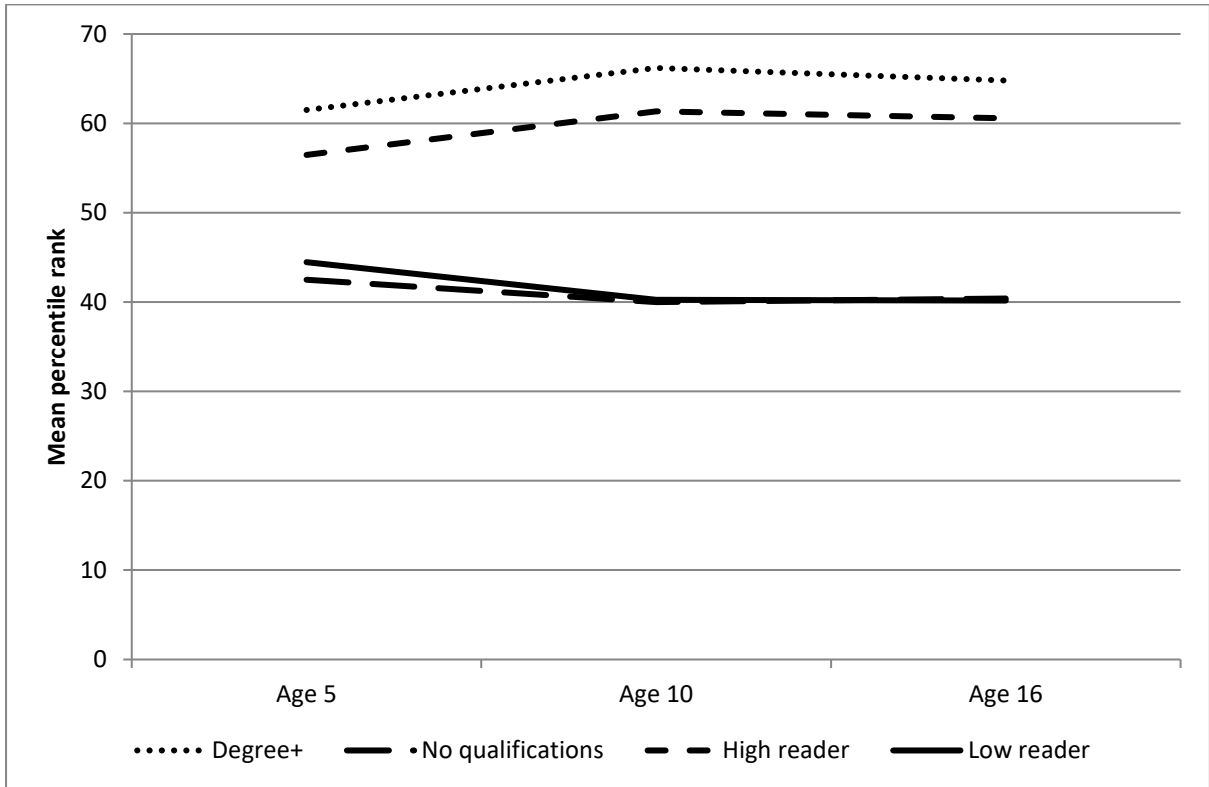
Figure 1: Vocabulary trajectories



N=3583, data imputed where missing.

Age 5: English Picture Vocabulary; Age 10: the vocabulary subscale of the Pictorial Language Score; Age 16: Vocabulary score.

Figure 2: Mathematics trajectories



N=3583, data imputed where missing.

Age 5: Copying Designs; Age 10: Friendly Maths test; Age 16: Arithmetic test.

Appendix Table A1: Odds ratios of Logit response models for BCS70 1986 school tests***Gender (reference: Men)**

Women 1.37*** (0.057)

Marital status (reference: Single)

Married 1.90*** (0.265)

Mother lives in London in 1970 (reference: not in London)

In London 0.60*** (0.043)

Parity (reference: 0)

1 0.83*** (0.042)

2 0.75*** (0.05)

3+ 0.55*** (0.045)

Lactation (reference: attempted)

No 0.94 (0.042)

Mother's age at Delivery (reference: less than 20)

20-24 1.25* (0.107)

25-29 1.35*** (0.122)

30-34 1.61*** (0.163)

35 or more 1.51*** (0.187)

Mother's age at completion of education (reference: 14 or less)

15 1.2 (0.133)

16 1.38** (0.162)

17 1.27 (0.166)

18 + 1.29 (0.166)

Father's social class (reference: SC 1)

SC2 1 (0.104)

SC3 non-manual 1 (0.108)

SC3 manual 0.84 (0.086)

SC4 0.76* (0.087)

SC5 0.55*** (0.079)

Other 0.82 (0.116)

Father's age at completion of education (reference: 14 or less)

15 0.99 (0.095)

16 1.04 (0.112)

17 1.31* (0.16)

18 or more 1.02 (0.116)

N 15270**pseudo R^2** 0.024Exponentiated coefficients; Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*With thanks to Tarek Mostafa. This logit response model can be compared to identical models for each wave of BCS70 in Mostafa and Wiggins (2014).

ⁱ We treat 'not working' as a separate category as NS-SEC is not available in BCS70 prior to 1980, so non-workers could not be categorised according to their previous NS-SEC.