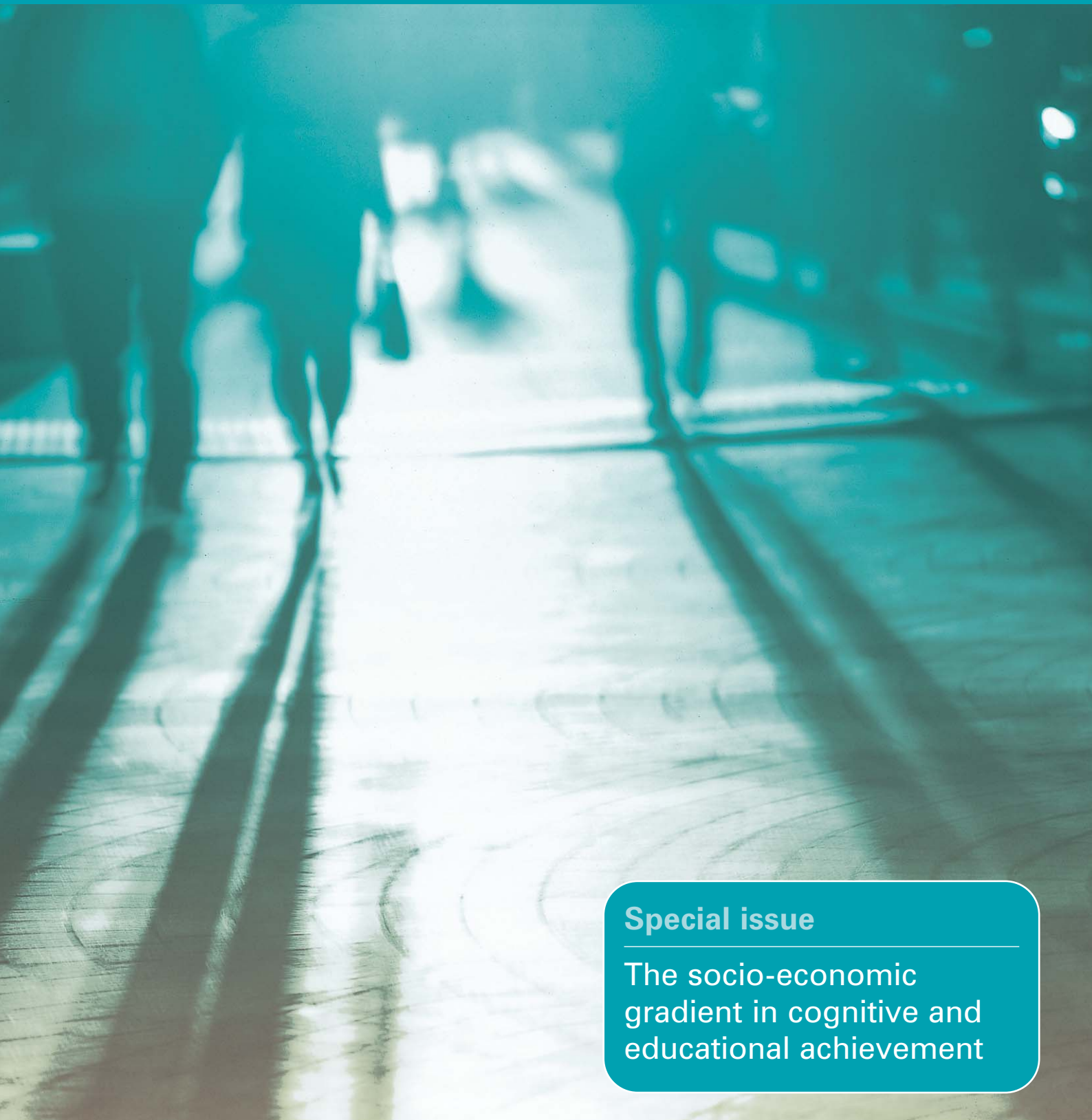


# Longitudinal and Life Course Studies: International Journal



## Special issue

The socio-economic  
gradient in cognitive and  
educational achievement

Published by

**SLLS** Society for Longitudinal  
and Life Course Studies

Supported by

**Longview**  
Promoting Longitudinal  
and Life Course Studies

# EDITORIAL BOARD

---

## EDITORIAL COMMITTEE

### Executive Editor

**John Bynner**, *Longview and Institute of Education, UK*

### Health Sciences

**Section Editor - Michael Wadsworth**, *Longview, MRC Unit for Lifelong Health and Ageing London, UK*

**Associate Editor - David Blane**, *Imperial College Medical School, UK*

### Social and Economic Sciences

**Section Editor - Robert Erikson**, *University of Stockholm, Sweden*

**Associate Editors - Paul Gregg**, *University of Bristol, UK*

- **John Hobcraft**, *University of York, UK*

- **Karl Ulrich Mayer**, *Yale University, USA*

### Statistical Sciences and Methodology

**Section Editor - Harvey Goldstein**, *University of Bristol, UK*

**Associate Editor - Bianca de Stavola**, *London School of Hygiene and Tropical Medicine, UK*

### Development and Behavioural Sciences

**Section Editor - Barbara Maughan**, *Institute of Psychiatry, Kings College, UK*

**Associate Editor - Lars Bergman**, *University of Stockholm, Sweden*

## BOARD MEMBERS

**Mel Bartley**, *University College London, UK*; **Paul Boyle**, *University of St. Andrews, UK*; **Nick Buck**, *University of Essex, UK*; **Richard Burkhauser**, *Cornell University, USA*; **Jane Costello**, *Duke University, USA*; **Tim Croudace**, *University of Cambridge, UK*; **George Davey-Smith**, *University of Bristol, UK*; **Lorraine Dearden**, *Institute for Fiscal Studies, UK*; **Ian Deary**, *University of Edinburgh, UK*; **Glen Elder**, *University of North Carolina, USA*; **Peter Elias**, *University of Warwick, UK*; **Leon Feinstein**, *Cabinet Office, UK*; **Antony Fielding**, *University of Birmingham, UK*; **Andy Furlong**, *University of Glasgow, UK*; **Frank Furstenberg**, *University of Pennsylvania, USA*; **John Gray**, *University of Cambridge, UK*; **Rebecca Hardy**, *University College London, UK*; **Walter Heinz**, *University of Bremen, Germany*; **Marjo-Ritta Jarvelin**, *Imperial College London, UK*; **Heather Joshi**, *Institute of Education, UK*; **Kathleen Kiernan**, *University of York, UK*; **Harvey Krahn**, *University of Alberta*; **Di Kuh**, *University College London, UK*; **Carli Lessof**, *National Centre for Social Research, UK*; **Dean Lillard**, *Cornell University, USA*; **Jean Martin**, *University of Oxford, UK*; **Steve Machin**, *London School of Economics, UK*; **Robert Michael**, *University of Chicago, USA*; **Sir Michael Marmot**, *University College, London*; **Scott Montgomery**, *Karolinska University Hospital, Sweden*; **Jeylan Mortimer**, *University of Minnesota, USA*; **Brian Nolan**, *University College Dublin, Ireland*; **Lindsay Paterson**, *University of Edinburgh, UK*; **Ian Plewis**, *University of Manchester, UK*; **Chris Power**, *Institute of Child Health, UK*; **David Raffe**, *University of Edinburgh, UK*; **Steve Reder**, *University of Portland, USA*; **Marcus Richards**, *Medical Research Council, UK*; **Amanda Sacker**, *University of Essex, UK*; **Ingrid Schoon**, *Institute of Education, UK*; **John Schulenberg**, *University of Michigan, USA*; **Jackie Scott**, *University of Cambridge, UK*; **Rainer Silbereisen**, *University of Jena, Germany*; **Chris Skinner**, *University of Southampton, UK*; **Heike Solga**, *Social Science Research Centre, Germany*; **Håkan Stattin**, *Orebro University, Sweden*; **Fiona Steele**, *University of Bristol, UK*; **Alice Sullivan**, *Institute of Education, UK*; **Kathy Sylva**, *University of Oxford, UK*; **Gert Wagner**, *German Institute for Economic Research, Germany*; **Chris Whelan**, *Economic & Social Research Institute, Ireland*; **Richard Wiggins**, *Institute of Education, UK*; **Dieter Wolke**, *University of Warwick, UK*

---

## Open Journal System

The **LLCS** journal is produced using the [Open Journal System](#), which is part of the [Public Knowledge Project](#). OJS is open source software made freely available to journals worldwide, for the purpose of making open access publishing a viable option for more journals. Already, more than 8,000 journals around the world use this software.

## Copyright Notice

Authors who publish with Longitudinal and Life Course Studies agree to the following terms:

- Authors retain copyright and grant the Journal right of first publication with the work, simultaneously licensed under a [Creative Commons Attribution License](#) that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.
- Following first publication in this Journal, Authors are able to enter into separate, additional contractual arrangements for the non-exclusive distribution of the journal's published version of the work (e.g., post it to an institutional repository or publish it in a book), with an acknowledgement of its initial publication in this journal, provided always that no charge is made for its use.
- Authors are permitted and encouraged to post their work online (e.g. in institutional repositories or on their own website) prior to and during the submission process, as it can lead to productive exchanges, as well as earlier and greater citation of published work.

# CONTENTS

---

## PAPERS

- 1-18** Children's educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood  
*Alissa Goodman, Paul Gregg, Elizabeth Washbrook*
- 19-40** The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study  
*Lorraine Dearden, Luke Sibieta, Kathy Sylva*
- 41-58** The socio-economic gradient in child outcomes: the role of attitudes, behaviours and beliefs  
*Paul Gregg, Elizabeth Washbrook*
- 59-76** The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16  
*Haroon Chowdry, Claire Crawford, Alissa Goodman*
- 77-93** Explaining the socio-economic gradient in child outcomes: the inter-generational transmission of cognitive skills  
*Claire Crawford, Alissa Goodman, Robert Joyce*

**94-97** **NEWS, EVENTS AND RESOURCES**

## SPECIAL SUPPLEMENT

### SLLS 2010 Conference Abstracts

The Editors are most grateful to the Nuffield Foundation for supporting the journal through its development phase.

Published by the Society for Longitudinal and Life Course Studies

Address: 90 Mansfield Road, London, NW3 2HX

Email: [info@longviewuk.com](mailto:info@longviewuk.com)

<http://www.longstudies.longviewuk.com/index.shtml>

# AUTHOR GUIDELINES SUMMARY

---

## Submission of Papers

All papers, written in the English language, should be submitted via the LLCS website as a Microsoft Word 2003 file, or in a compatible format. If there is a good reason why this is not possible, authors are requested to contact the [Journal Manager](#) before submitting the paper. All subsequent processes involving the author are carried out electronically via the website.

## Preparation of Texts

**Length.** The paper should normally be approximately 5000 words, with longer papers also accepted up to a maximum of 7000 words. This word count excludes tables, figures and bibliography.

**Font and Line Spacing.** Please use Calibri (or similar sans serif) font, 12pt, with 1.5 line spacing in all main text, single space in figure or table captions.

**Page Layout.** All text should be justified to the left hand margin (all margins of at least 2.5cm) with no indents at the start of paragraphs and a line space separating paragraphs. All headings and sub-headings used within the body of the paper should also be left justified. Please do **NOT** use automatic text formatting.

**Weblinks.** To help our readers to look up cited references or other information available on the web, authors should ensure that all such references are activated.

**Ensuring an Anonymous (Blind) Review.** Please submit papers with a full detailed title page. Once a paper has been submitted to LLCS via the website, it will be 'anonymised' by the removal of all author name(s) and institution names on the title page, and any identifying electronic document properties will also be removed. Authors do not need to remove their name(s) from the main text or references but any reference to their work or themselves should be expressed in the third person.

**Abstract.** The abstract (no subheads or paragraphs) should be no more than 250 words (not part of the main word count).

**Keywords.** These should be included just below the author list (minimum 3 maximum 10).

**Abbreviations.** Words to be abbreviated should be spelt out in full the first time they appear in the text with the abbreviations in brackets. Thereafter the abbreviation should be used.

**References.** Please use the Harvard Style and refer to examples in the full Guidelines

*Authors not complying with these reference guidelines will be asked to make all the necessary alterations themselves, if their paper is accepted for publication.*

**Notes.** As a general rule, supplementary notes should be avoided, but if thought to be essential, they should *not* appear on the page as 'Footnotes', but instead, be included as Endnotes.

**Supplementary material.** Supplementary material may be uploaded with the submission, and if the paper is published, will be visible to the reader via a link on the RHS of the screen.

**Submissions containing graphs, tables, illustrations or mathematics.** All graphs, tables and illustrations should be embedded in the submitted text, and have clear, self-explanatory titles and captions. Mathematical expressions should be created in Word 2003 (or a compatible package) with equation editor 3.0, unless the author has good reason to use other software, in which case please contact the [Journal Manager](#). All biological measures should be reported in SI units, as appropriate, followed, in the text, by traditional units in parentheses.

**Author Citation.** If the paper is accepted for publication, a version of the paper with all authors cited in full on the title page will be used. Only individuals who have contributed substantially to the production of the paper should be included.

**Copy Editing.** All accepted manuscripts are subject to copy editing, with reference back to author with suggested edits and queries for response. Particular care should be taken by the author to check references.

**Proofs.** The corresponding author will be asked to view a layout proof of the article on the website and respond with final amendments within 3 days.

# Children's educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood in the UK

**Alissa Goodman**

Institute for Fiscal Studies

**Paul Gregg** [P.Gregg@bristol.ac.uk](mailto:P.Gregg@bristol.ac.uk)

Centre for Market and Public Organisation, University of Bristol

**Elizabeth Washbrook**

Centre for Market and Public Organisation, University of Bristol

(Received November 2010 Revised January 2011)

**Key words:** education, inter-generational transmission, socio-economic gap

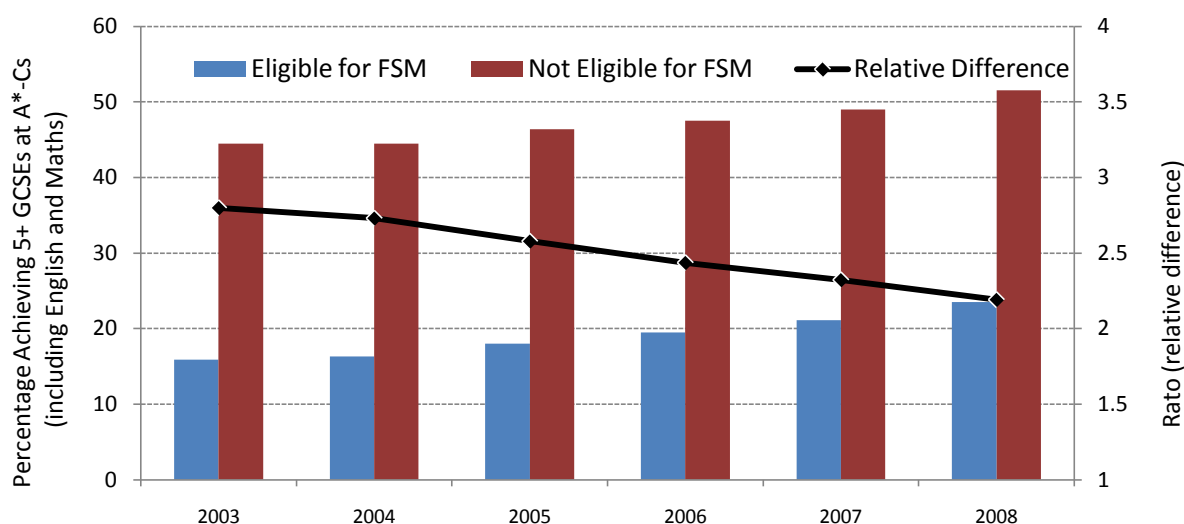
JEL codes: I21

## 1. Introduction to this Special Issue of Longitudinal and Life Course Studies (LLCS)

It is well known that children growing up in poor families emerge from our schools with substantially lower levels of educational attainment. The only regularly published government statistics are for attainment gaps between those on free school meals (FSM), broadly an indicator of receipt of major workless benefits, and the rest of the population. As of 2008, less than a quarter of children from families eligible for free school meals obtained the widely used UK benchmark of sufficient achievement at the secondary school level for continuing in academic education to age 18 and potentially going on to university (defined as five General Certificates of Secondary Education – GCSEs – at grades C or above including English and

maths). This compares with just over a half of their richer peers, not eligible for free school meals. As Figure 1 shows, the proportion of both groups achieving this benchmark has risen in recent years. This growth has been slightly faster amongst the FSM-eligible group, such that the relative gap between these groups has fallen over this period (as indicated by the black line). Whilst gaps in achievement at age 16 by family income have started to close over the last decade (Gregg and Macmillan 2010) these gaps remain large, and since educational qualifications are such a strong determinant of later life income and opportunities, such achievement gaps create a major obstacle to social mobility, which is of strong public concern.

**Figure 1. Percentage of children achieving 5+ GCSEs at A\*-C (including English and maths) by FSM eligibility, 2003-2008.**



**Sources:** 2003-2007, Department for Children, Schools and Families (2009), *Deprivation and Education: the evidence on pupils in England, Foundation Stage to Key Stage 4*, <http://www.dcsf.gov.uk/research/data/uploadfiles/DCSF-RTP-09-01.pdf>; 2008, DCSF Departmental Report 2009, <http://publications.dcsf.gov.uk/eOrderingDownload/DCSF-Annual%20Report%202009-BKMK.PDF>

Policymakers and commentators on policy thinking have long struggled to understand precisely what the sources of these educational inequalities are, and in turn to find policies that will reduce them. The influential work of Feinstein (2003, 2004) has shown how attainment gaps appear early in childhood but then continue to widen through childhood. The large magnitude of the gaps early in life have led many to argue for an increased focus on the early years (for example, Esping-Andersen 2005; Carneiro and Heckman 2003). However, a number of studies have highlighted how increasing the school leaving age results in substantial earnings gains for those forced to continue education, who are largely drawn from lower social backgrounds (e.g. Meghir and Palme 2005 or Harman and Walker 1995). This highlights the potential for policy development before and during the school years and the support for post-compulsory learning in addressing socio-economic attainment gaps.

Our aim in this LLCS Special Issue (which comprises this overview paper and 4 studies) is to assess the empirical relevance of a particular set of factors for explaining the socio-economic gradient in cognitive and educational

achievement over the course of childhood. The factors in which we are interested can be grouped together under the broad umbrella term 'transmission mechanisms'. This diverse set of influences ranges, for example, from parenting styles during the very earliest stages of life and parental cognitive and social abilities in general, through to parental aspirations for educational success in the primary school years, and teenage engagement in risky and positive behaviours during adolescence. What unites them is they are all factors that have been proposed as 'proximal' influences on children's developmental outcomes, in the language of ecological models of child development (e.g. Bronfenbrenner and Morris 1998), as distinct from 'distal' influences such as socio-economic and demographic characteristics. Loosely speaking, proximal factors are those that drive the observed association between a distal factor and an outcome. Children from disadvantaged backgrounds perform worse on achievement tests because the contexts, environments and interactions experienced by such children differ from those experienced by better-off children. If all such contexts and environments are observed and accounted for by

the researcher, then socio-economic background by definition will have no 'direct' association with the outcome.

As a set of observational studies, the research documents the strength of a wide variety of mechanisms that potentially generate the observed pattern of outcome differences between disadvantaged and more affluent children. It is well-known that work of this kind cannot establish the causal nature of the relationship between mechanisms and outcomes, as the possibility of correlated unmeasured factors is always present. We address this as far as possible, through the use of rich conditional models that proxy unmeasured influences with indicators of distal family characteristics, school performance and prior test scores. Evaluations of existing policy initiatives reviewed in this paper provide additional evidence on the likelihood that the mechanisms we identify are amenable to policy intervention.

Section 2 introduces the four data sources used in the chapters, and charts the observed relationship between parental socio-economic background and educational outcomes at different ages. Section 3 provides details of the common measure of socio-economic position used in the four following studies. Section 4 discusses the conceptual framework and gives an overview of how it is operationalized, while Section 5 gives a general formulation of the empirical models estimated in each paper. Section 6 gives a summary of the key findings and considers their implications for policy formation. Section 7 investigates whether evidence from recent policy initiatives supports a causal interpretation of the associations we identify, and Section 8 offers a brief conclusion.

## 2. The four studies

In order to study these factors, we are fortunate enough to be able to make use of four new and rich sources of data, capturing groups of children growing up in the UK today (described in Table 1). These studies surveyed children and their families at various points in time from early childhood through to mid-adolescence, and all contain high quality information on children's cognitive achievement at different ages, and the processes, environments and contexts experienced by family members. The developmental stage covered by each survey is different, meaning that different types of transmission mechanisms will be relevant in each case, but to facilitate comparability, we impose common definitions on some of the key concepts of interest and employ a common modelling framework. In terms of definitions, we derive an index of socio-economic position (SEP) using a common methodology in all four studies, and divide children in each survey into quintiles (fifths) on the basis of this measure. Hence the relative position of the most- and least-disadvantaged fifth of children can be compared in a systematic way across studies. Second, all outcomes measures are converted to the common metric of percentile scores or ranks. As is shown in Table 1, the outcomes measured differ across childhood stages, and this technique provides a way to assess the relative size of the SEP gap as children age. The table also provides a brief introduction to the datasets used in each of the four studies, the ages over which children are followed in our analysis and the key outcome variables. Further details of the specific studies and definitions of other variables are provided in each article.

Table 1. Overview of the datasets

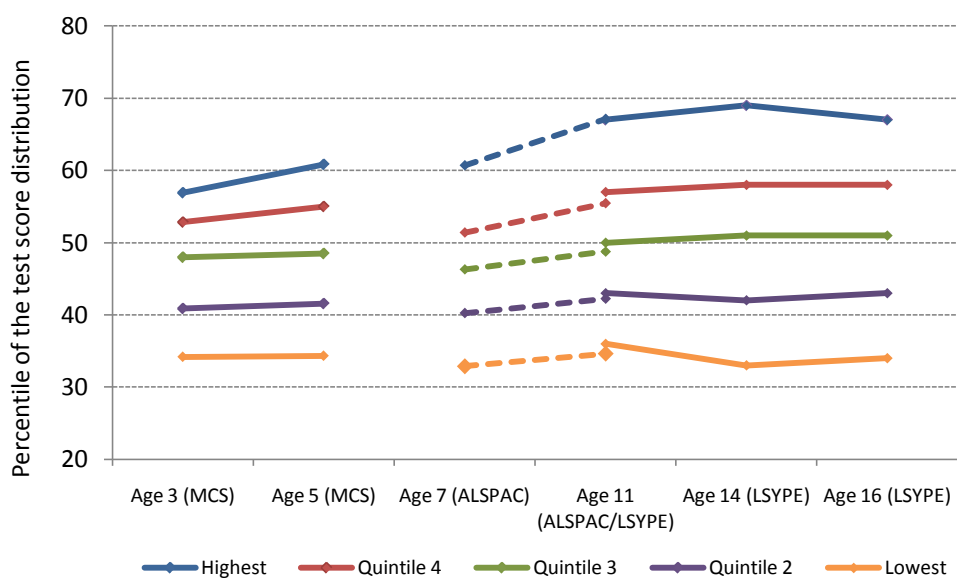
	<b>Study 1: Pre-school</b>	<b>Study 2: Primary school</b>	<b>Study 3: Secondary school</b>	<b>Study 4: Across the ages</b>
<b>Data source</b>	Millennium Cohort Study (MCS)	Avon Longitudinal Study of Parents and Children (ALSPAC)	Longitudinal Study of Young People in England (LSYPE)	Children of the British Cohort Study (BCS)
<b>Coverage</b>	Nationally representative sample of children born in the United Kingdom	Census of all children born in the old Avon area of England (a region in the southwest of England located around the cities of Bristol and Bath)	Nationally representative sample of 14-year olds in England	Children in 2004 of a nationally representative sample of adults born in one week in 1970
<b>Dates of birth of cohort members</b>	September 2000 to January 2002	April 1991 to December 1992	September 1989 to August 1990	June 1987 to April 2005
<b>Total cohort members</b>	19,517	13,988 children alive at 1 year	15,770	11,083
<b>Sample size used in analysis</b>	11,054	7,972	13,343	3,416
<b>Focal outcome measure</b>	British Ability Scales (BAS) Naming Vocabulary score, age 5	Average of standardized Key Stage 2 (KS2) scores in English, maths and science, age 11	Average of point scores in eight best GCSE (or vocational equivalent) exam results (KS4), age 16	British Ability Scales (BAS) Vocabulary and Early number concepts, age 3-5 BAS Word reading, Spelling and Number skills, age 6-16
<b>Prior outcome measures</b>	(BAS) Naming Vocabulary score, age 3	Average of standardized Key Stage 1 (KS1) scores in reading, writing and maths, age 7	Average of standardized Key Stage 2 (KS2) scores in English, maths and science, age 11 Average of standardized Key Stage 3 (KS3) scores in English, maths and science, age 14	



Figure 2, drawn from the three longitudinal cohorts, summarizes the average percentile achievement score of children in each SEP quintile group (to be defined more fully below) across ages and cohorts. The picture is that educational deficits emerge early in children's lives, even before entry into school, and widen throughout childhood. Even by the age of 3, there is a considerable gap in cognitive test scores between children in the

poorest fifth of the population, compared to those from better off backgrounds. The poorest children average at the 34<sup>th</sup> percentile of overall attainment, whilst the richest are at the 57<sup>th</sup> percentile on average, a gap of 23 percentile points. This gap gets wider as children enter and move through the schooling system, especially in the primary school years. At age 14 this gap is 36 percentile points, before narrowing a little by age 16.

**Figure 2. Cognitive achievement outcomes by socio-economic position quintile, across surveys and ages.**



**Notes:** Children in each survey are divided into fifths, ranked according to a constructed measure of socio-economic position based on their parents' income, social class, housing tenure, and a self-reported measure of financial difficulties. The chart plots the average cognitive achievement measures for each group from the ages of 3 through to 16.

Figure 2 conveniently summarizes the patterns we wish to explain in three of the papers, but it is important to not to interpret it in terms of changes in a common outcome for a single group of individuals as they age. Cohorts vary over those born in 1989/90, 1991/2, and 2000/1, and over children born in the Avon region, England or the whole United Kingdom. Further, the outcomes vary from a test of purely verbal ability for children age 3 and 5, to "Key Stage" results in the core subjects of English, maths and (sometimes) science for children age 7 through 14, to GCSE results in eight different subjects (including vocational courses) at age 16. Nevertheless is notable that magnitudes of the SEP gaps at age 5 in the Millennium Cohort Study (MCS) cohort line up closely with those at age 7 in the Avon Longitudinal Study of Parents And Children

(ALSPAC) cohort, and similarly for the gaps in age 11 outcomes in the ALSPAC and Longitudinal Study of Young People in England (LSYPE) cohorts. The Children of the British Cohort Study (BCS) dataset differs from the three longitudinal cohorts, in that it surveys a group of children at a single point in time when they are all different ages. The appropriate methods for dealing with data in this format are discussed fully in the paper, but it is clear that the BCS study can only estimate an 'average' SEP gradient in the cognitive outcome that is invariant to the child's age. The unique addition for the BCS study is that both the mother and her children have undertaken similar cognitive tests in their childhood and hence offer an inter-generational perspective.

### 3. Definition of socio-economic position

For each of our project strands we have constructed an index of socio-economic position (SEP) that is designed to be as common as possible across strands. Its definition is described in detail here in order to avoid repetition in the substantive articles. The index seeks to capture the longer-term material resources of the household, and is constructed from the following variables:

- Log equivalised household income (averaged across as many points in time as possible, depending on the survey used)
- Reported experience of financial difficulties
- Mother's and father's occupational class
- Housing tenure

The index is constructed using principal-components analysis, and individuals are then placed into quintiles (fifths) of the population ranked by this measure. This approach is likely to give a more accurate classification of the family's long-term social position than measures taken at a single point in time (which will exhibit greater fluctuation), or that capture only one aspect of the family's material resources (such as paternal occupation). This is particularly important in a comparative study of this kind because the individual SEP indicators are measured at different stages of the life course in different studies. It is an approach that recognises that the resources or 'capitals' that convey advantage or disadvantage are multi-dimensional, and that the best and least well-off families exhibit clusters of a number of different kinds of characteristic (Galobardes, Lynch and Davey Smith 2007). The benefits of an approach that recognizes cumulative risk are illustrated by the fact that the outcome differentials between different SEP groups are larger than the differentials between groups defined by income or occupational class alone. Hence the combined measure discriminates those children likely to perform well or poorly, better than any single component indicator. An important exclusion from the combined measure, however, is parental educational qualifications. These are retained as separate control variables, enabling us to explore the distinction between education as an indicator of non-material parental resources – such as knowledge and cognitive ability – and material resources like earnings capacity.

In order to document SEP differentials that are comparable across studies, it is important to classify children according to their position in the underlying population rather than the (possibly non-

representative) estimation sample. Survey attrition and item non-response mean that disadvantaged children are likely to be under-represented in the final samples. Defining SEP groups on the basis of these samples, then, risks drawing the quintile boundaries too "high" relative to the population (and perhaps misclassifying differently across studies). For this reason, we conduct an imputation procedure that includes all children in the definition of the SEP quintile groups, even if they are subsequently dropped from the analysis due to missing data on outcomes or mechanisms. The imputation is conducted only on the SEP component variables and with the single aim of defining more representative quintile groups and consistency across the four studies in this regard.

The imputation procedure (the 'ice' command in Stata10) uses switching regression, an iterative multivariable regression technique, that predicts the likely values of missing items on the basis of the non-missing data (for details see van Buuren, Boshuizen and Knook 1999). Typically multiple imputation (MI) is used as an integral part of the analysis of interest and involves the creation of multiple datasets, each of which is analyzed separately before the averaging of the resulting series of parameter estimates. Since our aim is only to approximate the 'true' population quintile boundaries, we use only a single round of imputation that fills in likely values of missing SEP components on the basis of those observed. Maternal and paternal education are used in the imputation procedure to improve the prediction of missing values. The MI procedure, as we use it, gives us a single complete set of SEP indicators for every individual sampled. We then conduct polychoric principal components analysis (PCA) to combine the indicators into a summary index. This data reduction technique adapts standard principle components analysis in a manner that is appropriate for dealing with discrete variables such as parental occupation and housing tenure (see Kolenikov and Angeles 2004). It extracts a single component or index from the data, such that the index accounts for the maximum variation possible in the underlying indicators.

Although we investigate outcome differences over the full range of the SEP quintile groups, we focus our results on one key statistic: the difference in mean outcomes between the poorest 20% and the richest 20% of children according to the SEP index. We also explore results comparing the lowest and middle SEP quintiles and although the gaps in outcomes are

smaller, as we would expect, our conclusions regarding the explanatory power over different mechanisms are virtually identical.

#### 4. Conceptual framework

The finding that family income and poverty have strong consequences for child development, though to varying degrees and across different contexts, is well established (Blow et al 2006; Brooks-Gunn and Duncan 1997; Duncan and Brooks-Gunn 1997; Duncan and Brooks-Gunn 2000; Gregg and Machin 1998; Haverman and Wolfe 1995; Mayer 1997; Sylva et al 2008). This set of studies focuses on the mechanisms by which social and economic disadvantage may translate into child outcomes. As such, our quantitative analysis is related to a number of theoretical literatures, that hypothesise different routes through which advantage and disadvantage may be transmitted from parents to children.

The developmental psychology literature provides our primary conceptual framework for studying the effects of parental beliefs, attitudes and practices on children's cognitive and socio-emotional development. Bronfenbrenner and Morris (1998, 996) state that, "Throughout the life course, human development takes place through processes of progressively more complex reciprocal interaction between an active, evolving bio-psychological human organism and the persons, objects, and symbols in its immediate external environment. To be effective, the interaction must occur on a fairly regular basis over extended periods of *time*. Such enduring forms of interaction in the immediate environment are referred to as *proximal processes*." They further state that these proximal processes vary systemically with individual characteristics and contexts.

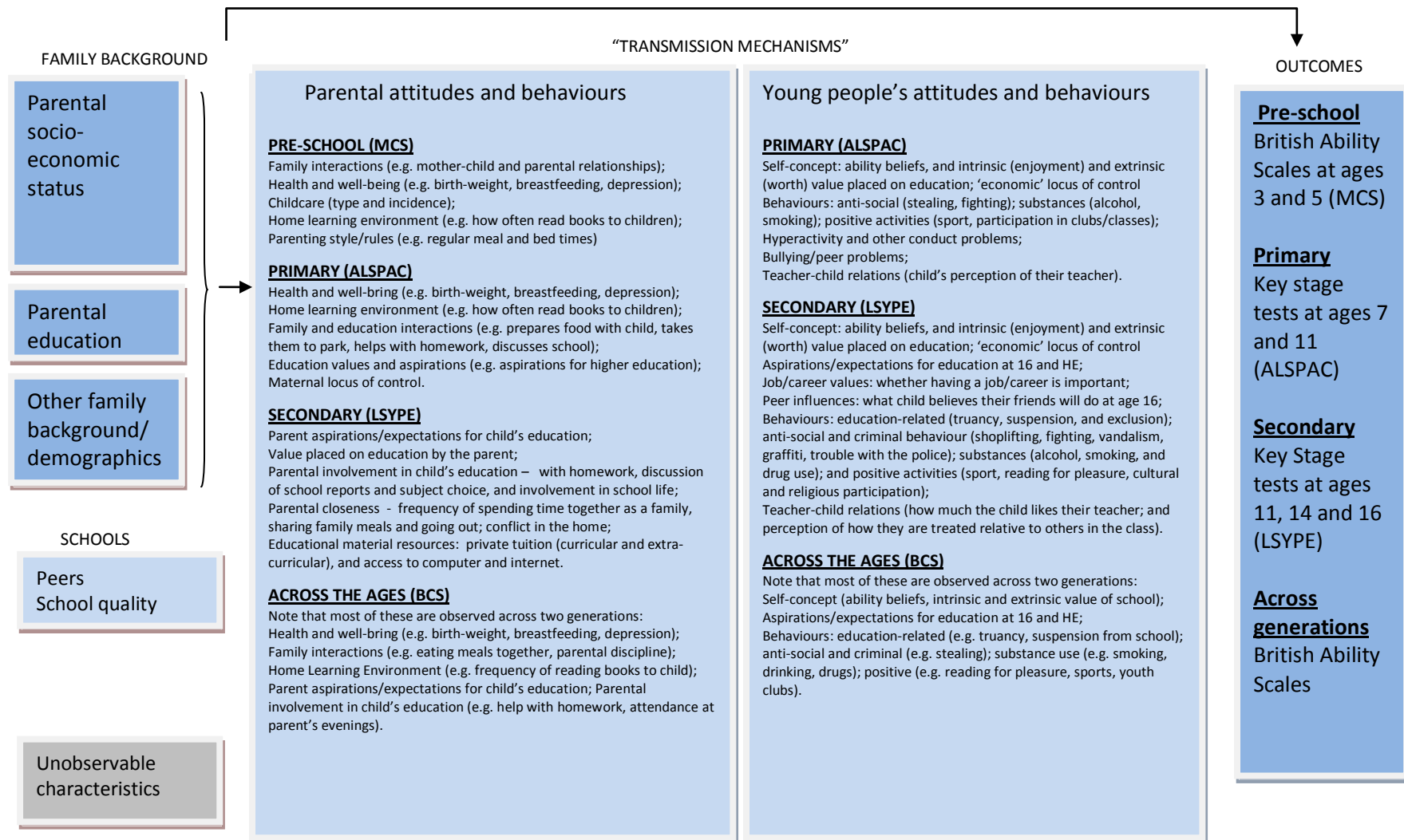
The sociological literature examines how family beliefs, attitudes and practices can be construed as social and cultural capital. For example, Bourdieu's work examines the role played by social and cultural capital in reproducing patterns of social and economic advantage and disadvantage (Bourdieu 1977a, 1977b; Bourdieu and Passeron 1977). Under the social capital theory, social relationships and networks create a resource which families can draw upon (Croll 2004). Cultural capital reflects the idea that "cultural experiences in the home facilitate children's adjustment to school and academic achievement, thereby transforming cultural resources into what [Bourdieu] calls cultural capital" (Lareau 1987, 74).

The economics literature has generally focused on theories of parental investment. For instance, in the Becker-Tomes model, parents invest in their children's education because they care about their children's future well-being, investing up until the point that marginal benefit equals marginal cost (Becker and Tomes 1986). Under this simple optimising theory, parental income should not influence child outcomes under the assumption that there are no credit constraints. Given that it seems unlikely that all families will be able to borrow against future earnings, poorer families may well not be able to invest optimal amounts (for more information on credit constraints see Carneiro and Heckman 2002). Beyond credit constraints, other economic models suggest that a lack of income may place significant strains on poorer families, preventing them from providing a rich home-learning environment, or reducing the quality of parenting (for a review of such models see Mayer 1997).

In the past, the developmental psychological literature has relied on observation, questionnaire and interview methods. It usually relies on research with small samples of about 100 families, and rich datasets. The sociological and economic literatures have usually relied on secondary analysis of existing datasets, many of which contain information on a limited number of variables. The current study attempts to use four datasets, that are both large in terms of sample sizes and rich in terms of variables, to bridge these three theoretical approaches/literatures.

The potential transmission mechanisms between socio-economic background and educational achievement identified by the literature are vast in scope, a scope mirrored by the range of explanatory variables available in our datasets. This presents the researcher with a trade-off between a framework that considers as many influences as possible in a simple and even-handed way, and one in which the inter-relationships between a smaller subset of variables are modelled explicitly, using theoretical insights from a specific branch of the literature. The latter is the approach most commonly used to address the question of socio-economic differentials, and it is vital for understanding the complex inter-relationships between parental and child characteristics and their evolution over time. Our analysis takes the former approach, which can be seen as complementary to the wealth of more narrowly-focused studies.

Figure 3. Transmission mechanisms between distal characteristics and cognitive achievement outcomes



The central panels in Figure 3 give some indication of the breadth of factors covered by the term 'transmission mechanisms' across the four studies. Many different typologies for organizing influences on children's development are possible, but it is extremely difficult to find a scheme that applies naturally to all the mechanisms in which we are interested, over the different stages of childhood. We choose the single broad-brush distinction between parental attitudes and behaviours and young people's attitudes and behaviours, as one that can be applied consistently across studies. Within each study, the mechanisms considered are then grouped into specific sub-categories, such as the 'home learning environment' experienced in the pre-school period or the 'anti-social behaviours' engaged in by adolescents.

Figure 3 also shows that we make a clear conceptual distinction in our analysis, between the transmission mechanisms listed in the central panels, and a set of other 'distal' characteristics listed beneath parental socio-economic status on the left. We do not consider factors such as parental education, family composition and the school attended by a child to be proximal in the same sense as the transmission mechanisms of interest. Instead, like socio-economic status, they are structural features of children's environments that are associated with disparities in child outcomes which can, in theory, be fully explained by differences in intermediary processes. Measures of other distal factors besides SEP are useful, because their inclusion provides a check on how far all the relevant transmission mechanisms are captured by our data, and can be used to proxy for unobserved proximal factors. If all the relevant proximal processes for children's development are measured and controlled in our analysis, then distal factors

such as single parenthood and parental education should have no remaining association with the child outcome. In fact we find that many such distal factors remain strong and significant predictors of outcomes even in the fully controlled models, implying that they are systematically related to unmeasured factors that are consequential for children's development. Their inclusion in the models, therefore, helps to 'mop up' the residual unexplained variation in child outcomes. The interpretation of their contribution to outcomes, however, is crucially different from that of the proposed transmission mechanisms, as their 'direct' effects via any of the observed mechanisms have been controlled away. Their remaining association with outcomes, then, is only a partial one and reflects an 'indirect' association via residual unknown mechanisms that ideally would be observed and controlled.

## 5. Empirical strategy

In order to avoid repetition within articles, this section formally sets out the model used in the three longitudinal studies and that forms that baseline in the inter-generational study. Whatever the stage of childhood, it is clear that the inter-relationships between different groups of mechanisms will be many and complex. Parental attitudes and behaviours early in life will influence the attitudes and behaviours adopted by the child, which will then in turn affect parents' attitudes and behaviours later in life, and so on. We make no attempt to unpick this type of intermediate relationship. Every transmission mechanism in the analysis is considered simultaneously, so the associations identified are conditional or net associations only. Specifically, our analysis makes use of the coefficients from the (within-study) least squares regression:

$$Outcome_{it} = \alpha Mechanisms_i + \beta Distal\_chars_i + \gamma SEP_i + \varepsilon_{it} \quad (1a)$$

The dependent variable is the outcome score of child  $i$  at age  $t$ , the age of the child at the end of the observation period in each dataset. (This is age 5 in the pre-school years study, 11 in the primary years study and age 16 in the secondary years study. The pre-school years study also estimates some versions in which age 3 scores are treated as the final outcome.) The outcome is regressed

simultaneously on all the mechanism variables listed in Figure 3 (which may be measured at different points in the child's life), a set of distal characteristics such as parental education and family composition and a set of dummy variables for the five SEP quintile groups, omitting the lowest quintile group from the regression as the reference case. The  $\varepsilon_{it}$  term is an individual-specific error that

is uncorrelated with any observed child characteristics.

The  $\alpha$  coefficients therefore capture the correlation between a factor of interest and the outcome holding all other observable factors constant. An insignificant estimate of a particular  $\alpha$  should not be taken to imply that the factor is necessarily inconsequential for children's development. It is possible that the factor operates entirely through its 'knock-on' impact on other influences that are controlled in the model. The 'all else constant' assumption means that care is needed in interpreting the contribution of a single variable when other highly correlated variables are also present in the model. The joint contribution of a group of variables, that together measure a common concept, can be thought of providing a more 'realistic' measure of association than any single marginal effect.

The concept of 'a correlation holding all other observable factors constant' also highlights that the

$\alpha$  coefficient will pick up the influence of any unobservable factors that are correlated with both the factor of interest and the outcome. This classic problem of omitted variables bias is essentially unavoidable in a regression-based study of this kind, and reminds us that the  $\alpha$ 's cannot be considered estimates of causal effects. We note however, that the inclusion of a rich set of distal family characteristics and the SEP dummies in equation (1a), means that the source of any bias must be uncorrelated not only with any of the included transmission mechanisms, but also with the structural features of families and schools captured by these proxies.

In three of our studies, longitudinal data on the same children at different ages allow us to explore the timing of the developmental process and its association with different transmission mechanisms in more detail. In these studies we also estimate 'value-added' regressions in which a lagged outcome score is added to equation (1a).

$$Outcome_{it} = \alpha^V Mechanisms_i + \beta^V Distal\_chars_i + \gamma^V SEP_i + \delta^V Outcome_{it-1} + \varepsilon_{it}^V \quad (1b)$$

Here the  $t-1$  subscript does not indicate a lag of a fixed period of time, but rather indexes the child's age at start of the observation period in the dataset in question. In comparison with the  $\alpha$ 's from equation 1a, the  $\alpha^V$ 's capture the association of the mechanism with the child's trajectory between age  $t-1$  and age  $t$ , rather than the level of the outcome at age  $t$ . These estimates help to disentangle how far the influence of each mechanism is specific to the developmental stage in question, and how far it has already been 'embedded' in attainment at the start of the period. These estimates provide a stringent test of the predictive power of the mechanisms because they net out the part of the causal effect of the mechanism that has manifested itself in the outcome at  $t-1$ , as well as its correlation with any later unobserved confounders that are associated with prior achievement.

The estimates from equations 1a and 1b provide estimates of the independent association of every mechanism variable with the outcome, but this does not tell us their empirical importance in explaining the raw socio-economic gap. For this, a second step is required that captures the association between each mechanism and SEP. A factor can only be important in predicting the socio-

economic gap if it both differs systematically between socio-economic groups and is associated with the child outcome. Its total "contribution" as an explanatory factor is the product of these two associations, and so may differ from the impression given by the coefficients in the outcome regressions alone. One way to interpret this contribution is as the predicted difference in the outcome if average differences in the mechanism in question between SEP groups were eliminated.

In this second step, we quantify the SEP-mechanism association using coefficients from the following regression:

$$Mechanism_{ij} = \theta_j SEP_i + \mu_{ij} \quad (2)$$

The  $\theta_j$  coefficients capture the mean differences in the  $j$ th mechanism variable between the omitted lowest SEP quintile group and the other four groups. Large and significant estimates of the  $\theta_j$ 's, therefore, indicate strong social grading in the factor of interest. The term  $\mu_{ij}$  is the within-group residual of child  $i$  – the deviation from the mean value of the  $j$ th mechanism variable in the SEP quintile to which the child belongs – and as such is uncorrelated with SEP.

To provide a complete set of estimates between the observed variables in the model, we require a further regression for the relationships between SEP and the  $k$  distal characteristics, identical in format to equation 2:

$$Distal\_char_{ik} = \pi_k SEP_i + \omega_{ik} \quad (3)$$

$$Outcome_{it} = \left\{ \gamma + \sum_j \alpha_j \theta_j + \sum_k \beta_k \pi_k \right\} SEP_i + \left\{ \sum_j \alpha_j \mu_{ij} + \sum_k \beta_k \omega_{ik} + \varepsilon_{it} \right\} \\ \equiv BSEP_i + \eta_{it} \quad (4)$$

The idiosyncratic error terms – the  $\mu_{ij}$ 's,  $\omega_{ik}$ 's and  $\varepsilon_{it}$ 's – are uncorrelated with SEP by construction, so the unconditional regression coefficients on SEP – the B vector – can be decomposed into the sum of terms shown in the first set of brackets. The product term  $\alpha_j \theta_j$  measures the contribution of the  $j$ th mechanism to this raw gap. To illustrate, the analysis in Gregg and Washbrook (this issue, pp 41-58) estimates that a 1 kg difference in birth weight is associated with a 2.0 percentile rank difference in the Key Stage 2 outcome score at age 11, holding all observable variables constant (Table 3). This is the  $\alpha_j$ . Children in the highest SEP quintile group are, on average, 0.11 kg heavier at birth than children in the lowest quintile (Table 2). This is the  $\theta_j$ . The product of the two terms gives the contribution of birth weight differences to the socio-economic gradient: 2.0 multiplied by 0.11, or 0.22 percentile points, less than 1 percent of the overall 31.3 point gradient

The vectors of coefficients estimated from the three sets of equations can be brought together in a simple decomposition that summarizes the hundreds of associations between SEP, the intermediary variables and the outcome in a single set of figures. To see this, we can substitute 2 and 3 into 1a:

(Table 4). This example also serves to illustrate why large significant coefficients from the outcome equation 1a may be misleading as to the importance of a particular mechanism for the socio-economic gap. Although birth weight is significantly associated with positive outcomes, disadvantaged children weigh only slightly less on average than the most advantaged children, so eliminating the socio-economic gap in birth weight predicts only a minor change in the socio-economic gradient.

In order to construct the value-added equivalent of the decomposition, we must estimate one further regression:

$$Outcome_{it-1} = \rho^V SEP_i + \varphi_{it-1} \quad (5)$$

The value-added equivalent of 4 can then be derived by substituting 2, 3 and 5 into 1b to get:

$$Outcome_{it} = \left\{ \gamma^V + \sum_j \alpha_j^V \theta_j + \sum_k \beta_k^V \pi_k + \rho^V \delta^V \right\} SEP_i + \left\{ \sum_j \alpha_j^V \mu_{ij} + \sum_k \beta_k^V \omega_{ik} + \delta^V \varphi_i + \varepsilon_{it}^V \right\} \\ \equiv BSEP_i + \eta_{it} \quad (6)$$

Here, B again represents the unconditional difference in the mean outcomes of children in different SEP groups. The term  $\rho^V \delta^V$  captures the extent to which differences in prior outcomes can explain differences in current outcomes. In the value-added model the product terms  $\alpha_j^V \theta_j$  now capture the contribution of the  $j$ th mechanism to

the SEP gap conditional on outcomes at  $t-1$ . To return to the birth weight example given above, a 1 kg difference in birth weight is associated with a 0.5 percentile point difference in the age 11 outcome in the value-added model, compared with a 2 point difference in the levels model. This estimate of  $\alpha_j^V$  is multiplied by the same 0.11 kg estimate of  $\theta_j$  to

give a contribution of 0.055 percentile points to the overall socio-economic gradient when age 7 outcomes are held constant.

The difficulties involved in interpreting a single  $\alpha$  coefficient discussed above, also affect the interpretation of a single product term  $\alpha_j\theta_j$ . For this reason we sum over the contribution of a group of related variables when presenting the results of the decomposition. Overall, the decomposition technique gives us a parsimonious impression of the relative importance of different types of measured factors in explaining the socio-economic gradient. Note that the importance of a set of factors in the framework depends entirely on the magnitudes of the underlying associations, rather than on the precision with which they are estimated.

An alternative approach (and one that we employed in preliminary analyses) is to observe how the coefficients on the SEP quintile dummies change (or are mediated) as variables are progressively added to equation 1a. A drawback of this approach is that the results depend crucially on the order in which variables are added to the model, as earlier additions will pick up some of the effects of later ones to the extent they are correlated. Without a strong "time-ordering" of variables the sequence is arbitrary, and only multiple replications with different combinations can determine the sensitivity of the findings to a particular ordering of the introduction of controls. Since our aim is provide parsimonious estimates that are comparable across studies and life course stages, the decomposition method outlined above, which draws conclusions from a single conditional model, has a distinct advantage.

## 6. Summary and interpretation

Throughout the papers in this Special Issue we have explored how children from poor backgrounds typically show lower educational attainment compared to children from better off backgrounds, and why this gap widens throughout much of childhood. We begin our story at the very earliest stages of childhood, and follow young people up until the age of 16, when they potentially obtain their first formal qualifications. Our main analysis splits childhood into three periods, broadly conforming to pre-school, primary and secondary phases of education, recognising that cognitive development and attainment within each period builds on learning

in the previous one(s). The papers show a wealth of simple evidence that, from the earliest of ages, poorer children experience much less advantageous environments at home than children from better off backgrounds, and that differences in these environments have a strong association with poor children's lower cognitive development in early childhood, and progressively poorer academic attainment through school. The differences we have found cover many different aspects of home life, from home-learning environments and parenting styles at a young age, to parents' aspirations and expectations for their child's future education during primary and secondary school, measures of family closeness, and the availability of material resources such as a computer and internet at home during the teenage years. At the same time we have also found that children from poor families typically display many more behavioural problems, at all ages, than children from better off backgrounds.

The research also highlights a number of key findings on the stability or otherwise of measured ability across generations and over the course of childhood. Our analysis of the BCS explores the inter-generational heritability of cognitive capabilities, with approximately one fifth of the gap between richest and poorest explained by a direct link between the cognitive skills of the parent and child, one that is unmediated by the rich set of environmental factors observed in our surveys. Such a relationship may reflect genetic inheritance or the inheritance of environmental disadvantage not captured by the variables we can observe. However, it does suggest that direct genetic heritability of cognitive ability can only be a small but non-trivial component of the socio-economic gradient in attainment. A second general point relates to the role played by prior attainment, as pupils age. When considering our value added models where prior attainment is included, we find that it accounts for between 40 and 60% of the variation in current attainment in each period. The contribution of prior attainment is lowest for age sixteen, when prior attainment is measured at age 11. Such differences may reflect measurement errors in attainment, the assessment of a wider range of subjects at age 16 than at prior ages, the fact that children develop at somewhat different ages or that environmental influences impact on children's progress between assessments. Whilst we cannot fully isolate the relative importance of these different explanations, the data certainly point to a potentially



important role for the last of these and hence for policy in reducing the socio-economic gradient.

The big question arising from our work is what it can tell us about policy formation. Will improved parenting skills in the very early stages of life lead to better outcomes at school, many years later? Will raising maternal aspirations for education, young people's self-esteem and ability beliefs have a similar effect? By the teenage years, can improving young people's own aspirations, reducing their involvement in risky behaviours and encouraging positive behaviours help to close the gap between the poorest children and those from better off backgrounds, and hence help to break the cycle of poverty across the generations?

The evidence presented offers three major areas in which policy may make a contribution to reducing educational inequalities. Of course these three broad areas do not operate in isolation from each other – each having extremely important feedbacks on the others.

*(i) Parents and the family home:*

- Improving the home learning environment in poorer families (e.g. books and reading pre-school, computers in teen years)
- Helping parents from poorer families to believe that their own actions and efforts can lead to higher educational outcomes
- Raising families' aspirations and desire for advanced education – from primary schooling onwards

*(ii) The child's own attitudes and behaviours, and their approach in taking forward their past experiences into learning:*

- Reducing children's behavioural problems; improving coping and management capabilities for risky behaviours, conduct disorder and ADHD
- Helping children from poorer families to believe that their own actions and efforts can lead to higher educational outcomes
- Raising children's aspirations and desire for advanced education – from primary schooling onwards

*(iii) The school's approach:*

- Schools could arguably be doing more to reduce inequalities in attainment between rich and poor, and potentially have a very significant role to play in counteracting the effects of the big inequalities in family backgrounds and home environments that our study has revealed.

When relating the findings to policy questions in more detail, it is necessary to sound a strong note of caution. While our models generally include prior attainment and long-run background factors as controls - helping us to isolate the effects of specific age-related factors - our research is nevertheless based on detailed statistical correlations, rather than robust trials. This means that we have not established robust causal relationships from this work. More generally, the measures of aspirations, attitudes and behaviours that we include in our model are likely to be indicative of wider processes operating within families and peer groups, and there are likely to be other unmeasured differences across families which our measures are partially capturing. The possibility of correlated unobservable characteristics, and of reverse causation, mean that taking our findings purely at face value could lead us to misplaced policy conclusions. Moreover, many of the aspects of parental and child attitudes and behaviours that we have considered are strongly related to each other. Hence it is not always appropriate to isolate one of these factors as a focus for intervention, when it might reflect a broader set of attitudes and beliefs that are not easily measured independently.

One way to throw light on the causality or otherwise of the associations identified in the research, is to look for corroborating evidence from the implementation of existing policies. UK governments have introduced a number of policies based around these broad areas aimed at closing the attainment gaps between rich and poor. In order to understand how successful these policies have been in reducing the gap in school attainment between rich and poor children, we need address several key questions. Are these factors - namely early environments, attitudes, aspirations and the like – malleable, and have these policies actually been successful at improving them? Do such improvements raise poor children's attainment in the way that is hoped? The following section reviews the current policy evidence base in the context of our findings, and highlights areas in which further evaluation studies are needed.

## 7. Policy Interventions

We begin by discussing programmes designed to influence parenting, the home learning environment, and early years' childcare and education provision. Sure Start is now a national

programme that aims to reach all families, with more intensive support for the more needy. The early evidence on Sure Start among children at age 3 was rather mixed (NESS Research team, 2008), although we understand that the evidence for age 5, which will soon be available, is more encouraging. More targeted programmes, by their nature, are more straightforward to evaluate and there is clearer, positive evaluation evidence on some of these. For example, the introduction of the Family Nurse Partnership in 30 pilot sites in the UK – aimed at improving very early parenting skills, and parent and child health - is backed up by randomised control trials, showing the effectiveness of this programme in the US in improving children's long-term behavioural and cognitive outcomes (up to thirteen years after involvement in the programme; Olds et al 1998). The parenting elements of the Incredible Years Programme, operating in various guises around the UK, is another well-evaluated parenting programme showing success in improving child behaviours among children at a young age (Hutchings et al 2007 and Bywater et al 2009). Some other parent-centred programmes, while not yet subject to fully robust outcome evaluations, appear quite promising. These include Family Intervention Projects, which address the problems of a small number of families with severe behavioural problems, tackling what is typically a complex web of mental and physical health problems, substance misuse and domestic violence (White et al 2008).

Three features of the parenting-based work discussed above are worth drawing out. The first is that the majority of parenting support programmes are aimed at pre-school aged children. Whilst there is a clear and obvious reason for this, our research highlights the ongoing potential for improved parenting to reduce inequalities in child development, certainly into the primary years and perhaps to a lesser extent into the secondary school period. Second, the best evidence we have on programmes being successful is for high intensity (and costly) programmes concentrated on the most needy families and children. While intensive programmes that focus on helping small numbers of children most in need tend to have the strongest evidence base behind them, educational disadvantage affects a very large number of children from low income families, but with lower intensity than those at the extreme, and it may be that policy needs to focus more on these (although Sure Start is

a major exception here). Finally, the evaluation evidence tends to be clearer about the positive impact of these programmes on children's social and emotional well-being, and health, but is generally much less clear about their impacts on children's long-term cognitive development, and educational attainment. While both are clearly important, if one is trying to reduce *educational* inequalities, then this latter point is clearly a relevant concern. The question mark over whether such programmes improve cognitive development and raise educational attainment chimes with our own findings. For example, Dearden, Sibieta and Sylva (this issue, pp 19-40) highlighted that although children from poor families typically experience much less advantageous early caring environments, most aspects of the home environment (except for specifically the home *learning* environment) were not directly responsible for the big gaps in cognitive development we sought to explain. However they were important for explaining differences in children's social and emotional well-being.

There are also a number of mainly school-based programmes aimed at raising children's aspirations and tackling behavioural and emotional issues. For example, one major voluntary programme for primary and secondary schools is the Social and Emotional Aspects of Learning (SEAL), which emphasises the importance of social skills such as empathy, self-awareness, and self-regulation. Aimhigher seeks to raise aspirations for Higher Education among young people, while various programmes under the National Behaviour and Attendance Strategy seek to improve behaviour within the school context; Aiming High and, within this, Extended Schools Services, aim to promote youth engagement in positive activities. Many of these programmes – such as SEAL, and various elements of the government's strategy towards behaviour and attendance - emphasise the importance of the whole-school ethos in improving young people's attitudes and behaviours, as well as individual- or small-group work.

Our reading of the evidence on these types of programmes is that in general, their effectiveness is much less robustly established than the parenting-focused programmes we discussed further above. As such, their benefit remains unproven. One exception is Aimhigher, or more specifically the Excellence Challenge element, where robust evaluation findings on attainment are positive.

Targeted at young people in urban, deprived schools, it was found that one school year's exposure to the programme in Year 11 (age 15-16) led to pupils scoring 2.5 points higher at GCSE (equivalent to 2.5 grades improvement on the current scale) and being 3.9 percentage points more likely to report that they intended to participate in higher education (Emmerson et al 2005). While Aimhigher thus appears very useful, it starts in the secondary phase of schooling. In this issue we find that aspirations are also an important potential influence on attainment even by the age of 11 (see Gregg and Washbrook, this issue, pp 41-58), suggesting that activities aimed at raising aspirations in primary school might also be valuable.

By contrast, our reading from various evaluations of SEAL suggests that this approach is as yet unproven – since in general, clear benefits have not been very robustly established. For example, in the one independent evaluation that has involved a control group design (Hallam et al 2006), statistically significant positive impacts were found for some social and emotional outcome measures, but many more outcomes did not appear affected by the interventions, and indeed there were a number of important outcomes that appeared adversely affected by some interventions. One intervention, 'Going for Goals' did show a more consistent positive impact on the children involved, though no impact was found on young people's motivation, the main aspect of learning that is supposedly addressed by this intervention. Additionally, to our knowledge this (or any other) programme's impact on young people's sense that their destiny can be shaped by their own actions (locus of control) has not been tested – though the findings from this issue suggest that this may be important.

Stronger evidence is also required on the effectiveness of the government's strategies towards behaviour improvement (the National Behaviour and Attendance Strategy), and on positive activities (the Aiming High strategy, including Extended Schools services). The Behaviour Improvement Programme (BIP) was one specific intervention within this umbrella, which was subject to a formal evaluation and found positive benefits on young people's school attendance (Hallam et al 2005). However, other formal evaluation work on the BIP, based on the LSYPE, found no discernible impacts of the programme on young people's likelihood of truancy, or on any other of a detailed set of attitudes and

behaviours, or on attainment at age 14 (Chowdry et al 2009).

More evidence is also needed regarding the promotion of positive activities (including the Aiming High strategy and the connected Extended Schools services). Aiming High is the previous government's ten year strategy aiming to increase young people's participation in constructive leisure activities. While there was piecemeal evaluation of some elements of the Aiming High strategy, such as the national evaluation of the Positive Activities for Young People, there has been no overall evaluation of the effectiveness of the approach. Under the connected Extended Schools Services (ESS) programme, councils set up activities in and around schools for the evenings, weekends and during holidays. Services offered include study support; play/recreation, sport, music, arts and crafts and other special interest clubs; volunteering and business and enterprise activities; childcare; parenting support; specialist services such as speech and language therapy; and community access to facilities including adult learning, ICT and sports facilities. While the formal national evaluation of ESS is yet to report, evaluation of a predecessor programme found some evidence for positive impacts on young people's behaviour and learning (Cummings et al 2007).

There are also a number of more intensive initiatives and teaching programmes in schools designed to directly improve the learning outcomes of children and young people in particular need of help, many of whom are from disadvantaged backgrounds. These initiatives include Special Educational Needs provision (SEN), and very intensive programmes in primary school such as Every Child a Reader, Every Child Counts and Every Child a Writer. The basic effectiveness of programmes such as Reading Recovery, the intervention at the core of the Every Child a Reader programme, in helping young children struggling to read to catch up with their peers has been robustly established in a number of different studies. However, uncertainty remains as to whether such gains are sustained in the longer-term, and the *cost-effectiveness* of these very expensive, intensive one-to-one teaching programmes has been both asserted (KPMG, 2009) and questioned (Policy Exchange, 2009). Other programmes such as the Literacy and Numeracy hours have also been backed up by positive evaluation findings, and suggest that the positive benefits are found more among children

from low income families (Machin and McNally 2004).

One set of issues not touched upon yet in this policy discussion is whether (i) the level of resources channelled towards pupils from low income backgrounds in schools, (ii) the funding mechanisms for delivering these, and (iii) the school structures into which such resources are channelled, are likely to be effective in reducing the gap in educational attainment between rich and poor children. Funding provided to schools is already biased in favour of more deprived schools (Chowdry et al 2007a). Leading up to the 2010 UK general election, both the Conservatives and the Liberal Democrats campaigned on the idea of a disadvantaged pupil premium in the school funding system in England, and a commitment to introduce a pupil premium was then included in the Coalition's programme for government. The main aim of the pupil premium is to narrow the achievement gap between children coming from rich and poor families, by targeting resources even more heavily towards schools with a high proportion of disadvantaged pupils, and reducing any disincentive that schools might have to recruit such pupils. In a recent report, Chowdry, Greaves and Sibieta (2010) examined the rationale for a pupil premium and concluded that current evidence suggests a pupil premium is only likely to have a modest impact on the achievement gap. Furthermore, as a result of planned cuts to existing funding, the pupil premium is only likely to represent a net increase in real-terms funding for a small number of schools with large numbers of deprived pupils – only 1 in 8 schools are likely to see increases in real-terms funding of 5% or more in total over the next four years<sup>1</sup>. The vast majority of schools are likely to see net cuts in their funding over the next four years, with the least-deprived schools likely to see real-terms cuts of about 10%. To the extent that the pupil premium is likely to narrow the achievement gap, it seems only likely to do so through reducing the level of cuts imposed on relatively deprived schools.

Another set of policies of particular note are those that are designed to incentivise or force young people (particularly those from poor backgrounds) to remain longer in formal education. Education Maintenance Allowances (EMAs; introduced nationwide in 2004) were designed to encourage

more young people from low income backgrounds to remain in full-time education beyond 16, through a means-tested payment of up to £30 per week, made to young people aged 16-18. Robust evaluation evidence suggests that the financial incentive works: there have been positive impacts on staying-on rates, retention, and achievement (Chowdry et al 2007b). The coalition government has announced that EMAs are to be abolished. Yet bigger changes in this area are imminent, with the forthcoming raising of the minimum education and training participation age. New legislation means that in the academic year 2013/14, young people will have to remain in some sort of education and training until the age of 17, and in 2014/2015 until the age of 18. This will largely impact on young people from poorer backgrounds, who are the most likely to leave school and training before 18 under the current system. While previous legislation to increase the *school* leaving age has generally been shown to raise attainment and have positive economic returns, it remains to be seen whether this particular extension, which increases the minimum leaving age by a further two years, and also includes jobs with formal training, will have a similar effect.

## 8. Conclusion

Overall, our results suggest that the broad area of aspirations, attitudes and behaviours of parents and children have the potential to be an important area for policy intervention. Our results are not causal, and so policy trials are needed to demonstrate the degree of malleability of mediator factors and subsequent impact on attainment. The interventions described above, which have variable quality of evidence of effectiveness, support the sense that this area is ripe for such policy trials, building on best practice to date. Two areas we regard as promising are interventions designed to reach a broader range of children than the acutely deprived or low achieving targeted in many programmes, and interventions that expand the age range of children and the involvement of their parents in other programmes. Our hope is that the evidence in this Special Issue will contribute to policy thinking and innovation, particularly around interventions targeted at children beyond the early years.

## Acknowledgments

The Joseph Rowntree Foundation has supported this project as part of its programme of research on education and poverty. We very gratefully acknowledge the support of JRF, and the productive guidance of Helen Barnard and Chris Goulden. We are also very thankful to members of our advisory group: Tim Crosier, Naomi Eisenstadt, Leon Feinstein, Zoe Ferguson, Tony Foot, Leslie Gutman, Lisa Harker, Andrew Ledger, Mark Newman, Ivan Turok, Anna Vignoles, Jane Waldfogel and Stephen Witt. We also gratefully acknowledge the role of the Leverhulme Trust in providing support for Washbrook's time, the ESRC Large Grant RES-060-23-0011 for support of Gregg's time, and DCSF for funding Chowdry, Crawford and Goodman on a parallel project on the LSYPE. Particular thanks also go to Debra Hertzberg and Lindsey Macmillan for research assistance. We would also like to thank the ALSPAC team for extensive help with data development. All views expressed are those of the authors. Contributors to the material this paper draws on are Haroon Chowdry, Claire Crawford, Lorraine Dearden, Robert Joyce, Luke Sibieta, Kathy Sylva and Elizabeth Washbrook.

## References

- Becker, G and Tomes N. (1986) Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3), S1-S39.
- Blow L, Goodman A, Kaplan G, Walker I and Windmeijer F. (2004) How Important is Income in Determining Children's Outcomes? A Methodology Review of Econometric Approaches. <http://www.ifs.org.uk/docs/methodology.pdf>.
- Bourdieu P. (1977a) Cultural Reproduction and Social Reproduction. In J. Karabel and A.H. Halsey. eds. *Power and Ideology in Education*. Pp 487-511. Oxford University Press, New York.
- Bourdieu P. (1977b) *Outline of a Theory of Practice*. Cambridge University Press, Cambridge.
- Bourdieu P and Passeron JC. (1977) *Reproduction in Education, Society and Culture*. (translated R Nice). Sage, Beverly Hills, Ca.
- Bronfenbrenner, U and Morris PA. (1998) The Ecology of Developmental Processes. In W. Denton (series ed) and RM Lerner (vol. ed) *Handbook of Child Psychology*, vol.1 Theory (5th edn). Wiley, New York.
- Brooks-Gunn J and Duncan G. (1997) The effects of poverty on children and youth. *The Future of Children*, 7, 55-71.
- van Buuren S, Boshuizen HC and Knook DL. (1999) Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine* 18, 681-694.
- Bywater T, Hutchings J, Daley D, Whitaker C, Yeo ST, Jones K, Eames C and Tudor Edwards R. (2009) Long-term effectiveness of a parenting intervention for children at risk of developing conduct disorder. *British Journal of Psychiatry* 195, 1-7. doi: 10.1192/bjp.bp.108.05631.
- Carneiro P and Heckman JJ. (2002) The evidence on credit constraints in post-secondary schooling. *Economic Journal*, 112(482), 705-734.
- Carneiro P and Heckman JJ. (2003) Human Capital Policy. in Heckman JJ and Kreuger AB (eds) *Inequality in America*. MIT Press, Boston, Ma.
- Chowdry H, Muriel A and Sibieta L. (2007a) *Level Playing Field? The implications of school funding*. CfBT Education Trust Research Report.
- Chowdry H, Dearden L and Emmerson E. (2007b) *Education Maintenance Allowance: Evaluation with Administrative Data. The impact of the EMA pilots on participation and attainment in post-compulsory education*. Learning and Skills Council. <http://readingroom.lsc.gov.uk/lsc/National/nat-emaevaluationadministrativedata-jan2008.pdf>
- Chowdry H, Greaves E and Sibieta L. (2010) The Pupil Premium: assessing the options. Institute for Fiscal Studies Commentary 113, (<http://www.ifs.org.uk/publications/4776>).
- Croll P. (2004) Families, Social Capital and Educational Outcomes. *British Journal of Educational Studies* 52, 390-416.
- Cummings C, Dyson A, Muijs D, Papps I, Pearson D, Raffo C, Tiplady L and Todd L, with Crowtheri D. (2007) *Evaluation of the Full Service Extended Schools Initiative: Final Report*. Dept for Education and Schools, London.
- Duncan G and Brooks-Gunn J. eds. (1997) *Consequences of growing up poor*. Russell Sage Foundation, New York.
- Duncan G and Brooks-Gunn J. (2000) Family Poverty, Welfare Reform, and Child Development. *Child Development*, 71, 188-196.

- Emmerson C, Frayne C, McNally S and Silva O. (2005) *The early impact of Aimhigher: Excellence Challenge on pre-16 outcomes: an economic evaluation*. DfES Research Report No. 652, DfES Publications, Nottingham.
- Esping-Andersen G. (2005) Social Inheritance and Equal Opportunity Policies. In S Delorenzi, J Reed and P Robinson. eds. *Maintaining momentum: Promoting Social Mobility and Life Chances from Early Years to Adulthood*. Institute for Public Policy Research, London.
- Feinstein L. (2003) Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 70, 73-97.
- Feinstein L. (2004) Mobility in Pupils' Cognitive Attainment During School Life. *Oxford Review of Economic Policy*, 20, No. 2: Education, 213-229.
- Galobardes G, Lynch J and Davey Smith D. (2007) Measuring socio-economic position in health research. *British Medical Bulletin*, 81-82, 21-37.
- Gregg P and Machin S. (1998) Childhood disadvantage and success or failure in the youth labour market. *Centre for Economic Performance Discussion Paper* 397.
- 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, 259-280.
- Hallam S, Rhamie J and Shaw J. (2006) *Evaluation of the primary behaviour and attendance pilot*. DfES, London.
- Harmon C and Walker I. (1995) Estimates of the Economic Return to Schooling for the United Kingdom. *The American Economic Review*, 85, 1278-1286.
- Haveman R and Wolfe B. (1995) The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 23, 1829-1878.
- Hutchings J, Bywater T, Daley D, Gardner F, Whitaker C, Jones K, Eames C and Edwards RT. (2007) Parenting Intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. <http://www.bmj.com/cgi/content/abstract/334/7595/678>
- Kolenikov S and Angeles G. (2004) *The Use of Discrete Data in Principal Component Analysis: Theory, Simulations, and Applications to Socioeconomic Indices*. Working Paper of MEASURE/ Evaluation project, No. WP-04-85, Carolina Population Center, UNC.
- KPMG Trust. (2009) The long-term costs of literacy difficulties. 2nd edition. [http://www.everychildachancetrust.org/ecar/pubs/long\\_term\\_costs\\_of\\_literacy\\_report.pdf](http://www.everychildachancetrust.org/ecar/pubs/long_term_costs_of_literacy_report.pdf)
- Lareau A. (1987) Social Class Differences in Family-School Relationships: The Importance of Cultural Capital. *Sociology of Education*, 60, 73-85.
- Machin S and McNally S. (2004) *The Literacy Hour*. Centre for the Economics of Education discussion paper 43. <http://cee.lse.ac.uk/cee%20dps/ceedp43.pdf>
- Mayer S. (1997) *What money can't buy: The effect of parental income on children's outcomes*. Harvard University Press, Cambridge, MA.
- Meghir C and Palme M. (2005) Educational Reform, Ability and Parental Background. *American Economic Review*, 95, 414-424.
- NESS Research team. (2008) *The impact of Sure Start Local Programmes on three year olds and their families*. Dept for Children, Schools and Families, London.
- Olds D, Henderson Jr. CD, Cole R, Eckenrode J, Kitzman H, Luckey D, Pettit L, Sidora K, Marros P and Powers J. (1998) Long-term Effects of Nurse Home Visitation on Children's Criminal and Antisocial Behavior: 15 Year Follow-up of a Randomised Controlled Trial. *Journal of the American Medical Association*, 280, 4.
- Policy Exchange. (2009) Every Child a Reader: An example of how top-down education reforms make matters worse. [http://www.policyexchange.org.uk/images/publications/pdfs/Every\\_Child\\_a\\_Reader.pdf](http://www.policyexchange.org.uk/images/publications/pdfs/Every_Child_a_Reader.pdf)
- Sylva K, Melhuish E, Sammons P, Siraj-Blatchford I and Taggart B. (2008) *Effective Pre-school and Primary Education 3-11 Project (EPPE 3-11): Final Report from the Primary Phase: Pre-school, School and Family Influences on Children's Development during Key Stage 2 (Age 7-11)*. DCSF, London.
- White C, Warrenner M, Reeves A and La Valle I. (2008) *Family Intervention Projects: An Evaluation of their Design, Set-up and Early Outcomes*. National Centre for Social Research, Report No. DCSF-RBW047.

## Endnote

<sup>1</sup> <http://www.ifs.org.uk/publications/5311>

# The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study

**Lorraine Dearden**

Institute for Fiscal Studies and Institute of Education, University of London

[ldearden@ifs.org.uk](mailto:ldearden@ifs.org.uk)

**Luke Sibieta**

Institute for Fiscal Studies

**Kathy Sylva**

University of Oxford

(Received July 2010 Revised November 2010)

## Abstract

*This paper shows that there are large differences in cognitive development between children from rich and poor backgrounds at the age of 3, and that this gap widens by the age of 5. Children from poor backgrounds also face much less advantageous “early childhood caring environments” than children from better off families. For example we identify differences in poor children’s and their mothers’ health and well-being (e.g. birth-weight, breast-feeding, and maternal depression); family interactions (e.g. mother-child closeness); the home learning environment (e.g. reading regularly to the child); parenting styles and rules (e.g. regular bed-times and meal-times), and experiences of childcare by ages 3 and 5. Differences in the home learning environment, particularly at the age of 3 have an important role to play in explaining why children from poorer backgrounds have lower test scores than children from better off families. However, a much larger proportion of the gap remains unexplained, or appears directly related to other aspects of family background (such as mothers’ age, and family size) that are not mediated through the early childhood caring environment.*

**Key words:** home learning environment, early child development, socio-economic gap

## 1. Introduction

Children growing up in poor families emerge from our schools with substantially lower levels of educational attainment. These educational deficits emerge early in children’s lives, even before entry into school, and widen throughout childhood (Feinstein 2003, 2004). In this article, we examine some of the routes through which family socio-economic position (SEP) affects cognitive development in the early years up to age 5.

Other articles in this Special Issue examine the socio-economic gap in the primary years (Gregg and Washbrook), the secondary years (Chowdry, Crawford and Goodman) and inter-generational factors (Crawford, Goodman and Joyce). The overall theoretical and empirical approach taken in these

four articles is discussed in detail in Goodman, Gregg and Washbrook, also in this Special Issue.

In seeking to explain the socio-economic gap in early child cognitive outcomes, we focus on a range of parenting behaviours, health and well-being factors, and family interactions. We refer to these factors collectively as the “early childhood caring environment”. These have seldom been measured together in large datasets, preventing comprehensive empirical analysis of their roles as predictors and consequences of poor child outcomes. In order to conduct this analysis we will follow children over time within the Millennium Cohort Study (MCS), a recent and detailed study of children born at the turn of the millennium. Here

we focus on cognitive outcomes, but in our more detailed working paper (Dearden, Sibieta and Sylva, 2010) we also examine socio-emotional development, the mediation of other gradients and the determination of the early childhood caring environment.

The rest of this paper proceeds as follows: section 2 describes our theoretical approach, and section 3 describes the data and summary statistics, including the socio-economic gradients in early child outcomes and the early childhood caring environment. In section 4, we use multivariate regression techniques to decompose the socio-economic gaps in child outcomes as per the empirical methodology set out in Goodman, Gregg and Washbrook in this Special Issue. Finally, section 5 concludes the paper.

## 2. Theoretical Approach

In this paper, we use the same theoretical approach described in detail in Goodman, Gregg and Washbrook in this volume. Our starting point is the socio-economic gradient in child outcomes in the early years. We then explore some of the potential transmission mechanisms through which socio-economic disadvantage may lead to poor child outcomes. In doing so, we attempt to explore the role played by a set of 'proximal' factors that may drive the relationship between poor child outcomes and 'distal' influences such as socio-economic disadvantage and other aspects of family background. Generally speaking, proximal factors are those that drive the observed association between a distal factor and an outcome (as can be seen in Figure 1).

In this paper, we refer to the full set of proximal influences we observe as 'the early childhood caring environment'. This includes the nature of family interactions, health and well-being factors, childcare arrangements, the home learning environment (HLE), and parenting style and rules. These factors are inter-related, and seem to be the most likely to influence child development between birth and the

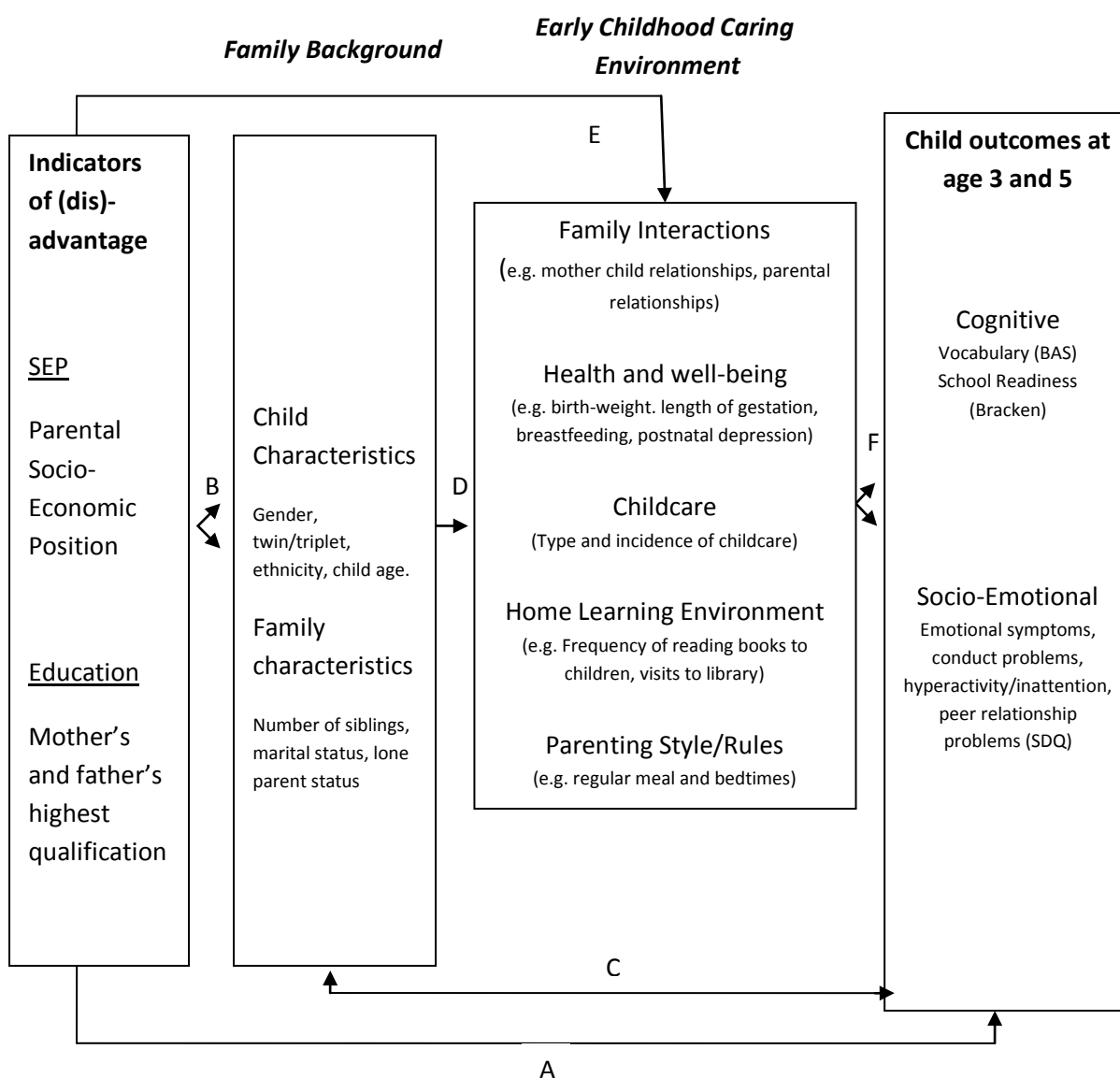
age of five. These factors will mediate the SEP gradient both via the effects of SEP on the early childhood caring environment (arrow E), and the early childhood caring environment on child outcomes (arrow F). However, these factors might also mediate the effects of other distal factors such as family background (arrow C) via a combination of arrows D and F.

Our choice of transmission factors reflects the diverse social science literature on the determinants of educational success, as a well as data availability. Empirical research in this area is far from definitive, but a number of potential pathways are frequently discussed in the literature (e.g. the home environment, quality of child care, parent-child relationships etc), many of which have been included in the present theoretical framework.

Several studies have found that differences in the home environment, as measured by the HOME scale (which includes items on household resources, such as reading materials and toys, and parental practices, such as discipline methods), account for a substantial portion of the effect of income on the cognitive development of pre-school children and on the achievement scores of elementary school children (Brooks-Gunn et al 1993; Duncan, Brooks-Gunn and Klebanov 1994; Korenman, Miller and Sjaastad 1995). Similar findings have been obtained by the Effective Provision of Pre-school Education (EPPE) Project in the United Kingdom, which found that although other family factors are also highly significant, the early years HLE (i.e. frequency of self-reported parental involvement in a range of activities, such as reading, library visits, playing with letters/numbers etc.) exerts a significant and independent influence on several cognitive (i.e. attainment at ages 3, 5 and 11) and behavioural (i.e. self-regulation, pro-social behaviour, and hyperactivity) outcomes. However, the EPPE HLE index is only moderately correlated ( $r=0.33$ ) with family socio-economic status (SES) or mother's qualification levels (Sylva et al 2008; Melhuish et al 2008).



Figure 1. Simple model linking financial position to early education outcomes



Another important pathway involves the health and well-being of the child and parents (e.g. birth-weight, gestation, breast-feeding patterns and indicators of post-natal depression). A 1990 analysis in the USA indicated that the poverty-related health factors such as low birth weight, elevated blood lead levels, anaemia, and recurrent ear infections and hearing loss contributed to the differences in IQ scores between poor and non-poor four-year-olds (Goldstein 1990). The findings suggest that the cumulative health disadvantage experienced by poor children on these health measures may have accounted for as much as 13% to 20% of the difference in IQ between the poor and non-poor children during the 1970s and 1980s (Goldstein 1990). Parents who are poor are also likely to be less healthy themselves, both emotionally and

physically (Adler, Boyce, Chesney, Folkman and Syme 1993), and this could result in impaired parent-child interactions and fewer home learning experiences (Bornstein 1995). For example, a study conducted by the USA National Longitudinal Survey of Youth (NLSY) found that currently poor mothers spanked their children more often than non-poor mothers, and this harsh behaviour was an important component of the effect of poverty on children's mental health (McLeod and Shanahan 1993).

A third possible pathway is through the care young children receive outside the home, as much research has shown that high-quality, developmentally appropriate child care in the pre-school years is associated with enhanced social, emotional, and linguistic competence (Howes 1988; Hofferth and

Phillips 1991; NICHD Early Child Care Research Network 1997, 1998; Ramey and Ramey 1998). In addition, randomized trials have demonstrated that intensive early childhood programmes for poor children can increase verbal ability and reasoning skills through early elementary school (Belfield, Nores, Barnett and Schweinhart 2006; Brooks-Gunn et al 1994; Burchinal, Campbell, Bryant, Wasik and Ramey 1997; Garces, Thomas and Currie (2002); Lazar and Darlington 1982; Ludwig and Miller 2007; Ramey and Ramey 1998).

One crucial mechanism we are unable to consider is the inter-generational transmission of cognitive ability and social skills. To the extent that parental ability and social skills are also correlated with the measures of the early childhood environment we do observe, their omission may lead to biased conclusions about the extent to which the observed factors can explain the socio-economic gaps in educational attainment. Crawford, Goodman and Joyce in this Special Issue use the British Cohort Study to examine the influence of parental cognitive ability and social skills on children's cognitive skills. They conclude that parental cognitive ability (measured during the parent's childhood) is a statistically significant predictor of children's cognitive skills, and that it explains about one sixth of the SEP gap in those skills, even after controlling for a rich set of demographic, attitudinal and behavioural factors. However, the addition of these parental characteristics to their model does very little to alter their impression of the relative importance of other proximal factors.

### 3. Data and descriptive statistics

The Millennium Cohort Study (MCS) began as a longitudinal study of approximately 18,000 children

born in the UK in 2000. The first sweep of the study was conducted when MCS children were about 9 months old. This over-sampled individuals from ethnic minorities and individuals living in disadvantaged areas of the country. Three further sweeps of data have since been collected when the children were aged about 36 months (sweep 2), when they were about 5 years old (sweep 3) and when they were 7 years old (sweep 4). Further sweeps of data will be collected at future key milestone ages. For our analysis, we chose to sample those who responded to the first three surveys and those where the mother is the main respondent<sup>1</sup>. We also excluded individuals who had missing data for some key characteristics, such as education, measures of the HLE, child outcomes and ethnicity. This left us with approximately 11,100 observations.

We proxy for each child's socio-economic background using two measures: a socio-economic position index and parent(s) highest educational qualifications. As described in detail in Gregg et al (2010), we construct our index of socio-economic position by performing principal-component analysis on a number of potential proxies for socio-economic background (average log equivalised income at sweeps 2 and 3, mother's class, father's class, housing tenure and whether the family have experienced financial difficulties). We then take the first principal component and use this as an index of socio-economic position. The factors used are summarised in Table 1 below, across quintiles of this index of socio-economic position. The data is weighted to take account of both the sampling design and non-response, as it is throughout this paper.

**Table 2. Factors included in the index of socio-economic position**

	SEP	SEP	SEP	SEP	SEP	Gap
Avg. Log Equivalised Income (MCS2 and MCS3)	4.93	5.26	5.55	5.87	6.39	1.46**
Mother – Professional/Managerial (MCS1)	1.4%	7.4%	22.7%	50.6%	84.1%	82.6 ppts**
Mother – Intermediate Occupation (MCS1)	4.6%	14.9%	30.9%	32.7%	14.6%	10 ppts**
Mother – Small Employer (MCS1)	1.2%	6.0%	7.5%	6.6%	1.1%	-0.1 ppts
Mother – Lower Supervisory (MCS1)	7.0%	8.7%	8.0%	4.1%	0.1%	-6.8 ppts**
Mother – Routine Occupation (MCS1)	85.8%	63.1%	31.0%	6.1%	0.1%	-85.6 ppts
Father – Professional/Managerial (MCS1)	2.6%	12.8%	28.7%	55.6%	94.7%	92.1 ppts**
Father – Intermediate Occupation (MCS1)	1.4%	3.9%	8.2%	9.0%	2.8%	1.4 ppts**
Father – Small Employer (MCS1)	7.1%	17.4%	19.9%	15.3%	2.2%	-4.9 ppts**
Father – Lower Supervisory (MCS1)	18.1%	25.4%	21.8%	14.0%	0.2%	-17.9 ppts**
Father – Routine Occupation (MCS1)	70.8%	40.5%	21.5%	6.1%	0.1%	-70.7 ppts**
Rented Privately in one or more sweep	14.0%	12.2%	7.5%	2.9%	0.5%	-13.4 ppts**
In Social Housing in one or more sweep	86.0%	31.4%	4.5%	0.5%	0.0%	-86 ppts**
Experienced financial difficulties	41.9%	19.2%	12.3%	5.4%	0.1%	-41.8 ppts**

\*\* and \* indicate significant differences between Q1 and Q5 at the 1% and 5% levels, respectively.

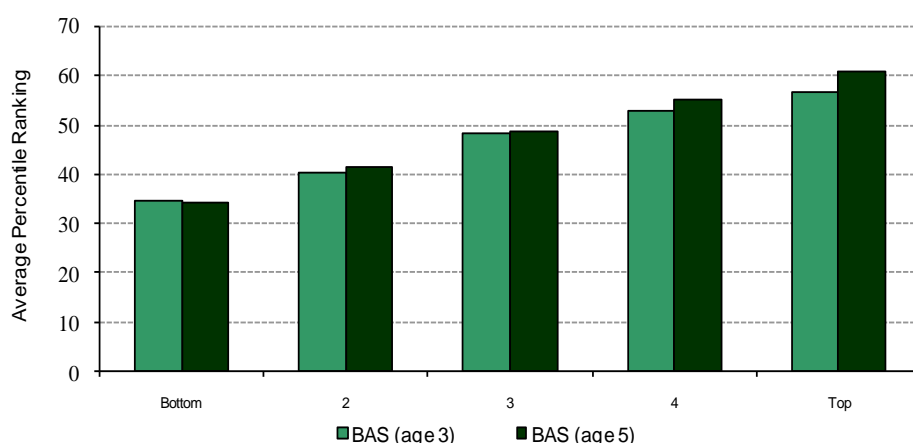
We focus on socio-economic differences in a specific measure of cognitive development: the British Ability Scales (BAS) (Early Years version) Naming Vocabulary. In our more detailed working paper (Dearden et al 2010), we also examine differences in the Bracken School Readiness measure, as well as differences in socio-emotional development as captured by the Goodman Strength and Difficulties questionnaire.

The BAS is well recognised as an excellent measure of children’s vocabulary, and is highly correlated with other language measures as well. It is also one of the best predictors of children’s general intelligence, which is highly stable, as there may well

be a strong genetic component to both cognitive and linguistic skills. On the other hand, a measure of school readiness, such as the Bracken School Readiness measure, is much more amenable to positive change through early childhood education programmes. For ease of interpretation, we have converted all outcome measures into percentile ranks.

Figure 2 shows the average percentile rankings of BAS scores at ages 3 and 5 for these five SEP quintiles. At age 3, the gap between the top and bottom quintiles for the BAS stands at about 22 percentile points, but by age 5, the gap has widened to over 26 percentile points.

Figure 2. BAS Scores by SEP Quintile



In Table 3, we investigate the dynamics of the BAS scores in a bit more detail. As we would expect, children in lower SEP quintiles are much more likely than those in higher quintiles to be in the bottom 40% of BAS scores at age 3. However, lower quintiles are also much less likely to escape the bottom 40% by age 5, as only around 25% of these children were able to do so (compared to about 60% in the top quintile group). Again, as expected, children from lower quintiles are less likely to be in the top 40% of

BAS scores at age 3, but those who attained that level of vocabulary skills are more likely than children in higher quintiles to drop out later<sup>ii</sup>. So not only do children from lower SEP quintiles have lower average outcomes at age 3, but they are also more likely to stay in the bottom 40% of achievers and more likely to drop out of the top 40%. The aim of the rest of this paper is to try and explain the socio-economic gaps in child outcomes at age 3 and age 5, as well as the widening of the gap between ages 3 and 5.

Table 3. Cognitive ability at ages 3 and 5, by parents’ SEP

Proportion of children:	SEP Q1	SEP Q2	SEP Q3	SEP Q4	SEP Q5
Bottom 40% at Age 3 (BAS)	58.5	50.0	37.8	31.5	27.0
Escape from bottom 40% by Age 5 (BAS)	25.4	33.1	41.8	53.7	60.7
Top 40% at Age 3 (BAS)	24.6	33.4	45.8	52.1	58.3
Drop out of top 40% by Age 5 (BAS)	57.9	48.3	40.8	34.1	29.8
Bottom 40% at Age 5 (BAS)	61.9	51.8	39.8	30.3	22.3
Top 40% at Age 5 (BAS)	21.7	30.6	42.2	51.8	59.5

### Transmission Mechanisms

Here we describe the groups of proximal influences that we have chosen to examine as potential transmission mechanisms.

**Parental Education** consists of variables that measure parent(s) highest educational qualification. We measure this as the highest equivalent NVQ level for both mothers and, where present, fathers. Dummy variables are included for other qualifications that could not be categorised into an NVQ level.

**Family Background** consists of variables pertaining to characteristics of the child and the family. Child characteristics include: the sex of the child; their age; whether they are a twin or triple; whether the child was in special care unit just after birth; and their ethnicity. Characteristics of the rest of the family include: mother's age at birth (plus a quadratic term); a separate teenage mother dummy variable; parental employment at sweeps 1, 2 and 3; whether only English is spoken in the household at sweep 1; marital status and partnering of the parents at sweeps 1, 2 and 3; the number of siblings and older siblings at sweeps 2 and 3.

The next groups of variables are those we collectively refer to as measure of the early childhood caring environment.

**Family Interactions** are measured by indicators of relationship quality at sweeps 1 and 2, based on the number of questions where respondents indicate a problem in their relationship with their partner (using a shortened version of the Golombok Rust Inventory of Marital State with 7 items at sweep 1 and 4 items at sweep 2, e.g. Are you on the brink of separation?), which is standardised to have a mean of zero and a standard deviation of one; interviewer-assessed measure of mother-child closeness at sweep 2, defined by number of items where mother and child appeared close out of a total of six items (e.g. whether the mother spontaneously praised their child during the interview), which is then standardised to have a mean of zero and standard deviation of one); mother-child relationship and conflict problems, measured as the number of problems identified on two subscales of the Pianta index (7 item parent-child relationship scale, e.g. do you share an affectionate/warm relationship with your child, and the 7-item parent-child conflicts scale, e.g. Does your child easily become angry with you), which are both standardised to have a mean of zero and standard deviation of one; and, finally,

whether mothers and fathers felt that they spent plenty of time with their children at the second sweep.

**Health and Well-Being** is measured by the number of cigarettes smoked by the mother per day before and after becoming pregnant; units of alcohol mothers reported drinking whilst pregnant and at the time of sweep 1; length of gestation (days); birth-weight (kg); did the mother try to breast-feed; age at which breast-feeding ceased (0-26 weeks); still breast-feeding after 26 weeks; whether the mother was suffering from post-natal depression during the child's early infancy (defined as agreeing to four of the nine statements from a shortened version of the Malaise inventory, e.g. do you often get into a violent rage?); mother reported indicators of infant temperament with three sub-scales relating to children's mood (5-tems, e.g. Is your child your child content during interruptions of milk or solids), the regularity of day-to-day routines (e.g. does your child nap for about the same length from day to day) and adaptability to new situations (5-items, e.g. Is your child still wary or frightened of strangers after 15 minutes); the mother's height (cm), weight (kg) and body mass index before birth; and the father's height (cm), weight (kg) and body mass index at sweep 1.

**(Non-Maternal) Childcare** indicators are constructed from sweeps 2 and 3 information to show whether parents mainly used informal care settings (grand-parents, other relatives, other non-relatives), and whether children had attended nursery school/class, playgroup, pre-school or childminder by the ages of 3 (sweep 2) and 5 (sweep 3).

**Home Learning Environment** at age 3 is measured by an index constructed from the reported frequency with which parents undertake 7 specific activities with their children (frequency of reading to child, library visits, play with ABCs/letters, teaches numbers/shapes, songs/nursery rhymes, drawing/painting). These 7 items are added together to give an index with a maximum score of 49, which is then divided into five equally-sized quintiles. We construct a similar index of the home-learning environment at age 5, which is based on a similar set of parenting activities (frequency of reading to child, tells stories to their child, takes them to an outdoor playground, parenting play with toys with their children, parents play sport with their children, songs/nursery rhymes, drawing/painting). These measures of the HLE are comparable to those used in

Melhuish et al (2008). We also separately control for self-reported parenting competency; frequency of reading to child (which is also included in HLE index). Including frequency of reading in addition to the overall HLE index allows us to see whether reading has a differential effect to the others elements. All measures are available at both sweeps 2 and 3.

**Parenting styles/rules** measures whether: the parents have lots of rules (at sweep 2); parents strictly enforce rules (sweep 2); children have regular bed-times (sweeps 2 and 3); children have regular meal times (sweeps 2 and 3); family eats breakfast together (sweep 3); children watch more than 3 hours of TV a day (sweeps 2 and 3); children play computer for more than 1 hour per day (sweep 3).

Table 4 then shows socio-economic differences in parental education, family background and these measures of the early childhood caring environment. As one might expect, mothers and fathers in the bottom quintile are less highly qualified than those from higher quintiles. For example, less than 7% of mothers in the bottom quintile have educational qualifications equivalent to NVQ level 4 or higher (i.e. university or equivalent), as compared with over 70% of the highest quintile).

Children from lower SEP quintiles are generally more likely to be from ethnic minorities (though Indian children are most likely to come from higher SEP quintiles). There are bigger differences across the SEP quintiles when we move on to examine differences in characteristics of the rest of the family. Mothers in the lowest SEP quintile were, on average, younger at the time of birth than those from higher quintiles. Mothers (and fathers, where present) are more likely to have been in employment across the three sweeps of the MCS. Nearly three quarters of mothers from the richest SEP quintile were in work at the time of the third sweep, as compared with about 30% of the poorest quintile. Children from the poorest quintile are less likely to speak only English at home as compared with richest quintile.

Looking at measures of family size and structure, we observe that children in the bottom SEP quintile are more likely to have come from lone-parent families than those from higher quintiles at the first MCS sweep (about 37%, as compared with just under 1% of the top SEP quintile). If two parents are present, parents in the bottom SEP quintile are also more likely to be cohabiting than those from higher quintiles (about a third at the first sweep, as compared with a little over 12% of the richest

quintile). Children from the bottom SEP quintile, on average, also have a greater number of siblings (numbers shown in table are for the third sweep).

The table then moves on to examine differences in our measures of the early childhood caring environment. The first panel relates to family interactions. Lower SEP quintiles seem to experience more mother-child relationship and conflict problems than richer SEP quintiles. Where they are together, there is also a greater degree of relationship problems amongst parents in poorer SEP quintiles than in the richest SEP quintile.

Mothers in the richest SEP quintile are more likely to try breast-feeding than those from poorer quintiles (90% of the richest quintile compared with a little over 50% of the poorest quintile). If they do breast-feed, mothers from the richest SEP quintile are likely to breast-feed for longer as well. However, mothers in the bottom quintile do not, on average, report consuming more units of alcohol during and after pregnancy. Mothers in the bottom quintile are, however, more likely to smoke during and prior to pregnancy. Mothers from the poorest quintile also have a shorter length of gestation and their children have lower birth-weights<sup>iii</sup>. Children from poorer families are also less likely to display a happy mood, appear less adaptable to new situations and display less regularity in daily routines. It is also noteworthy that whilst 7% of mothers from the richest quintile suffered from maternal depression during early infancy, about 22% of mothers from the poorest quintile did so. Parents from the richest quintile also report being taller at birth than parents from the poorest quintile, and report being slightly heavier. When we examine differences in Body Mass Index, we find that mothers and fathers in the poorest SEP quintiles are both more likely to be obese and more likely to be under-weight, i.e. are less likely to be classed as normal weight.

Examining childcare patterns up to sweep 2, we find that children from the poorest SEP quintile are more likely to have been to a nursery school or class than those from richer SEP quintile, but less likely to have been to a pre-school, childminder, day nursery or crèche. Children from the middle SEP quintile are those most likely to have been to a playgroup. These patterns continue up to sweep 3, though (as we would expect) noticeably more children have been to a nursery school or class between sweeps 2 and 3. At both sweeps, poorer parents are more likely to be

relying on sources of informal childcare, grandparents in particular.

The next part of the table examines socio-economic differences in the home-learning environment. We find that children from the poorest SEP quintiles are least likely to be in the richest quintile of our HLE index at sweeps 2 and 3. They are less likely to be read to every day at sweeps 2 and 3. Nearly 80% of children in the richest SEP quintile are read to every day at sweep 2, compared with 43% of the poorest SEP quintile.

The last block of the table shows that children from the richest quintile are more likely to have lots of rules and for these to be strictly enforced at sweep 2. They are also the most likely to have regular bed-times or meal times at sweep 2 (about 92% of these children have a regular bed-time

compared with around 69% of those from the poorest SEP quintile). These differences continue up to sweep 3, though are less dramatic. Children from the poorest SEP quintile are also more likely to watch more than 3 hours of TV a day at sweeps 2 and 3 than those from higher quintiles, are more likely to play more than one hour of computer a day at sweep 3.

Therefore, there are a wide variety of socio-economic differences across a range of factors, from number of siblings, birth-weight and reading frequency, to regularity of bedtimes, mother's age at birth and childcare patterns. But which of these factors explain the socio-economic gaps in child outcomes at ages 3 and 5 that we saw earlier. This is the focus of the next section.

**Table 4 - Socio-economic differences in family and child characteristics/behaviours**

	SEP Q1 (Low)	SEP Q2	SEP Q3	SEP Q4	SEP Q5 (High)	Gap (High – Low)
<b>Parental Education (MCS1)</b>						
Mother - no qualifications	32.2%	15.8%	6.8%	2.4%	0.8%	-31.3 ppts**
Mother NVQ level 1	16.8%	12.4%	6.6%	3.3%	0.9%	-15.8 ppts**
Mother NVQ level 2	34.4%	41.3%	37.2%	25.9%	13.5%	-20.8 ppts**
Mother NVQ level 3	10.2%	15.3%	19.4%	17.7%	11.0%	0.8 ppts
Mother NVQ level 4/5	6.5%	15.1%	30.1%	50.7%	73.7%	67.2 ppts**
Mother - other qualifications	3.4%	3.2%	1.2%	1.4%	0.5%	-2.8 ppts**
Father - no qualifications	44.8%	31.9%	18.1%	9.3%	3.9%	-40.9 ppts**
Father NVQ level 1	12.3%	8.5%	6.2%	3.9%	1.1%	-11.1 ppts**
Father NVQ level 2	25.6%	32.0%	31.4%	23.7%	11.4%	-14.1 ppts**
Father NVQ level 3	10.8%	15.1%	18.3%	17.6%	11.4%	0.6 ppts
Father NVQ level 4/5	6.5%	12.5%	26.0%	45.5%	72.2%	65.6 ppts**
Father - other qualifications	0.0%	0.1%	0.0%	0.0%	0.0%	0 ppts
<b>Child Characteristics</b>						
Male	48.0%	49.1%	48.9%	51.4%	49.7%	1.6 ppts
Child's age at MCS3 (years)	5.21	5.19	5.20	5.20	5.20	-0.01
Multiple birth	2.1%	3.0%	2.3%	3.0%	2.8%	0.7 ppts
Special care unit	8.7%	8.8%	9.5%	9.6%	9.3%	0.6 ppts
MCS1 White	85.5%	86.7%	92.9%	94.3%	94.0%	8.5 ppts**
MCS1 Indian	0.6%	2.3%	1.8%	1.6%	1.5%	0.8 ppts**
MCS1 Pakistani	3.1%	4.3%	1.5%	1.0%	0.4%	-2.7 ppts**
MCS1 Bangladeshi	1.3%	0.5%	0.2%	0.2%	0.0%	-1.2 ppts**
MCS1 Black Caribbean	1.9%	1.1%	0.6%	0.4%	0.4%	-1.4 ppts**
MCS1 Black African/Other	2.2%	1.2%	0.4%	0.3%	0.3%	-1.8 ppts**
MCS1 Other ethnicity	0.7%	1.1%	0.7%	0.6%	0.7%	0 ppts
MCS1 Mixed ethnicity	4.8%	2.7%	2.0%	1.8%	2.7%	-2.1 ppts**

(table 4 cont'd)

<b>Family Characteristics</b>						
Mother age at MCS1 (years)	24.9	27.4	29.7	31.0	32.2	7.29**
Mother worked at MCS1 or MCS2	32.7%	59.8%	74.8%	77.7%	81.2%	48.5 ppts**
Mother worked at MCS3	30.5%	53.8%	71.6%	76.3%	74.2%	43.7 ppts**
Father worked at MCS1 or MCS2	77.7%	95.5%	98.9%	99.4%	99.7%	22.0 ppts**
Father worked at MCS3	49.7%	71.4%	83.7%	89.0%	93.6%	43.8 ppts**
Only English spoken at home	91.2%	90.1%	94.2%	94.4%	94.5%	3.3 ppts**
Lone parent at MCS1	36.7%	16.6%	6.6%	2.9%	1.3%	-35.3 ppts**
Teenage mother	13.9%	6.4%	1.6%	0.2%	0.0%	-13.8 ppts**
Cohabiting at MCS1	34.3%	33.2%	22.7%	18.8%	12.2%	-22.0 ppts**
Lone parent at MCS2	41.4%	18.8%	8.3%	4.3%	2.5%	-38.9 ppts**
Cohabiting at MCS2	25.2%	23.1%	16.1%	12.3%	8.1%	-17.0 ppts**
Lone parent at MCS3	41.4%	21.8%	11.1%	6.7%	4.3%	-37.1 ppts**
Number of siblings at MCS3	1.60	1.42	1.30	1.21	1.14	-.46**
Number of older siblings at MCS3	1.09	0.98	0.85	0.73	0.61	-.48**
<b>Family Interactions</b>						
Mother-child relationship problems (sds)	0.21	0.04	-0.03	-0.11	-0.11	-.31**
Mother-child conflict problems (sds)	0.10	-0.01	-0.01	-0.03	-0.04	-.14**
Interviewer assessed measure of closeness (sds)	-0.23	-0.05	0.04	0.10	0.14	.37**
Relationship problems at MCS1 (sds)	0.31	0.17	0.09	-0.03	-0.07	-.37**
Relationship problems at MCS2 (sds)	0.32	0.15	0.08	-0.01	-0.06	-.38**
Mother spends plenty of time with child (MCS2)	78.7%	71.7%	62.5%	58.3%	51.1%	-27.6 ppts**
Father spends plenty of time with child (MCS2)	33.3%	20.5%	20.9%	18.4%	13.5%	-19.8 ppts**
<b>Health and Well-Being</b>						
Tried to breast-feed Child	50.6%	64.8%	73.7%	82.9%	89.9%	39.2 ppts**
Age at which breast-feeding stopped (weeks)	9.69	11.74	13.35	15.26	16.75	7.06**
Still breast-feeding at MCS1	10.2%	17.8%	24.4%	32.3%	38.8%	28.6 ppts**
Alcohol units consumed by mother in pregnancy	0.45	0.43	0.34	0.39	0.48	2.8 ppts
Mother alcohol consumption at wave 1 (units)	2.06	2.03	2.17	3.06	3.97	1.9**
Smoking by mother during pregnancy (# cigs)	4.61	2.47	1.19	0.50	0.29	-4.32**
Smoking by mother after pregnancy (# cigs)	8.60	5.69	3.45	1.96	1.19	-7.4**
Gestation length in days	276.1	276.8	277.2	277.3	277.5	1.43**
Birth Weight (kg)	3.26	3.33	3.39	3.42	3.45	0.18**
Infant temperament - mood (sds)	0.03	-0.01	0.06	-0.01	-0.08	-0.11**
Infant temperament - regularity (sds)	-0.32	-0.09	0.09	0.11	0.20	0.51**
Infant temperament - adaptability (sds)	-0.13	-0.05	0.03	0.02	0.13	0.26**
Mother suffered post-natal depression	21.8%	14.5%	11.8%	8.3%	6.7%	-15.1 ppts**
Mother height at birth (cm)	163.15	163.80	164.19	164.63	165.49	2.33**
Father height at birth (cm)	177.15	177.45	178.49	179.13	179.74	2.59**
Mother weight at birth (kg)	63.15	64.07	64.65	64.39	63.90	0.75 ppts
Father weight at birth (kg)	80.24	81.98	84.21	84.25	83.79	3.55**
Father under-weight	1.8%	0.8%	0.2%	0.4%	0.3%	-1.4 ppts**
Father normal-weight	37.0%	35.3%	32.7%	34.9%	41.4%	4.4 ppts**
Father over-weight	27.1%	33.0%	42.0%	45.9%	45.5%	18.3 ppts**
Father obese	11.3%	11.8%	12.9%	11.6%	9.0%	-2.3 ppts*
Mother under-weight	8.9%	6.6%	3.4%	3.4%	2.6%	-6.3 ppts**
Mother normal-weight	54.4%	57.9%	63.8%	67.5%	72.3%	17.9 ppts**
Mother over-weight	17.3%	20.9%	19.9%	18.7%	16.8%	-.5 ppts
Mother obese	10.9%	9.7%	9.5%	7.7%	5.6%	-5.2 ppts**

(Table 4 cont'd)

<b>Childcare</b>						
Cared for by grandparents MCS2	17.0%	24.3%	31.4%	32.6%	25.7%	8.7 ppts**
Cared for by other relatives MCS2	4.5%	7.3%	5.7%	5.0%	3.7%	-.8 ppts
Cared for by friends/neighbours MCS2	1.7%	0.9%	1.1%	1.1%	2.0%	.3 ppts
Has been to nursery school/class MCS2	27.3%	21.8%	21.1%	22.0%	23.7%	-3.5 ppts**
Has been to playgroup MCS2	25.2%	30.0%	31.1%	30.5%	25.7%	.4 ppts
Has been to pre-school MCS2	9.8%	13.6%	18.8%	17.7%	17.5%	7.7 ppts**
Has been to childminder MCS2	3.1%	5.6%	7.4%	11.3%	11.2%	8.1 ppts**
Has been to day nursery or creche MCS2	7.1%	8.8%	11.9%	17.3%	23.6%	16.5 ppts**
Cared for by grandparents MCS3	22.0%	27.5%	29.7%	32.6%	27.5%	5.5 ppts**
Cared for by relatives MCS3	10.2%	10.3%	7.8%	6.3%	3.5%	-6.7 ppts**
Cared for by friends MCS3	5.0%	6.2%	6.4%	8.0%	11.4%	6.4 ppts**
Has Been to nursery school/class MCS3	66.4%	60.1%	53.1%	53.7%	49.4%	-16.9 ppts**
Has Been to playgroup MCS3	20.6%	25.5%	26.9%	25.9%	21.6%	.9 ppts
Has Been to pre-school MCS3	13.9%	21.7%	28.5%	30.7%	29.9%	15.9 ppts**
Has Been to childminder MCS3	3.0%	4.2%	5.8%	7.4%	5.6%	2.5 ppts**
Has Been to day nursery or creche MCS3	6.2%	7.6%	10.6%	14.7%	20.4%	14.2 ppts**
<b>Home-Learning Environment</b>						
Bottom HLE quintile at MCS2	31.8%	25.7%	22.1%	20.7%	17.5%	-14.3 ppts**
2nd HLE quintile at MCS2	16.8%	18.7%	17.8%	17.0%	16.5%	-.3 ppts
3rd HLE quintile at MCS2	19.5%	20.5%	21.2%	20.5%	22.8%	3.3 ppts**
4th HLE quintile at MCS2	17.4%	16.6%	17.4%	20.3%	20.5%	3.1 ppts**
Top HLE quintile at MCS2	14.5%	18.5%	21.5%	21.5%	22.7%	8.1 ppts**
Read to everyday at MCS2	43.1%	51.8%	63.3%	72.4%	78.6%	35.5 ppts**
Read to some days at MCS2	45.5%	40.1%	32.9%	25.2%	20.0%	-25.4 ppts**
Bottom HLE quintile at MCS3	27.2%	25.2%	24.2%	19.8%	18.2%	-8.9 ppts**
2nd HLE quintile at MCS3	18.5%	16.8%	19.8%	18.3%	18.4%	-.1 ppts
3rd HLE quintile at MCS3	19.5%	21.1%	21.7%	23.7%	22.6%	3 ppts*
4th HLE quintile at MCS3	16.0%	18.5%	17.3%	20.3%	21.1%	5 ppts**
Top HLE quintile at MCS3	18.7%	18.5%	17.0%	17.9%	19.7%	.9 ppts
Read to everyday at MCS3	42.1%	47.7%	51.4%	57.0%	62.7%	20.5 ppts**
Read to some days at MCS3	48.8%	47.1%	43.8%	40.5%	36.1%	-12.7 ppts**
Mother rates herself as good parent MCS2	15.6%	20.8%	25.3%	28.5%	37.0%	21.3 ppts**
Mother rates herself as very good parent MCS2	26.9%	29.6%	28.0%	28.7%	26.7%	-.1 ppts
Father rates himself as good parent MCS2	17.5%	19.2%	25.9%	29.9%	36.7%	19.2 ppts**
Father rates himself as very good parent MCS2	27.1%	30.7%	31.7%	30.6%	30.4%	3.2 ppts
<b>Parenting Style/Rules</b>						
Lots of rules - MCS2	27.4%	28.0%	31.1%	32.9%	34.3%	6.8 ppts**
Strictly enforced rules - MCS2	42.7%	46.0%	47.2%	52.4%	58.0%	15.3 ppts**
Regular bed-times at MCS2	68.7%	76.0%	82.1%	87.7%	92.1%	23.4 ppts**
Regular meal-times at MCS2	85.1%	89.4%	94.1%	96.2%	98.0%	12.8 ppts**
Watches > 3 hours TV a day - MCS2	29.9%	22.0%	14.5%	8.6%	6.0%	-23.8 ppts**
Watches > 3 hours TV a day - MCS3	21.4%	16.6%	13.2%	11.3%	8.2%	-13.2 ppts**
Plays computer > 1 hour a day - MCS3	31.0%	25.0%	20.8%	16.6%	11.3%	-19.7 ppts**
Regular bed-times at MCS3	84.2%	88.6%	92.2%	95.2%	96.4%	12.2 ppts**
Regular meal-times at MCS3	88.0%	92.4%	94.8%	97.3%	97.6%	9.6 ppts**
Eat breakfast together at MCS3	86.6%	90.6%	93.9%	96.1%	97.1%	10.4 ppts**

\*\* and \* indicate significant differences between Q1 and Q5 at the 1% and 5% levels, respectively.



## 4. Explaining socio-economic gradients in child outcomes

In this section, we examine how much of the observed socio-economic differences in child cognitive outcomes at the ages of 3 and 5 can be explained by measures of the early childhood caring environment, including: family interactions; health and well-being; (non-maternal) childcare; the home-learning environment; and parenting style and rules; and how much remains related to distal factors such as socio-economic position and other aspects of family background.

### Empirical Methodology

In order to decompose the socio-economic gap in early cognitive outcomes into the contribution from different sources, we use the decomposition methodology described in detail in Goodman, Gregg and Washbrook in this Special Issue.

We first set out the raw differences in each educational outcome, according to socio-economic position of the parents, focusing on the gap between the top and bottom quintiles (top-bottom gap) as well as the gap between the middle and bottom quintiles (middle-bottom gap). All gaps are expressed in percentile point terms.

We then attempt to isolate the role of each factor in explaining the socio-economic gradient in child outcomes, after controlling for all observable characteristics, by decomposing the raw gap in child outcomes into the amount explained by each variable. This is calculated as the coefficient on each variable (taken from a regression including all observable characteristics) multiplied by unconditional differences in that variable across quintiles of our socio-economic position index. We group similar mediating factors together. The sum of the amount explained by all groups of potential mediating factors, represents the total amount explained by observable characteristics; the rest of the raw gap is thus unexplained.

The relationships we estimate are unlikely to be causal. For our estimates to be the causal impact of these factors, we would have to argue (among other things) that our measures of family characteristics, and in particular our mediating factors, cannot be affected by the child outcomes of interest. However, it is highly likely that factors, such as the child's home learning environment, are going to be influenced by the child's cognitive development. We do not attempt to take into account this possible

simultaneity. Furthermore, we would also have to argue that there are no other unobserved characteristics of the child or family, that influence these mediating factors, as well as the child outcomes we measure. Again, this is unlikely to be true and this means that our estimates of the impact of different mediating factors are likely to be biased.<sup>iv</sup>

While our work (along with most other work in this area) cannot robustly establish the presence of direct causal links between these factors, we are fortunate to have an extremely rich dataset at our disposal, that allows us to observe in great detail a wide range of family background variables, family health and well-being, as well as parenting activities, relationships and behaviours, that serve as plausible transmission mechanisms between child poverty and poor early educational attainment. This should give us some clues as to possible policy responses to address the socio-economic gaps in early child outcomes, as well as avenues for future research.

We start by examining BAS vocabulary scores at age 3, before moving on to age 5 BAS vocabulary scores (with and without controls for prior scores at age 3). The full regression results on which these decompositions are based are shown in Table A1 in the Appendix.

### Age 3 - BAS (Vocabulary)

In the first row of Table 5, we show the raw gap at age 3 in BAS average percentile ranks between the top and bottom quintiles of our index of socio-economic position (21.9 ppts), as well as the gap between the middle and bottom quintiles (13.5 ppts). Both are statistically significant at the 1% level. The next two rows then show that about 69% of the socio-economic gaps in BAS scores at age 3 can be explained by differences in observable characteristics; whilst 31% remains unexplained. The rest of the table then decomposes the 69% explained by all observable characteristics, into how much is explained by differences in the early childhood caring environment, and how much remains related to parental education and family background (as well as a separate group for missing dummy variables). Within each group, we also further decompose the amount explained by each group into the amount explained by individual (or very similar) variables. We observe that:

- **Parental Education** explains about one sixth of the socio-economic gaps in BAS scores,

after controlling for all other observable characteristics. This is the amount that remains related to differences in parental education, which cannot be accounted for by differences in the early childhood caring environment.

- **Family Background** characteristics still explain about a quarter of the socio-economic gaps in BAS scores at age 3, conditional on all other observable characteristics. Looking at the detailed results, we see that 15-17% of the gaps are explained by the fact that mothers of children from higher quintiles are, on average, older than those from lower quintiles. Similarly, the facts that the top quintiles contains greater proportions of working fathers and that these children have fewer siblings, explain 7-10% and 6-8% of the socio-economic gaps, respectively. The age of children explains about 10% of the gaps, which means that the poorer children must have been tested at slightly later ages, since age has a positive effect on BAS percentiles.
- **Family Interactions** explain 4-5% of the socio-economic gaps, with the largest single amount being explained by greater levels of mother-child closeness among families in higher quintile.
- **Health and Well-Being** factors similarly explain about 3-5% of the gaps, with most of the positive contribution coming from

breast-feeding and birth-weight, offset by some negative factors (smoking before pregnancy and parental height/weight).

- **Childcare** patterns explain about 1% of the top-bottom gap, but about 4% of the middle-bottom gap.
- **The Home-Learning Environment** explains one sixth of the socio-economic gaps, most of which is done by the HLE index and reading frequency at age 3.
- **Parental Style/Rules** explains very little of the socio-economic gaps at age 3 in BAS percentile ranks.

Therefore, this decomposition shows that about a quarter of the socio-economic gaps in BAS scores at age 3 can be explained by differences in the early childhood caring environment, with the home-learning environment explaining about one sixth, and family interactions and health/well-being factors explaining about 5% each. However, around a quarter of the gap remains related to family background characteristics, mostly down to the fact that in higher quintiles, mothers are older, children have fewer siblings and fathers are more likely to be in work. And about a sixth remains related to differences in parental education, controlling for other observable characteristics. Around 30% of the gap cannot be explained by differences in any of the observable characteristics.

Table 5. Explaining socio-economic gaps in cognitive ability at age 3 (BAS)

	Percentile point gap		As % total gap	
	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1
<b>Raw Gap</b>	<b>21.90</b>	<b>13.54</b>	<b>100%</b>	<b>100%</b>
<b>Total: Explained</b>	<b>15.02</b>	<b>9.66</b>	<b>69%</b>	<b>71%</b>
<b>Total: Unexplained</b>	<b>6.88</b>	<b>3.88</b>	<b>31%</b>	<b>29%</b>
<b>Amount Explained by Factors:</b>				
<b><i>Parental Education</i></b>	<b>3.81</b>	<b>2.02</b>	<b>17%</b>	<b>15%</b>
<b><i>Family Background</i></b>	<b>5.19</b>	<b>3.72</b>	<b>24%</b>	<b>27%</b>
Gender	-0.11	-0.06	-1%	0%
Age of child	-1.89	-1.42	-9%	-10%
Twin/triplet	0.00	0.00	0%	0%
Special care unit after birth	-0.01	-0.01	0%	0%
Ethnicity	0.55	0.39	3%	3%
Only English spoken at home	0.38	0.35	2%	3%
Country of residence	-0.06	0.02	0%	0%
Mother works	0.39	0.34	2%	3%
Father works	1.46	1.31	7%	10%
Mother's age at birth	3.35	2.35	15%	17%
Marital/partner Status	-0.52	-0.38	-2%	-3%
Siblings	1.67	0.84	8%	6%
<b><i>Family Interactions</i></b>	<b>0.83</b>	<b>0.62</b>	<b>4%</b>	<b>5%</b>
Mother-child closeness	0.99	0.73	5%	5%
Parental harmony	0.00	0.00	0%	0%
Parental time	-0.16	-0.11	-1%	-1%
<b><i>Health and Well-Being</i></b>	<b>1.02</b>	<b>0.43</b>	<b>5%</b>	<b>3%</b>
Breast-feeding	0.72	0.40	3%	3%
Alcohol consumption	0.13	0.01	1%	0%
Smoking patterns	-0.13	-0.01	-1%	0%
Gestation length	-0.04	-0.04	0%	0%
Birth weight	0.54	0.39	2%	3%
Infant temperament	-0.01	-0.09	0%	-1%
Maternal depression	-0.09	-0.06	0%	0%
Parental height/weight	-0.10	-0.17	0%	-1%
<b><i>Childcare</i></b>	<b>0.33</b>	<b>0.52</b>	<b>1%</b>	<b>4%</b>
<b><i>Home-Learning Environment</i></b>	<b>3.60</b>	<b>2.18</b>	<b>16%</b>	<b>16%</b>
HLE and reading at Age 3	2.64	1.62	12%	12%
Self-reported parental competence	0.96	0.55	4%	4%
<b><i>Parenting Style/Rules</i></b>	<b>-0.04</b>	<b>-0.08</b>	<b>0%</b>	<b>-1%</b>
Amount/strictness of rules	0.02	-0.01	0%	0%
Regular bed times at age 3	0.47	0.27	2%	2%
Regular meal times at age 3	-0.05	-0.03	0%	0%
Watches lots of TV at age 3	-0.49	-0.32	-2%	-2%
<b><i>Missing Dummies</i></b>	<b>0.28</b>	<b>0.26</b>	<b>1%</b>	<b>2%</b>

Table 6. Explaining socio-economic gaps in cognitive ability at age 5 (BAS)

	No controls for prior ability				Controlling for prior ability			
	ppt terms		As % total gap		ppt terms		As % total gap	
	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1	Q5-Q1	Q3-Q1
<b>Raw Gap</b>	<b>26.57</b>	<b>14.19</b>	<b>100%</b>	<b>100%</b>	<b>26.57</b>	<b>14.19</b>	<b>100%</b>	<b>100%</b>
<b>Total Explained</b>	<b>19.87</b>	<b>11.94</b>	<b>75%</b>	<b>84%</b>	<b>23.58</b>	<b>14.05</b>	<b>89%</b>	<b>99%</b>
<b>Total Unexplained</b>	<b>6.70</b>	<b>2.25</b>	<b>25%</b>	<b>16%</b>	<b>2.99</b>	<b>0.14</b>	<b>11%</b>	<b>1%</b>
<b>Amount Explained by Factors:</b>								
<b>Prior Cognitive Ability</b>					<b>13.40</b>	<b>8.14</b>	<b>50%</b>	<b>57%</b>
<b>Prior Non-Cognitive Ability</b>					<b>0.31</b>	<b>0.21</b>	<b>1%</b>	<b>1%</b>
<b>Parental Education</b>	<b>6.68</b>	<b>3.09</b>	<b>25%</b>	<b>22%</b>	<b>4.09</b>	<b>1.89</b>	<b>15%</b>	<b>13%</b>
<b>Family Background</b>	<b>9.42</b>	<b>6.70</b>	<b>35%</b>	<b>47%</b>	<b>5.29</b>	<b>3.69</b>	<b>20%</b>	<b>26%</b>
Gender	-0.01	-0.01	0%	0%	0.04	0.02	0%	0%
Age of child	0.02	0.03	0%	0%	0.02	0.03	0%	0%
Twin/triplet	-0.03	-0.01	0%	0%	-0.02	-0.01	0%	0%
Special care unit after birth	0.00	0.00	0%	0%	0.00	0.00	0%	0%
Ethnicity	0.33	0.28	1%	2%	0.07	0.10	0%	1%
Only English spoken at home	0.44	0.40	2%	3%	0.27	0.25	1%	2%
Country of residence	0.04	0.03	0%	0%	0.07	0.04	0%	0%
Mother works	0.41	0.35	2%	2%	0.11	0.08	0%	1%
Father works	0.69	0.55	3%	4%	-0.30	-0.27	-1%	-2%
Mother's age at birth	4.07	2.86	15%	20%	2.52	1.75	9%	12%
Marital/partner Status	1.35	1.12	5%	8%	1.42	1.13	5%	8%
Siblings	2.11	1.11	8%	8%	1.10	0.58	4%	4%
<b>Family Interactions</b>	<b>0.77</b>	<b>0.60</b>	<b>3%</b>	<b>4%</b>	<b>0.14</b>	<b>0.16</b>	<b>1%</b>	<b>1%</b>
Mother-child closeness	0.74	0.55	3%	4%	0.16	0.12	1%	1%
Parental harmony	-0.01	0.01	0%	0%	-0.02	0.02	0%	0%
Parental time	0.04	0.04	0%	0%	0.00	0.02	0%	0%
<b>Health and Well-Being</b>	<b>-1.10</b>	<b>-1.11</b>	<b>-4%</b>	<b>-8%</b>	<b>-1.04</b>	<b>-0.89</b>	<b>-4%</b>	<b>-6%</b>
Breast-feeding	1.04	0.52	4%	4%	0.60	0.29	2%	2%
Alcohol consumption	0.15	0.07	1%	0%	0.10	0.06	0%	0%
Smoking patterns	-1.09	-0.73	-4%	-5%	-0.90	-0.63	-3%	-4%
Gestation length	-0.07	-0.07	0%	0%	-0.09	-0.08	0%	-1%
Birth weight	0.29	0.21	1%	1%	0.11	0.08	0%	1%
Infant temperament	0.26	0.13	1%	1%	0.17	0.09	1%	1%
Maternal depression	-0.03	-0.02	0%	0%	0.05	0.03	0%	0%
Parental height/weight	-1.64	-1.22	-6%	-9%	-1.08	-0.74	-4%	-5%
<b>Childcare</b>	<b>0.55</b>	<b>0.47</b>	<b>2%</b>	<b>3%</b>	<b>0.25</b>	<b>0.26</b>	<b>1%</b>	<b>2%</b>
<b>Home-Learning Environment</b>	<b>2.06</b>	<b>1.16</b>	<b>8%</b>	<b>8%</b>	<b>0.18</b>	<b>-0.03</b>	<b>1%</b>	<b>0%</b>
HLE and reading at Age 3	1.97	1.21	7%	9%	0.47	0.24	2%	2%
HLE and reading at Age 5	0.04	0.03	0%	0%	0.11	0.04	0%	0%
Self-reported parental competence	0.04	-0.08	0%	-1%	-0.40	-0.31	-1%	-2%
<b>Parenting Style/Rules</b>	<b>1.24</b>	<b>0.71</b>	<b>5%</b>	<b>5%</b>	<b>0.88</b>	<b>0.51</b>	<b>3%</b>	<b>4%</b>
Amount/strictness of rules	0.15	0.04	1%	0%	0.10	0.03	0%	0%
Regular bed times at age 3	0.31	0.18	1%	1%	0.06	0.03	0%	0%
Regular bed times at age 5	0.33	0.22	1%	2%	0.37	0.24	1%	2%
Regular meal times	0.34	0.24	1%	2%	0.08	0.06	0%	0%
Watches lots of TV/Computer	0.10	0.04	0%	0%	0.27	0.15	1%	1%
<b>Missing Data</b>	<b>0.27</b>	<b>0.32</b>	<b>1%</b>	<b>2%</b>	<b>0.09</b>	<b>0.12</b>	<b>0%</b>	<b>1%</b>

**Age 5 – BAS (Vocabulary) – Static Model**

Table 6 then shows the results of our decomposition methodology for analysing the socio-economic differences in BAS scores at age 5. As we have already shown, the top-bottom gap has widened to 26.6 percentile points, and the middle-bottom gap to 14.2 percentile points. The four columns under the heading “No controls for prior ability” show the results of our static decomposition. Specifically, it shows that 84% of the middle-bottom gap and 75% of top-bottom gap can be accounted for by differences in observable characteristics. The specific amounts explained by different variables are as follows:

- **Parental Education** still explains about a quarter of the socio-economic gaps, after controlling for other observable characteristics. This is more than was the case at age 3.
- **Family Background** characteristics still explain a further 35% of the top-bottom gap, but nearly 47% of the middle-bottom gap. The most important individual variables were mother’s age (15-20% of the gaps), number of siblings (8%), whether the father was in work (3-4%) and marital/partner status (6-9%). Only the latter was not important in explaining gaps at age 3.
- **Family Interactions** explain about 3-4% of the socio-economic gaps in BAS scores at age 5, with the most important variables again being those related to mother-child closeness. This is similar to what we found for BAS scores at age 3.
- **Health and Well-Being** factors explain a negative amount of the gap (i.e. they increase it) with positive contributions (from breast-feeding patterns) more than offset against negative ones (smoking patterns and parental height/weight).
- **Childcare** only explains about 2-3% of the socio-economic gaps.
- **The Home-Learning Environment** explains 8% of the socio-economic gaps. Interestingly, this is solely down to differences in the home-learning environment and reading frequency at age 3. The differences at age 5 are found to be unimportant.
- **Parenting Style/Rules** make a further small contribution of 5%, with small contributions from all of the individual variables.

**Age 5 - BAS (Vocabulary) – Value added model**

The last four columns of Table 6 show the results

of our decomposition methodology when we also include controls for prior cognitive ability and socio-emotional development (specifically, the Strengths and Difficulties questionnaire). This shows that observable characteristics explain 89% of the top-bottom gap and fully explain the middle-bottom gap. Specifically:

- **Prior cognitive ability** explains the majority of both socio-economic gaps. It explains 57% of the middle-bottom gap and 50% of the top-bottom gap.
- **Prior socio-emotional development**, however, only explains 1% of the gaps.
- **Parental Education** still explains a further 13-15% of the socio-economic gaps, conditional on all other observable characteristics and prior outcomes.
- **Family Background** characteristics still explain a further 20% of the top-bottom gap and 26% of the middle-bottom gap, conditional on other observable characteristics and prior outcomes. The individual variables making the largest contributions, are mother’s age at birth, marital/partner status and number of siblings. However, the importance of mother’s age at birth and number of siblings are halved as compared with the static decomposition at age 5, suggesting that part of their impact at age 5 is via higher levels of cognitive ability at age 3.
- **Family Interactions** only explain about 1% of the socio-economic gaps, conditional on other observable characteristics and prior outcomes.
- **Health and Well-Being** factors make a small negative contribution, overall. There are positive contributions from breast-feeding patterns, but negative ones from smoking patterns and parental height/weight.
- **Childcare** only contributes a further 1-2% to the socio-economic gap.
- **The Home-Learning Environment** hardly contributes anything at all, conditional on other observable characteristics and prior outcomes. Since it was important in the static decomposition, this suggests that differences in the home-learning environment only explain gaps at age 5 via improving cognitive ability at age 3.
- **Parenting Style/Rules** explain a further 3-4% of the gap, though no one variable appears to be particularly important.

## Summary of Results

In total, observable characteristics explain about 65-75% of the socio-economic gaps at age 3, and 77-87 % of socio-economic differences in cognitive outcomes at age 5.

At age 3, differences in the early childhood caring environment explain about 25% of the socio-economic gap. This can mostly be accounted for by differences in the home-learning environment and reading frequency. Other proximal factors, like family interactions and health and well-being, explain a small proportion of the socio-economic gap. However after conditioning on this rich set of proximal factors, it is family background factors that still explain the largest portion of socio-economic differences. Looking at individual family background factors, it is differences in mother's age, number of siblings and working patterns, that are found to explain the largest element of these socio-economic differences. The next largest contribution comes from the remaining effects of parental education that do not occur via the observable characteristics.

When we look at age 5 cognitive outcomes, we find that prior cognitive outcomes explain over 50% of the socio-economic differences at age 5, whilst prior socio-emotional development explains very little, if anything. The only other factors that explain a large proportion of the socio-economic gaps, after controlling for prior ability, are parental education and family background (again, mother's age, number of siblings and marital/partner status being most important). The influence of these items is much reduced compared with the static model, suggesting some of their impact occurs via their effect on cognitive ability at age 3. The home-learning environment is found to be unimportant in this decomposition, suggesting that it only explains age 5 outcomes through its influence on age 3 cognitive outcomes.

## Conclusions

In conclusion, we have confirmed that there are substantial differences in cognitive and socio-emotional development between children from rich and poor backgrounds even at the age of 3, and that this gap widens by the age of 5. Children from poor backgrounds also face much less advantageous "early childhood caring environments" than children from better off families. For example, we have observed significant differences in poor children's

and their mothers' health and well-being and the home learning environment.

We also find that differences in the home learning environment at age 3 have an important role to play, explaining about a sixth of the gap in cognitive outcomes between children from richer and poorer backgrounds. However, a much bigger proportion of the socio-economic gap remains directly related to other aspects of family background (such as mothers' age, and family size) that are not mediated through the early childhood caring environment, and a significant element remains explained.

It is noteworthy that it is the home-learning environment measured at age 3, that is found to be important in explaining outcomes at ages 3 and 5, the latter working through its impact on cognitive ability at age 3. The home-learning environment measured at age 5 is not estimated to impact on cognitive outcomes at age 5, or thus the gap in cognitive outcomes at age 5. This stresses the importance of early intervention. However, it is difficult to know with certainty whether policymakers can, a) change the home-learning environment, and (b) whether any shifts in the home-learning environment will reduce the gap in early child outcomes. This is partly because it is difficult to put a definite causal interpretation to our finding, and because the malleability of the home-learning environment to outside policy intervention is currently unknown. We therefore believe it to be essential that different methods of shifting the home-learning environment at early ages are trialled and evaluated in the UK at the earliest opportunity.

However, it is also worth noting that pre-determined factors still explain the largest element of the socio-economic divides in cognitive outcomes at age 3 and age 5, even when we condition on a rich set of measures of the early childhood caring environment. The most important factors being mother's age at birth, number of siblings, parental education and prior cognitive development (at age 3). With a view to closing socio-economic gaps in cognitive outcomes, these results underline the importance of early intervention, at least before age 3 and perhaps even prior to birth, if one believed the results that would suggest encouraging poorer mothers to delay the birth of their first child might narrow some of the socio-economic gap in early cognitive development.

Taken together, our findings suggest that policies to improve parenting skills and home learning environments in isolation cannot possibly eliminate the cognitive skills gap between rich and poor young children, though such policies could go some way towards reducing it. On the other hand, in our working paper we show that many aspects of the

early childhood caring environment do have a positive effect on children's social and emotional development, suggesting that policies aimed at improving health, parenting skills and the home-learning environment could have other important short- and long-term pay-offs.

## Acknowledgements

This work was carried out as part of a larger body of work funded by the Joseph Rowntree Foundation (JRF) through its programme of research on education and poverty. We gratefully acknowledge the support of JRF, and the generous and productive guidance of Helen Barnard and Chris Goulden. We are also thankful to members of the advisory group: Tim Crosier, Naomi Eisenstadt, Leon Feinstein, Zoe Ferguson, Tony Foot, Leslie Gutman, Lisa Harker, Andrew Ledger, Mark Newman, Ivan Turok, Anna Vignoles, Jane Waldfogel, and Stephen Witt. All errors remain the responsibility of the authors.

## References

- Adler NE, Boyce WT, Chesney MA, Folkman S and Syme SL. (1993) Socioeconomic inequalities in health: No easy solution. *Journal of the American Medical Association*, 269, 3140-3145.
- Belfield CR, Nores M, Barnett S and Schweinhart L. (2006) The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup, *Journal of Human Resources*, University of Wisconsin Press, 41, 162-190.
- Brooks-Gunn J, Duncan GJ, Klebanov PK and Sealant N. (1993) Do neighborhoods influence child and adolescent behavior? *American Journal of Sociology*, 99, 335-95.
- Brooks-Gunn J, McCarton C, Casey P, McCormick M, Bauer C, Bernbaum J, Tyson J, Swanson M, Bennett F, Scott D, Tonascia J and Meinert C. (1994) Early intervention in low birth weight, premature infants: Results through age 5 years from the Infant Health and Development Program. *Journal of the American Medical Association*, 272, 1257-1262.
- Burchinal MR, Campbell FA, Bryant DM, Wasik BH and Ramey CT. (1997) Early intervention and mediating processes in cognitive performance of children of low-income African-American families. *Child Development*, 68, 935-954.
- Dearden L, Mesnard A and Shaw J. (2006) Ethnic differences in birth outcomes in England. *Fiscal Studies*, 27, 17-46.
- Dearden L, Sibieta L and Sylva K. (2010) The socio-economic gradient in early child outcomes: evidence from the Millenium Cohort Study. *IFS Working Paper*.
- Duncan GJ, Brooks-Gunn J and Klebanov PK. (1994) Economic deprivation and early-childhood development. *Child Development*, 65, 296-318.
- Feinstein L. (2003) Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 70, 73-97.
- Feinstein L. (2004) Mobility in Pupils' Cognitive Attainment During School Life. *Oxford Review of Economic Policy*, 20, No. 2: Education.
- Garces E, Thomas D and Currie J. (2002) Longer-Term Effects of Head Start. *American Economic Review*, American Economic Association, 92, 999-1012.
- Goldstein H. (1990) Explaining socioeconomic differences in children's cognitive test scores. *Working Paper No. H-90-1*. Cambridge, MA: Malcolm Wiener Center for Social Policy, John F Kennedy School of Government, Harvard University.
- Goodman A and Gregg P. (eds) (2010) *Poorer children's educational attainment: how important are attitudes and behaviour?* Joseph Rowntree Foundation: York.
- Hofferth S and Phillips DA. (1991) Childcare policy research. *Journal of Social Issues*, 47 1-13.
- Howes C. (1988) Relations between early child care and schooling. *Developmental Psychology*, 24, 53-57.
- Korenman S, Miller JE and Sjaastad JE. (1995) Long-term poverty and child development in the United States: Results from the National Longitudinal Survey of Youth. *Children and Youth Services Review*, 17, 127-51.
- Lazar I and Darlington RB. (1982) Lasting effects of early education: A report from the consortium for longitudinal studies. *Monographs of the Society for Research in Child Development*, 47(2-3, Serial No. 195).
- Ludwig J and Miller DL. (2007) Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *Quarterly Journal of Economics*, MIT Press, 122, 159-208.

- Melhuish EC, Sylva K, Sammons P, Siraj-Blatchford I, Taggart B and Phan M. (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, 157-188.
- McLeod JD and Shanahan MJ. (1993) Poverty, parenting and children's mental health. *American Sociological Review*, 58, 351-66.
- NICHD Early Child Care Research Network. (1997) The effects of infant child care on infant-mother attachment security: Results of the NICHD study of early child care. *Child Development*, 68, 860-879.
- NICHD Early Child Care Research Network. (1998) Relations between family predictors and child outcomes: Are they weaker for children in child care? *Developmental Psychology*, 34, 1119-1128.
- Ramey CT and Ramey SL. (1998) Prevention of intellectual disabilities: Early interventions to improve cognitive development. *Preventive Medicine*, 27, 224-232.
- Sylva K, Melhuish E, Sammons P, Siraj-Blatchford I and Taggart B. (2008) Effective Pre-school and Primary Education 3-11 Project (EPPE 3-11): Final Report from the Primary Phase: Pre-school, School and Family Influences on Children's Development during Key Stage 2 (Age 7-11). Department for Children, Schools and Families, London.



## Appendix A

Table A1 – Full Specification Regression Results for BAS at age 3 and age 5

	BAS (age 3)	BAS (age 5)	
		Levels	Value Added
<b>Prior Ability</b>			
Bracken (age 3)	n/a	n/a	6.796***
BAS (age 3)	n/a	n/a	7.985***
SDQ (age 3)	n/a	n/a	0.384
<b>SEP Quintile</b>			
2nd SEP quintile	1.145	1.118	0.385
3rd SEP quintile	4.397***	2.254**	0.14
4th SEP quintile	6.331***	4.636***	1.397
Top SEP quintile	7.714***	6.696***	2.992**
<b>Parental Education</b>			
Mother NVQ level 1	1.92	1.372	0.492
Mother NVQ level 2	3.150***	2.344**	0.788
Mother NVQ level 3	3.183**	1.87	0.165
Mother NVQ level 4/5	5.141***	5.549***	2.650**
Mother - other qualifications	0.865	1.489	1.152
Father NVQ level 1	-0.767	1.34	0.98
Father NVQ level 2	0.496	1.997*	1.911*
Father NVQ level 3	1.298	3.582***	2.764***
Father NVQ level 4/5	1.794*	5.528***	3.903***
Father - other qualifications	5.473*	-10.889	-8.646
<b>Child Characteristics</b>			
Male	-6.764***	-0.543	6.796***
Child's age (months/100)	0.911***	-0.539*	7.985***
Multiple birth	-0.44	-3.917	0.384
Special care unit	-1.658	-0.384	2.441***
MCS1 Indian	-5.970**	1.207	-0.463
MCS1 Pakistani	-9.127***	-3.855*	-2.582
MCS1 Bangladeshi	-11.527***	-0.599	0.127
MCS1 Black Caribbean	-8.799***	-6.565**	2.611
MCS1 Black African/Other d	-2.936	-5.719**	1.691
MCS1 Other ethnicity	-9.867***	-3.637	6.148
MCS1 Mixed ethnicity	-0.811	-0.378	-2.657
<b>Family Characteristics</b>			
Mother's age at birth	0.373	1.661***	0.940**
Mother's age at birth squared	4.795***	-0.020***	-0.011*
Only English at Home	4.129***	13.051***	-5.678**
Lives in Wales	0.802	-2.936***	-2.239
Lives in Scotland	n/a	2.840**	-0.999
Lives in Northern Ireland	3.084**	0.941	8.078***
Mother worked at one of waves	n/a	1.119	-3.520***
Mother worked at wave 3		-0.307	0.395
Father worked at one of waves	1.329***	0.291	-0.634
Father worked at Wave 3	-0.016*	0.968	0.721
Lone parent at MCS1	0.665	-2.335	-2.318
Had baby in teens	0.226	-0.09	-0.562
Two cohabiting parents at MCS1	1.192	-0.963	-1.592**
Got married by MCS2	-0.063	1.3	1.251
Split up by MCS2	-2.493**	-2.173	-2.867*
New partner by MCS2	-1.088	0.505	2.079
Split up by MCS3	n/a	-1.689	-0.055
New partner by MCS3	n/a	0.941	-0.064
Number of siblings at MCS 2/3	0.199	-0.408	-0.189
Number of older siblings at MCS 2/3	-3.679***	-3.984***	-2.100***

**Family Interactions**

Mother-child relationship problems (sds)	-2.596***	-1.755	1.55
Mother-child conflict problems (sds)	-0.496***	2.389	2.778
Interviewer assessed measure of closeness (sds)	0.872***	-0.267	0.482
Relationship problems at MCS1 (sds)	-0.028	-4.535**	-1.946
Relationship problems at MCS2 (sds)	n/a	0.735**	0.065
Mother spends plenty of time with child - MCS2	1.008	-3.253	-2.91
Father spends plenty of time with child - MCS2	-0.713	-0.212	-0.933

**Health and Well-Being**

Tried to breast-feed Child	1.102	0.26	0.696
Age at which breast-feeding stopped (weeks)	0.002	0.091	2.733*
Still breast-feeding at MCS1	0.992	0.055	-0.316
Mother alcohol consumption during pregnancy (units)	0.159	-0.023	-0.108
Mother alcohol consumption (small amount)	0.581	1.681***	0.073
Mother alcohol consumption at Wave 1 (units)	0.033	-0.026	-0.347
Number of cigarettes smoked by Mother during (squared)	-0.279*	-0.195	-0.121
Number of cigarettes smoked by mother before (squared)	0.007	0.008	1.379**
Number of cigarettes smoked by mother before (squared)	0.222*	0.287**	-0.033
Gestation length in days	-0.005	-0.005	-0.076
Gestation length in days (squared)	0.639**	0.488	0.006
Gestation length in days (squared)	-0.001**	-0.001*	0.17
Birth weight (kg)	2.893***	1.584**	-0.003
Infant temperamen - mood (sds)	-0.619**	-0.518*	0.211
Infant temperament - regularity (sds)	-0.208	0.278	-0.001
Infant temperament - adaptability (sds)	0.724**	0.222	0.616
Mother suffered post-natal depression	0.83	0.223	0.279
Mother height at birth (cm)	-0.018	0.017	5.148
Father height at birth (cm)	-0.036*	-0.03	-0.322
Mother weight at birth (kg)	0.005	-0.027	-4.542
Father weight at birth (kg)	0.082*	0.039	0.007
Father under-weight	-1.113	2.407	-0.008
Father over-weight	-2.125**	-2.325***	-0.004
Father obese	-2.653	-0.647	-0.002
Mother under-weight	-0.961	-0.473	2.735
Mother over-weight	-1.431	-0.105	-0.998
Mother obese	-0.512	1.538	1.013

**Childcare**

Cared for by grandparents MCS2	1.351*	1.096	0.953
Cared for by other relatives MCS2	0.334	-1.01	-0.879
Cared for by friends/Neighbour MCS2	-0.233	5.077*	4.416*
Has Been to nursery school/class MCS2	0.243	-0.414	-2.104***
Has Been to playgroup MCS2	1.045	-1.454*	-2.168***
Has Been to pre-school MCS2	2.239**	-0.778	-2.113**
Has Been to childminder MCS2	1.094	1.763	1.087
Has Been to day nursery or creche MCS2	-0.621	1.624	0.382
Cared for by grandparents MCS3	n/a	0.491	-0.094
Cared for by relatives MCS3	n/a	0.55	0.917
Cared for by friends MCS3	n/a	-0.72	-0.635
Has Been to nursery school/class MCS3	n/a	-1.215	-0.219
Has Been to playgroup MCS3	n/a	1.274	2.051**
Has Been to pre-school MCS2	n/a	0.664	0.696
Has Been to childminder MCS3	n/a	1.869	2.733*
Has Been to day nursery or creche MCS3	n/a	-1.753	-0.316

**Home-Learning Environment**

2nd HLE quintile at MCS2	1.966**	1.144	-0.473
3rd HLE quintile at MCS2	3.239***	1.419	-0.687
4th HLE quintile at MCS2	4.045***	2.290**	-0.65
5th HLE quintile at MCS2	5.292***	3.847***	-0.515
Read to everyday at MCS2	7.487***	5.488***	0.812
Read to some days at MCS2	2.495*	1.605	-1.024
2nd HLE quintile at MCS3	n/a	0.241	0.731
3rd HLE quintile at MCS3	n/a	0.514	1.138
4th HLE quintile at MCS3	n/a	-1.213	-0.744
5th HLE quintile at MCS3	n/a	-0.054	1.307
Read to everyday at MCS3	n/a	-0.309	-0.186
Read to some days at MCS3	n/a	-1.212	-1.08
Mother rates herself as good parent – MCS 2/3	2.075***	1.366*	0.173
Mother rates herself as very good parent – MCS 2/3	0.68	-1.304*	-1.923***
Father rates himself as good parent – MCS 2/3	1.268*	-0.386	-0.999
Father rates himself as very good parent – MCS 2/3	1.240*	-1.083	-1.249*

**Parenting Style/Rules**

Lots of rules - MCS2	-0.875	-0.292	-0.024
Strictly enforced rules - MCS2	0.554	1.138**	0.643
Regular bed-times at MCS2	1.685**	1.326	0.251
Regular meal-times at MCS2	-0.277	0.914	0.307
Watches > 3 hours TV a day - MCS2	2.291***	0.394	-0.159
Watches > 3 hours TV a day - MCS3	1.145	0.146	-0.084
Plays computer > 1 hour a day - MCS3	n/a	-1.102	-1.140*
Regular bed-times at MCS3	n/a	2.699**	3.005***
Regular meal-times at MCS3	n/a	1.166	0.176
Eat breakfast together at MCS3	n/a	1.02	0.228
Observations	11054	11054	11054
R-Squared	0.23	0.21	0.36

\*\*\*, \*\* and \* indicate significance at the 1%, 5% and 10% levels, respectively.

**Endnotes**

<sup>i</sup> At the time of writing, the fourth sweep has only very recently become available.

<sup>ii</sup> Please note that the proportion of children in the “bottom 40%” is not exactly 40%. This is due to the fact that test scores are only semi-continuous.

<sup>iii</sup> This may be related to ethnic differences in birth weight. Dearden, Mesnard and Shaw 2006, show that ethnic minorities tend to have lower birth weights.

<sup>iv</sup> Without the help of any experimental variation in SEP, parental income, or indeed any mediating factors of interest in this project, we cannot rule out that there are unobservable factors (such as genetics, or typically unobserved attributes such as motivation of parent) that instead explain low incomes, poor achievement, and the potential transmission pathways we have highlighted. Structural Equation Models or Instrumental Variables Methods could be used to correct for this. However, the assumptions they rely on (e.g. exclusion restrictions and non-linearities) are not credible in this context and it is easy to show that results are very sensitive to the particular assumptions made.



# The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 11

Paul Gregg and Elizabeth Washbrook

CMPO, University of Bristol

[P.Gregg@bristol.ac.uk](mailto:P.Gregg@bristol.ac.uk)

(Received July 2010 Revised November 2010)

## Abstract

*This paper explores the correlates of the socio-economic gradient in children's educational performance through the primary school years. Thus it sits between the companion papers on pre-school cognitive outcomes and attainment in the secondary school years in this Special Issue. The poorest 20% of children score, on average 14 percentile points lower than the middle 20% in Key Stage 2 tests at age 11, and 31 percentile points lower than the richest 20%. Overall around one third of the attainment gaps by socio-economic background at age 11 are found to emerge after age 7. The evolution of attainment gaps over this period is found to be related a range of attitudes to education and behavioural patterns of the study children. Low maternal aspirations for the child's final educational attainment are strongly linked to the widening socio-economic gap during these years, over and above their influence on the child's own measured attitudes and behaviours.*

JEL codes: I21

**Key words:** education, inter-generational transmission, socio-economic gap

## 1. Introduction

Children growing up in less affluent families emerge from our schools with substantially lower levels of educational attainment. These educational deficits emerge early in children's lives, even before entry into school, and widen throughout childhood (see Feinstein 2003 and 2004). This paper focuses on the differences between socio-economic groups in academic performance at the time of entry into secondary education at age 11. We use data from the Avon Longitudinal Study of Parents and Children (ALSPAC), a cohort of around 14,000 children born in Avon in the early 1990s.<sup>1</sup> In ALSPAC, the 20% of children with the lowest socio-economic position score, on

average, 14 percentile points lower than the middle 20% in Key Stage 2 tests at age 11, and 31 percentile points lower than the richest 20%. Children from low socio-economic backgrounds perform less well on tests at age 7. Moreover, poor children who perform well at 7 are more likely than rich children to have slipped back by age 11, and poor children who perform badly at 7 are less likely to catch up over the period. Around one third of the attainment gaps by socio-economic background at age 11 emerge after age 7.

We aim to examine some of the routes through which family socio-economic position (SEP) affects educational attainment. Our focus is on a range of early parenting behaviours and on parent and child attitudes, behaviours and

beliefs in the primary years, that have seldom been studied in work on the causes and consequences of poverty. We aim to explore which attitudes and beliefs are important in influencing attainment at age 11 and patterns of educational development between the ages of 7 and 11, and the relative strength of these influences. Parental and child attitudes and beliefs are likely to be correlated with a range of other family background factors and hence it is difficult to disentangle the independent influence of the mechanisms of interest. We seek to minimise this problem in three ways. First, we adopt a distal/proximal modelling approach in which a range of family demographic measures ‘compete’ with our posited transmission mechanisms, to explain educational patterning by socio-economic background. Measures such as parental education, age and number of siblings are included to ‘mop up’ any correlated but unobserved influences on educational attainment. School characteristics are also included as proxies in this way. Second, we explore the extent to which attitudes and behaviours are associated with educational development over the four years prior to the start of secondary schooling, by including measures of attainment at age 7. The inclusion of prior attainment, focuses the estimates on development during the specified window, and gives an insight into how much of the influence of attitudes and behaviours are already crystallized in attainment at the start of junior school. A final approach is to include earlier measures of parenting and the home learning environment which influence attainment at age 5 (see Dearden et al in this Issue), and again help to isolate the role of transmission mechanisms during the primary school years. If early (pre-school) environments are the major determinants of educational trajectories, then omitting them from our analysis may falsely overstate the role of the attitudes and behaviours in middle childhood that are our focus. Despite these strategies, the danger that our estimates pick up the correlation of the

specified transmission mechanisms with unobserved factors remains. However, it should not be assumed that our estimates necessarily overstate the potential contribution of attitudes and behaviours to the socio-economic gap. Any mis-measurement of the complex psychological constructs in question will have the reverse effect, and tend to bias the estimated associations downward.

## 2. Modelling approach

The aim of our analysis is to better understand the observed relationship between a child’s socio-economic background and his or her educational performance at 11. In particular, the aim is to assess the importance of attitudes and aspirations, both of the parents and the child, on attainment. We explore the role of a diverse range of factors that potentially mediate this relationship, in the sense that they are correlated with family background and directly influence children’s educational development. The ALSPAC data is extremely rich, so we organise our data according to distinct concepts that have been identified in the literature. The broad distal-proximal modelling framework is laid out in the overview paper in this Special Issue by Goodman et al (2010), section 4 and in particular in Figure 3. Socio-economic position (SEP) is the principal family background indicator of interest. However, conceptually it is only one of a set of ‘distal’ factors, or aspects of family background, that together characterise the social and cultural resources available to the child. SEP, along with parental education, family structure and size, school quality and other distal variables, are assumed to shape the ‘proximal’ attitudes, behaviours and beliefs that are the focus of this study. We hypothesise that the values and behaviours of parents (both in the pre-school period and during the primary school years) and of the children themselves, are relatively more immediate or direct influences on educational achievement, and help to explain why we

observe differences in average Key Stage scores between different distal groups.

Our choice of potential mediating processes is informed by data availability and by a diverse social science literature on the determinants of educational success. Many of the concepts used in this study are laid out in the overview paper (Goodman et al 2010). In this study, we draw in particular on concepts informed by the expectancy-value theory of achievement motivation, proposed by Eccles, Wigfield and colleagues (e.g. Wigfield and Eccles 2000). Underlying this theory, is the idea that individuals' achievement-related choices, persistence and performance in an activity can be explained by their beliefs about how well they will do the activity, and the extent to which they value the activity. However, we do not restrict our attention solely to motivational constructs, but also consider the role of factors such as children's behavioural problems, or self-regulation, which have been linked to educational performance in a largely unrelated literature (see for example, McLeod and Kaiser 2004; Currie and Stabile 2006). A child's degree of self-regulation influences the ability to see intentions through, and thus interacts with achievement motivation in influencing the expected outcome. The relationship between the parents' and child's motivations and expectations is also likely to be complex. The parent will be aware, though potentially imperfectly, of the child's capabilities, and will adjust expectations accordingly. However, the parents' own experiences, and those of others they are aware of, may also shape their own expectations of the child and their efforts to influence the child's motivation and decision-making.

In our analysis, we present results on the relative importance of each transmission mechanism, with and without conditioning on Key Stage 1. The first set of results (the levels model) reveals how far each factor is associated with Key Stage 2 outcomes, without distinguishing how far its influence is manifested pre- or post-age 7. The second set

of results (the value-added model) holds child attainment at 7 constant, and hence reveals how far each factor is associated with improvement or deterioration of the child's performance during the primary school years. In effect, the value-added model measures the influence of covariates on relative progress between the ages of 7 and 11. It addresses the question of whether and where gaps would open up during the period even if all children started with the same level of academic ability. The formal representation of the estimation structure and the results decompositions are laid out in the overview paper, to reduce repetition within each paper in the series.

### 3. Data and measures

ALSPAC is a cohort study that recruited around 14,000 pregnant women who were resident in the Avon area of England whose expected date of delivery fell between 1<sup>st</sup> April 1991 and 31<sup>st</sup> December 1992. Study families were surveyed via high frequency postal questionnaires from the time of pregnancy onwards, and via a number of hands-on clinics in which ALSPAC staff administered a range of detailed physical, psychometric and psychological tests to the children. ALSPAC has been linked to the National Pupil Database (NPD), which contains school identifiers and results on national Key Stage school tests for all children in the public school system, and information of local deprivation at the small area level (the government-produced Indices of Multiple Deprivation, IMD). For details of the comparability of ALSPAC with national data and attrition and sample selection issues for data used in the study please refer to Appendix 1 in the Supplementary Files. <sup>ii</sup>

The key outcome variable in our analysis is derived from the results of Key Stage 2 assessments in English, maths and science, taken by all pupils in state schools in Year 6 (age 11). We construct an average measure of performance on the three tests, and express this total as a percentile score. We explore two specifications of the Key Stage 2 outcome in all

the main results in the paper: one without, and one with, a control for Key Stage 1 performance at age 7. Again, the Key Stage 1 measure is constructed as an average score, here over reading, writing and maths tests.

Our measure of socio-economic position is derived from data on a number of indicators: household income at age 2 to 3, age 7 and age 11 (3 variables), mother's and father's social class (2 variables), housing tenure at ages 2 to 5 (2 variables) and average reported experience of financial difficulties at ages 0 to 7 (1 variable). We extract the first principal component using the method described in Goodman et al (2010), a component that accounts for 48% of the variation in the component variables, and individuals are then placed into quintiles (fifths) of the population ranked by this measure. The aim is construct a long-term measure of the material resources of the household, one that incorporates the fact that deprivation is multi-dimensional and that socio-economic risk factors are likely to be cumulative.

Many of our measures of potential transmission mechanisms are taken from a mother-reported postal questionnaire when the child was 9 years old, and from a hands-on clinic when the child was 8. The timing on these measures is advantageous, because they occur between the Key Stage 1 and Key Stage 2 tests. This enables us to explore how different family environments affect the trajectory of a child's development, given their prior attainment. However, the data requirements are such that families must have remained in the study from birth to 9 years. This is quite a stringent requirement and there is substantial attrition, leaving us with a working sample of only 7972, about half of original cohort remaining in English state schools. We have used a number of techniques to ensure that our definitions of SEP groupings, and the scaling of the Key Stage outcome variables, are as representative as possible of the national population, rather than only those who remain in the sample (see the appendix A2). Missing values on the

explanatory variables are dealt with by mean replacement and the inclusion of a missing dummy.

The ALSPAC data contain information on a wide range of factors that may help to explain the poorer educational performance of socially-disadvantaged children. We distinguish two types of explanatory factors. First are 'distal' factors that describe the resources available to parents and children in a broad sense, and capture the structural features of the environments in which children are raised.

*Parental education* consists of variables that capture the mother's and father's highest educational qualification, measured just prior to the child's birth (Certificate of Secondary Education/no qualifications; Vocational/O-level; A-level; Degree).

*Demographic characteristics* consist of variables measuring child's gender; ethnicity (dummy for non-white); month of birth (scaled to September = 0); family structure at age 7 (resident biological father, step-father, or single parent); mother's age at birth (dummies for <20; 20-24; 25-29; 30-34; 35 or more); number of older siblings (none, 1, 2, 3 or more); number of younger siblings at age 9 (none, 1, 2 or more); mother's and father's employment status at age 4; an indicator for English as a second language; and mother's and father's self-rated health (on a scale of 1 to 4) at age 4.

*School characteristics* are the mean Key Stage 1 score of all pupils at the child's school; the mean value-added between Key Stages 1 and 2 of all pupils at the child's school (both standardized to unit variance); and the proportion of pupils in the school eligible for free school meals (FSM).

The next groups of variables are the attitudes, behaviours and beliefs of parents and children or the 'proximal' factors, that we expect to be shaped by the distal features of the family's circumstances, and that in turn directly affect the learning process.

*Pre-school environments* consists of variables relating to health, cognitive stimulation and childcare experiences, all



measured prior to age 5. These bear a reasonably close relationship with those used in the companion early years paper in this issue by Deaden et al (2010). Our analysis allows us to explore whether these early influences have an association with faster progress between 7 and 11, perhaps because they capture differences between parents that persist. Included variables are: birth weight (in kg); an indicator of gestation less than 37 weeks; indicators of breastfeeding duration (never, less than 3 months, 3 to 6 months, or more than 6 months); an indicator for whether the mother smoked in pregnancy; an indicator for post-natal depression (average Edinburgh Post-natal Depression Scale over six dates, from pregnancy to 33 months); quintile groups of a Home Learning Environment (HLE) index (average of 9 standardized items measured at age 3: frequency that child is taken to the library, frequency that mother reads and sings to the child; indicators for whether the mother tries to teach the child colours, the alphabet, numbers, nursery rhymes, songs, and shapes and sizes); indicators for whether the child is read to every day and has a regular sleeping routine at 3; indicators for whether the child regularly attended a day nursery or crèche before age 3, and attendance at a nursery at age 3 or 4.

*Other parental attitudes and behaviours* consider potential ways in which parents may influence the educational development of children during the primary school years. These include a measure of the mother's sense of control in life (the 12-item Adult Nowicki-Strickland Locus of Control scale, standardized to unit variance); the mother's rating of whether school was a valuable experience for her (5-point scale, standardized to unit variance); the mother's aspirations for her lifting, drinking alcohol or carrying a weapon, each standardized and averaged).

Other contexts besides the home may also impact on ability and motivation. Here we examine experience of bullying (frequency of occurrence of nine events in the last 6 months,

child's education (get good GCSEs then leave, take at least one A-level; go to university, other); the amount and quality of mother-child educational interactions likely to contain some component of cognitive stimulation (eight 5-point items such as, helps with homework and draws or paints with child, standardized and averaged); and the amount and quality of other mother-child non-educational interactions that may nevertheless be important in fostering family bonds and socio-emotional well-being in children (nine 5-point items such as takes child to the park and does active play with child, standardized and averaged).

*Child's attitudes and behaviours* are captured by three broad groupings of variables. Beliefs and values include the child's perceptions of their own ability (the scholastic competence sub-scale of Harter's Self-Perception Profile for Children, standardized to unit variance); their feelings about, or intrinsic valuation, of school (the average of 15 standardized items such as how much the child enjoys school, seems bored by school and enjoys different lessons); their sense of what is important in life, or extrinsic values (indicators of whether the child believes school results, hobbies and interests, and material possessions to be important in life); and their sense of how personal effort in general will impact on their lives (the 12-item Nowicki-Strickland Internal-External Scale, standardized to unit variance).

Behavioural difficulties, which may interfere with academic progress, are measured in terms of symptoms of hyperactivity, conduct problems and emotional problems (three sub-scales of the parent-report Strengths and Difficulties Questionnaire, each standardized to unit variance); and engagement in anti-social activities (11 binary indicators for whether the child has ever engaged in behaviours like shop-such as having belongings taken, being threatened and being called names, standardized and averaged); experience of other peer problems (the Strengths and Difficulties Questionnaire peer problems sub-scale); participation in out-of-school activities,

such as sports, singing or drama groups (eight items capturing the frequency of participation, standardized and averaged); other pro-social behaviours (the Strengths and Difficulties pro-social sub-scale); and teacher-child relations (six items such as whether the child is frightened of the teacher and whether he/she believes the teacher thinks his/her work is good, standardized and averaged).

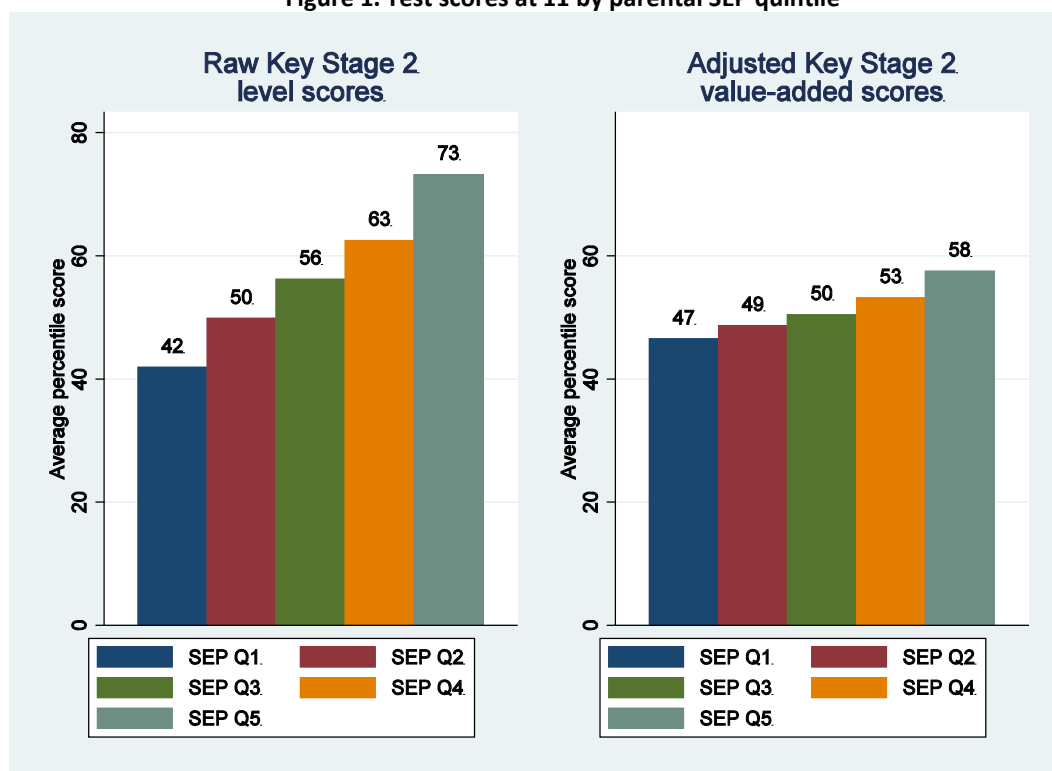
## 4. Results

### 4.1 Socio-economic differences in child and family characteristics

The left hand panel of Figure 1 shows the average test scores of children in our working sample at age 11 (KS2), by quintile of SEP. The

scores reflect the child's percentile in the Avon distribution and range from 1, for the lowest performing children, to 100 for the highest performing. If there were no systematic differences in attainment by SEP, each group would have an average score of 50.5. Deviations from this number show how far children from different socio-economic backgrounds tend to over- or under-perform relative to the average. It is clear that there are substantial differences in educational performance that run throughout the socio-economic scale. The mean score of the most advantaged children is 31 percentile points higher than that of the most disadvantaged.

Figure 1. Test scores at 11 by parental SEP quintile



**Note:** the right hand panel presents an 'adjusted' gap, showing the average percentile score by SEP quintile, assuming all children scored the same at age 7. Such estimates are derived by predicting each individual's Key Stage 2 percentile in the situation where all pupils scored equally (i.e. at percentile 50.5) at Key Stage 1, based on a 'value-added' regression of the following form:  $KS_{11} = \alpha + \lambda SEP_i + \theta KS_{17} + \varepsilon_{11}$ .

Although there are some differences around the middle of the SEP distribution, it is noticeable that the largest gaps occur at the tails, with the poorest children falling far behind, and the richest children pulling strongly

ahead. The right hand panel of Figure 1 explores how these patterns are affected by controlling for prior attainment at Key Stage 1 (KS1; age 7). The gradients here are noticeably smaller than in the left panel, and show that a

large fraction of the educational inequality observed at age 11 in fact reflects differences that are already apparent by age 7. However substantial differences remain, and suggest that the poorest children fall a further 11 percentile points (almost one-third of the raw 31 point gap) behind the richest children between 7 and 11, even if they started the period with the same levels of attainment.

Table 1 explores the trajectories of children from different socio-economic groups in more detail. The first two rows focus on low-achieving children at age 7. The lower SEP children in our sample are much more likely to fall into this group – 54% of the bottom quintile scored at the 40<sup>th</sup> percentile or below, compared with only 16% of the top quintile. They are also much less likely to escape from the low-achieving group by age 11. Less than a quarter of the poorest children who scored in the bottom 40% at age 7, are able to escape by age 11, whereas more than half of those in the top SEP quintile are no longer there four years

later. The next two rows show a parallel pattern for the high-achieving children at 7. Low SEP children are much less likely to be scoring above the 60<sup>th</sup> percentile at this age (25% of the poorest compared with 64% of the richest), and those who do are more likely to have dropped out of this high-achieving group by age 11 (28% of the poorest fall back compared with only 8% of the richest high-achievers). It is the combination of these two factors – that low SEP children start behind at age 7 and that high achieving children from poorer families do not progress as well as higher SEP children and are often overtaken by less able more affluent children during the primary school years – that results in the patterns shown in the final two rows of the table. At 11, children in the bottom SEP quintile are nearly five times as likely to be low-achievers, and two-and-a-half times less likely to be high-achievers, as children in the top SEP quintile.

**Table 1. Educational performance at 7 and 11, by parental SEP**

	SEP	SEP	SEP	SEP	SEP
Proportion of children:	Q1	Q2	Q3	Q4	Q5
Bottom 40% at KS1	0.54	0.41	0.33	0.27	0.16
Escape from bottom 40% by KS2 <sup>a</sup>	0.23	0.29	0.32	0.37	0.54
Top 40% at KS1	0.25	0.34	0.42	0.50	0.64
Dropout of top 40% by KS2 <sup>a</sup>	0.28	0.26	0.18	0.14	0.08
Bottom 40% at KS2	0.51	0.38	0.31	0.24	0.11
Top 40% at KS2	0.28	0.37	0.47	0.57	0.73

<sup>a</sup> Proportion exiting group, conditional on being in group at KS1.

Top and bottom 40% refer to the proportions in the population of all children in the Avon area. Proportions in the working sample do not equal exactly 0.40 because the working sample is positively selected, and because Key Stage 1 scores are only semi-continuous. 7972 observations.

The characteristics of low income families differ from those of their more affluent counterparts along many dimensions. Table 2 documents the socio-economic gaps in each of our potential transmission mechanisms. The first panel focuses a number of dimensions of the pre-school environment that plausibly affect children's health and development. Breast feeding, maternal smoking and post-natal depression are all very strongly graded by SEP, and there are also some differences in average birth weight and the likelihood of being born pre-term. Less advantaged parents tend to engage in fewer reading and teaching behaviours with their three-year-olds, although this is not universally true: 13% of the poorest families are in the highest home learning environment quintile, while 14% of the most

affluent families are in the bottom quintile. Finally, we see that exposure to centre-based child care before the age of 3 was relatively rare in this cohort, but much more common among the better off, while attendance at nursery at ages 3 to 4 was most common among the best and worst off, with middle SEP children showing the lowest participation rates. We might imagine that the quality of child care settings would differ with family income, for example because the most advantaged can afford expensive private nurseries<sup>iii</sup>. Research suggests that quality is the key factor that determines the influence of childcare on children's development (NICHD ECCRN and Duncan 2003), but unfortunately we are not able to measure it here, and can only explore differences in exposure to particular types.

**Table 2. Means of mechanism variables, by quintile of socio-economic position**

	Poorest SEP quintile (Q1)	2 <sup>nd</sup> SEP quintile (Q2)	Middle SEP quintile (Q3)	4 <sup>th</sup> SEP quintile (Q4)	Richest SEP quintile (Q5)	Q5-Q1
<b>Parental attitudes and behaviours</b>						
<b>A. Pre-school environments</b>						
Birth weight (kg)	3.34	3.40	3.44	3.44	3.46	0.11 kg**
Gestation < 37 weeks	6.4%	6.4%	5.4%	5.6%	4.4%	-1.9 ppt
Breast fed: Never	46.3%	34.1%	27.3%	21.1%	10.6%	-35.7 ppt**
Breast fed: < 3 mths	24.9%	25.9%	26.4%	25.9%	18.0%	-6.9 ppt**
Breast fed: 3-6 mths	11.3%	14.5%	18.0%	17.6%	18.7%	7.4 ppt**
Breast fed: 6 mths +	17.5%	25.5%	28.3%	35.4%	52.7%	35.2 ppt**
Mother smoked in pregnancy	56.5%	38.3%	23.2%	18.3%	10.7%	-45.9 ppt**
Mother had post-natal depression	27.3%	18.6%	12.0%	10.1%	8.1%	-19.2 ppt**
HLE at 3: Lowest quintile	30.6%	24.0%	19.6%	15.7%	13.8%	-16.9 ppt**
HLE at 3: Second quintile	23.3%	22.9%	22.0%	21.0%	17.3%	-6.0 ppt**
HLE at 3: Middle quintile	18.6%	20.6%	19.6%	21.1%	19.6%	1.0 ppt
HLE at 3: Fourth quintile	14.0%	15.1%	19.0%	21.6%	23.4%	9.4 ppt**
HLE at 3: Highest quintile	13.4%	17.5%	19.8%	20.6%	25.9%	12.5 ppt**
Child read to daily at 3	52.3%	59.7%	63.7%	66.8%	71.4%	19.1 ppt**
Child has regular sleep routine at 3	85.2%	89.7%	91.9%	94.7%	96.5%	11.3 ppt**
Centre-based child care pre-age 3	7.2%	9.3%	9.8%	15.7%	24.2%	16.9 ppt**
Nursery age 3 to 4	40.4%	38.1%	35.7%	44.0%	53.6%	13.2 ppt**

(Table 2 cont'd)

**B. Other attitudes and behaviours**

Mother's locus of control (scale)	-0.64	0.28	0.01	0.27	0.61	1.26 sd**
Mother found school valuable (scale)	-0.39	0.21	-0.01	0.1	0.35	0.74 sd**
Mother hopes child will get good GCSEs	21.9%	16.2%	11.3%	6.4%	1.4%	-20.5 ppt**
Mother hopes child will get at least 1 A-level	19.0%	20.4%	19.0%	14.2%	6.6%	-12.4 ppt**
Mother hopes child will go to university	36.5%	41.4%	50.2%	62.4%	80.5%	44.0 ppt**
Mother hopes other for child	22.7%	22.1%	19.5%	17.0%	11.5%	-11.2 ppt**
Mother-child interactions: Education (scale)	0.03	0.01	0.02	0.00	-0.03	-0.06 sd**
Mother-child interactions: Non-educational (scale)	-0.03	0.01	0.02	0.01	-0.01	0.02 sd
<b>Child's attitudes and behaviours</b>						
Ability beliefs (scale)	-0.12	0.11	-0.06	0.07	0.11	0.24 sd**
Locus of control (scale)	-0.31	0.16	-0.1	0.07	0.28	0.59 sd**
Enjoyment of school (intrinsic values, scale)	-0.03	0.03	0	0.04	0.02	0.06 sd
School results important in life (extrinsic values)	50.6%	55.8%	61.3%	62.9%	66.6%	16.0 ppt**
Hobbies important in life (extrinsic values)	66.6%	70.8%	75.2%	77.3%	83.5%	16.9 ppt**
Possessions important in life (extrinsic values)	75.0%	76.6%	76.7%	76.5%	78.5%	3.6 ppt*
Anti-social behaviours (scale)	0.1	0.03	-0.01	-0.02	-0.03	-0.13 sd**
Hyperactivity (scale)	0.27	0.15	-0.02	-0.08	-0.16	-0.42 sd**
Emotional symptoms (scale)	0.17	0.06	-0.02	-0.03	-0.1	-0.27 sd**
Conduct problems (scale)	0.33	0.11	-0.04	-0.08	-0.14	-0.47 sd**
Experience of bullying (scale)	0.08	0	0	0	-0.06	-0.14 sd**
Pro-social behaviours (scale)	-0.04	0	0.04	0.04	-0.05	-0.01 sd
Peer problems (scale)	0.32	0.08	-0.04	-0.07	-0.12	-0.44 sd**
Leisure/out-of-school activities (scale)	-0.23	0.16	-0.07	0.04	0.22	0.45 sd**
Teacher-child relations (scale)	-0.04	0.01	0.02	0.02	0.01	0.05 sd*

Full sample contains 7972 observations. Means defined over non-missing responses only. Q1 denotes the lowest SEP quintile, Q2 the second lowest SEP quintile, and so on; HLE denotes the home learning environment index; ppt denotes percentile points; scale denotes the average of a number of standardized item scores; std score denotes a variable standardized to mean 0, standard deviation 1 on the maximum available sample; sd denotes standard deviations. \*\* and \* indicate significance at the 1 and 5% levels respectively.

The second panel of Table 2 highlights differences in other parental attitudes and behaviours that are not specific to the pre-school period. Low SEP mothers tend to have a much more external locus of control (a sense that luck or fate, rather than their own actions, are what

matters in life), and tend to view their own schooling experiences as having been less valuable than more advantaged mothers. Most strikingly, there are very large differences in their educational aspirations for their children when they are age 9. 81% of mothers in the richest

quintile hope their child will go to university, compared with only 37% of mothers in the lowest quintile. Interestingly, these differences do not appear to be mirrored in differences in the frequency and variety of mother-child interactions at 9. We find only very small differences in mothers' reports of how often they make things or read with the child, help with homework, etc (educational interactions), and in how often they take the child to the park or shopping, prepare food with the child, etc (non-educational interactions). Hence the marked socio-economic differences we see in pre-school parenting behaviours seem to have narrowed by the mid-primary school years.

The final panel in Table 2 shows how children's own attitudes, behaviours and beliefs differ with socio-economic position. There are notable differences in the degree to which low-income children express negative views at the ages of 8 to 9. They tend to regard themselves as scholastically less able, they are less likely to view school results as important in life (as reported by the mother) and tend to have a more external locus of control. This sense that luck or chance determines outcomes, rather than one's own efforts, is also found among low income mothers, but the gradient is much less marked in the children than in their mothers. In contrast, we find little differences in the extent to which children enjoy school or value it for its own sake, and in the likelihood of believing material possessions to be important in life. Low income children are much more likely to exhibit behavioural problems in terms of hyperactivity, conduct problems and peer relations, including being a victim of bullying, whereas pro-social (cooperative) behaviours and teacher-child relations differ less with family background. Finally there is evidence of marked differences in participation in out-of-school leisure activities such as sports, clubs and classes.

#### **4.2 Explaining the socio-economic gaps in educational outcomes at 11**

So far we have shown that parental and child aspirations, attitudes and behaviours

differ across socio-economic position of the family. These analyses, however, do not show how important attitudes and behaviours are to attainment overall, nor do they identify which aspects of our range of measures of parent and child behaviours and beliefs contribute most to the transmission of socio-economic position. In this section, we explore which factors do the heavy lifting in explaining SEP attainment gradients in children, and the relative importance of specific attitudes, behaviours and beliefs. The explanatory power of a given variable depends on two associations: the extent to which it varies by socio-economic group, and the extent to which it independently predicts educational outcomes (see the overview paper by Goodman et al. 2010). Factors that vary little with socio-economic position – such as parent-child interactions at age 9 – cannot play an important role in generating the social gradient in outcomes. This does not imply, however, that such factors are not consequential for educational achievement in general. Equally, a factor may be strongly socially graded, but if it has little association with outcomes, its role in explaining the gradient will be small. In the search for the key attitudes and behaviours that drive the observed SEP gaps, we need to identify factors that are both concentrated in disadvantaged families, and that strongly interfere with the development of children's learning. Table 2 gave the levels of patterning of variables by SEP, so next we focus on the relationship between behaviours and educational attainment in a fully conditioned model. This provides estimates of the direct effects of each of the individual variables, over and above other aspects of family background.

We explore this in two models (levels and value-added), reporting the marginal association of each individual measure of attitudes and behaviours with Key Stage 2, given all other aspects of family background. This analysis allows us to identify factors that are not, or only weakly, socially graded but that may nevertheless be consequential for

educational achievement. Table 3 shows selected coefficients with the focus on attitudes and behaviours rather than family background characteristics, from two fully-controlled regression models of Key Stage 2 (i.e. conditioning on all the variables listed in section 3), with the sole difference between the two being the inclusion of Key Stage 1 control in column 2. Looking first at the influence of pre-school environments, we find significant associations of the expected sign between Key Stage 2 and birth weight and gestation, breastfeeding, and the home learning environment at 3, which is notable given the rich set of other controls included in the model. Birth weight and the pre-school learning environment become insignificant when Key Stage 1 is added to the model, implying that their influence on academic ability is fully apparent by age 7. Two 'perverse' results are the positive association of pre-natal smoking, and the negative association of reading to the child daily at age 3 with Key Stage 2. Both of these factors appear uncorrelated with Key Stage 1, as the associations remain strong when prior ability is controlled. We cannot say conclusively why these patterns arise, but it perhaps an illustration of the dangers of an 'over-controlled model' discussed in the overview chapter, where strong co-linearities between the covariates result in the identification of individual marginal effects from rare and unrepresentative observations. Here, early reading is strongly related to the home learning environment and suggests that this is the key factor. The home learning environment is highly significant and materially important to attainment at age 11, but this is entirely captured by attainment at age 7 and does not contribute to the value-added models. This is strong evidence for this being an age-specific factor, rather than a marker for time-invariant, unobserved, positive parental characteristics.

Looking next at maternal attitudes and behaviours during primary school, we see that maternal locus of control is a significant

predictor of age 11 outcomes, because it is associated both with higher ability at 7 and faster progress between 7 and 11. The magnitude of the association between maternal aspirations and child outcomes is, however, far more dramatic. Holding all else constant, children of mothers who hope they will go to university, score 13.7 percentile points higher at Key Stage 2 than children of mothers who want them to get good GCSEs then leave. This can be contrasted the gap of 4.5 between the lowest and the highest SEP quintiles estimated in the same regression. More than half of the effect of maternal aspirations is absorbed when we control for Key Stage 1, but maternal aspirations remain one of the biggest single predictors of progress between 7 and 11. It is not the case, then, that maternal aspirations simply reflect the child's revealed level of ability from school tests two years previously, although they are related. These findings are particularly strong given that they are direct effects, that is, net of any indirect effects on outcomes via children's own attitudes and behaviours.

Virtually all of our variables capturing children's attitudes, behaviours and beliefs have independent, statistically significant, if modest associations with Key Stage 2 outcomes. Children's beliefs in their own scholastic ability and their locus of control, are associated partly with faster progress between 7 and 11, but also reflect higher ability as measured at age 7. This may be because educational success promotes positive attitudes, but it may also be that positive attitudes begin earlier in life, and that this in part determines attainment at 7. A positive intrinsic valuation or enjoyment of school is also associated with higher scores at 11, but this effect is entirely explained by the higher prior achievement of children with positive valuations of school. Extrinsic values, however, or beliefs that something is important in life, continue to influence learning after the age of 7. Children who believe that school results or hobbies and interests are important in life,

score better at both ages, while those that believe material possessions to be important tend to fall behind. We find further evidence that hobbies and interests are associated with educational achievement, as the frequency of participation in leisure activities (such as sports, singing/drama lessons and groups such as Scouts) is a significant predictor of the outcome, although again only because it is positively correlated with prior achievement.

Children's social and behavioural skills are indeed key factors explaining educational success. Hyperactivity or attention problems are particularly associated with adverse outcomes, as are conduct problems and anti-social behaviours (fighting, stealing, substance abuse, etc). In all these cases controlling for prior attainment does not eliminate the effect, which would appear to be evidence against a hypothesis that they are simply correlated with low cognitive ability. Instead our results provide support for the idea that behavioural difficulties interfere with the learning process over the course of primary school and lead to under-performance at age 11. Interestingly, we find no association between emotional

symptoms (which relate more to depression, anxiety and internalising behaviours) and performance at either 7 or 11. And pro-social and peer relations have the opposite effects than expected. Children who exhibit strongly pro-social behaviours (e.g. volunteering to help others, sharing readily with other children) and also those who experience fewer problems with peers (e.g. tending to be liked by other children and spending time with them) score worse at Key Stage 1 and progress more slowly thereafter than more self-contained solitary children. These results are interesting because they highlight that different types of social and behavioural problems differ strongly in their association with educational performance.

Overall it is worth noting that the fit of these models is high, especially when age 7 attainment is included. In the model for age 11 attainment, just under half the total variance is explained, and including lagged age 7 attainment raises this to 70%. This suggests that the models being described do capture, to a large degree, the variance in child attainment, even if it is inevitable that some factors will be poorly measured.

**Table 3. Estimated effects of attitudes, behaviours and beliefs on Key Stage 2 scores**

Variable	Regression coefficient	
	(1)	(2)
<b>Pre-school environments</b>		
Birth weight (kg)	2.0***	0.5
Gestation < 37 weeks	2.0*	1.5*
Breast fed: Never		Omitted
Breast fed: < 3 mths	0.1	0.1
Breast fed: 3-6 mths	0.8	0.9
Breast fed: 6 mths +	1.9***	1.9***
Mother smoked in pregnancy	1.3**	0.9*
Mother had post-natal depression	0.5	0.5
HLE at 3: Lowest quintile		Omitted
HLE at 3: Second quintile	1.3*	-0.2
HLE at 3: Middle quintile	4.1***	0.7
HLE at 3: Fourth quintile	4.6***	0.2
HLE at 3: Highest quintile	5.2***	0.4
Child read to daily at 3	-2.2***	-1.4***
Child has regular sleeping routine at 3	0.8	0



Variable	Regression coefficient	
	(1)	(2)
Centre-based child care pre-age 3	-0.3	-0.3
Nursery age 3 to 4	-0.3	-0.5
<b>Other parental attitudes and behaviours</b>		
Mother's locus of control (scale)	1.1***	0.5**
Mother found school valuable (scale)	0.2	0.2
Mother hopes child will get good GCSEs		Omitted
Mother hopes child will get at least 1 A-level	8.3***	3.3***
Mother hopes child will go to university	13.7***	5.7***
Mother hopes other for child	6.7***	3.0***
Mother-child interactions: Education (scale)	-4.7***	-0.8**
Mother-child interactions: Non-educational (scale)	-1.4**	-1.0**
<b>Child's attitudes and behaviours</b>		
Ability beliefs (scale)	3.4***	1.3***
Locus of control (scale)	4.0***	2.2***
Enjoyment of school (intrinsic values, scale)	2.0***	0.3
School results important in life (extrinsic values)	6.1***	2.0***
Hobbies/interests important in life	1.9***	1.3***
Material possessions important in life	-3.1***	-2.1***
Anti-social behaviours (scale)	-3.6***	-1.5***
Hyperactivity (scale)	-4.5***	-1.5***
Emotional symptoms (scale)	0	-0.1
Conduct problems (scale)	-1.3***	-0.9***
Experience of bullying (scale)	-1.4**	-1.1***
Pro-social behaviours (scale)	-1.8***	-0.8***
Peer problems (scale)	0.7**	0.5**
Participation in leisure/out-of-school activities	2.3***	0.7
Teacher-child relations (scale)	-0.1	-0.4
Observations	7972	7972
Adjusted R-squared	0.439	0.703

**Notes:** table contains selected coefficients from two OLS regression models of the determinants of Key Stage 2 (age 11) percentile score. Missing dummies are included, where appropriate, but their coefficient estimates are not shown. Both models also include controls for SEP quintile, demographic and family background characteristics, and primary school characteristics. The second model (shown on the right) additionally controls for attainment at Key Stage 1 (age 7). All other coefficient estimates can be found in Appendix A3. \*\*\*, \*\*, and \* indicate significance at the 1, 5 and 10% levels respectively.

Whilst Table 2 showed how socially graded each factor is in the population (i.e. the strength of its association with SEP), Table 3 tells us about the strength of the association between individual

mediating variables and children's attainment at age 11. Next we can combine the two pieces of information to provide a summary decomposition of the relative importance of each factor, in

accounting for the attainment gaps as described in the overview paper. Table 4 shows the results of this exercise. We focus our results on the difference in mean test scores, between the poorest 20% of children and those in the middle 20% of the SEP distribution (the middle-bottom gap), which is a summary measure of differences between those in poverty and average children. In contrast, the difference between the poorest and the richest 20% (the top-middle gap) is an indicator of the full extent of educational inequality at 11. The left panel of Table 4 presents results from the levels model, and the right panel presents results from the value-added model. For each set of results, we show the contribution of a particular factor to the raw middle-bottom and top-bottom gaps, first in terms of percentile point Key Stage 2 scores, and then as a percentage of the total raw gap.

The top panel gives an overall summary of the decompositions. Of the 14.3 percentile point middle-bottom SEP gap in level scores, we can

explain 10.9 (76%) using all our measured distal and proximal factors, leaving 3.4 percentile points (24%) – unexplained. And of the 31.3 percentile point top-bottom gap, 26.8 points (86%) are explained and 4.5 points (14%) unexplained. The right panel shows that the conditioning on Key Stage 1 increases the proportion of the middle-bottom gap that is explained to 87%. To a large extent this displaces the predictive power of child and family characteristics – 66% is attributed to prior ability, leaving 21% to other measured factors associated with the widening socio-economic gap. We see a similar pattern for the top-bottom gap, with 59% explained by prior ability differences, 34% attributed to other measured factors and just 7% unaccounted for. Hence in the value-added models, the factors considered explain the widening attainment gap, between the poorest and most affluent, somewhat more fully than they do the gap between the poorest and middle income children.

**Table 4. Breakdown of the bottom-middle and bottom-top SEP gaps in average Key Stage 2 scores**

Factor	Difference from Q1 attributed to factor							
	(1) Levels model				(2) Value-added model			
	Percentile		As % of total		Percentile		As % of total	
	Q3	Q5	Q3	Q5	Q3	Q5	Q3	Q5
<b>All</b>	<b>14.33</b>	<b>31.33</b>	<b>100%</b>	<b>100%</b>	<b>14.33</b>	<b>31.33</b>	<b>100%</b>	<b>100%</b>
<b>All measured factors (sum I to VIII)</b>	<b>10.9</b>	<b>26.82</b>	<b>76.1%</b>	<b>85.6%</b>	<b>12.53</b>	<b>29.13</b>	<b>87.4%</b>	<b>93.0%</b>
<b>Residual unexplained component</b>	<b>3.43</b>	<b>4.51</b>	<b>23.9%</b>	<b>14.4%</b>	<b>1.80</b>	<b>2.20</b>	<b>12.6%</b>	<b>7.0%</b>
<b>I. Key Stage 1</b>	-	-	-	-	<b>9.44</b>	<b>18.45</b>	<b>65.9%</b>	<b>58.9%</b>
<b>II. Parental education</b>	<b>2.98</b>	<b>7.83</b>	<b>20.8%</b>	<b>25.0%</b>	<b>1.32</b>	<b>3.97</b>	<b>9.2%</b>	<b>12.7%</b>
Mother's education	1.40	3.86	9.8	12.3	0.60	2.20	4.2	7.0
Father's education	1.58	3.97	11.0	12.7	0.73	1.77	5.1	5.6
<b>III. Demographic characteristics</b>	<b>0.80</b>	<b>1.31</b>	<b>5.6%</b>	<b>4.2%</b>	<b>0.06</b>	<b>0.32</b>	<b>0.4%</b>	<b>1.0%</b>
Female	0.00	0.00	0.0	0.0	0.03	0.08	0.2	0.2
Non-white	0.13	0.13	0.9	0.4	0.11	0.11	0.8	0.4
Family structure	-0.21	-0.24	-1.4	-0.8	-0.40	-0.49	-2.8	-1.6
Month of birth	-0.18	-0.05	-1.3	-0.2	0.02	0.00	0.1	0.0
Mother's age at birth	0.30	0.38	2.1	1.2	0.16	0.36	1.1	1.1
Number older siblings	0.22	0.45	1.5	1.4	0.06	0.17	0.5	0.5
Number younger siblings by 9	-0.03	-0.08	-0.2	-0.2	-0.02	-0.03	-0.1	-0.1
Twin	0.00	0.01	0.0	0.0	0.00	-0.01	0.0	0.0
English second language at 11	0.00	0.00	0.0	0.0	0.00	0.00	0.0	0.0
Parental employment at age 4	0.82	0.97	5.7	3.1	0.20	0.24	1.4	0.8
Parental health at age 4	-0.24	-0.25	-1.7	-0.8	-0.10	-0.11	-0.7	-0.4

(Table 4 cont'd)

<b>IV. School composition and quality</b>	<b>1.76</b>	<b>4.39</b>	<b>12.3%</b>	<b>14.0%</b>	<b>-0.03</b>	<b>1.46</b>	<b>-0.2%</b>	<b>4.6%</b>
Mean pupil KS1 (std score)	1.65	2.85	11.5	9.1	-0.77	-1.33	-5.4	-4.2
Mean VA KS1-2 (std score)	0.80	2.50	5.6	8.0	1.00	3.14	7.0	10.0
Proportion pupils FSM	-0.69	-0.96	-4.8	-3.1	-0.26	-0.36	-1.8	-1.1
<b>V. Pre-school environments</b>	<b>0.49</b>	<b>1.10</b>	<b>3.4%</b>	<b>3.5%</b>	<b>-0.19</b>	<b>-0.03</b>	<b>-1.3%</b>	<b>-0.1%</b>
Birth weight	0.18	0.22	1.3	0.7	0.05	0.06	0.3	0.2
Gestation	0.00	-0.01	0.0	0.0	0.00	-0.01	0.0	0.0
Breast feeding	0.24	0.68	1.7	2.2	0.25	0.69	1.7	2.2
Smoking in pregnancy	-0.37	-0.50	-2.6	-1.6	-0.25	-0.34	-1.7	-1.1
Post-natal depression	-0.07	-0.08	-0.5	-0.3	-0.08	-0.09	-0.5	-0.3
HLE	0.71	1.20	5.0	3.8	0.04	0.08	0.3	0.3
Read to daily at 3	-0.32	-0.50	-2.2	-1.6	-0.21	-0.32	-1.5	-1.0
Regular sleeping routine at 3	0.11	0.16	0.8	0.5	0.00	0.00	0.0	0.0
Centre-based care pre-3	-0.01	-0.04	0.0	-0.1	-0.01	-0.04	0.0	-0.1
Nursery age 3 to 4	0.01	-0.04	0.0	-0.1	0.01	-0.07	0.1	-0.2
<b>VI. Parent's attitudes and behaviours</b>	<b>2.61</b>	<b>6.46</b>	<b>18.2%</b>	<b>20.6%</b>	<b>1.11</b>	<b>2.69</b>	<b>7.7%</b>	<b>8.6%</b>
Mother found school valuable	0.06	0.13	0.4	0.4	0.05	0.11	0.4	0.3
Maternal locus of control	0.54	1.02	3.8	3.3	0.24	0.45	1.7	1.4
Mother hopes at least 1 A-level	0.14	-0.64	1.0	-2.0	0.06	-0.26	0.4	-0.8
Mother hopes university	1.95	6.03	13.6	19.3	0.81	2.52	5.7	8.0
Mother hopes other for child	-0.06	-0.42	-0.4	-1.3	-0.03	-0.19	-0.2	-0.6
Mother-ch interactions: ed	0.03	0.33	0.2	1.1	0.01	0.06	0.0	0.2
Mother-ch interactions: non-ed	-0.05	0.00	-0.3	0.0	-0.03	0.00	-0.2	0.0
<b>VII. Child attitudes and behaviours</b>	<b>2.74</b>	<b>6.14</b>	<b>19.1%</b>	<b>19.6%</b>	<b>1.07</b>	<b>2.54</b>	<b>7.4%</b>	<b>8.1%</b>
Ability beliefs	0.13	0.64	0.9	2.0	0.05	0.25	0.3	0.8
Locus of control	0.33	1.43	2.3	4.6	0.18	0.78	1.3	2.5
Enjoyment of school	0.04	0.04	0.3	0.1	0.01	0.01	0.0	0.0
School results important in life	0.80	1.25	5.6	4.0	0.27	0.42	1.9	1.3
Hobbies important in life	0.26	0.47	1.8	1.5	0.17	0.3	1.2	1.0
Possessions important in life	-0.24	-0.44	-1.7	-1.4	-0.16	-0.3	-1.1	-1.0
Anti-social behaviours	0.17	0.28	1.2	0.9	0.08	0.12	0.5	0.4
Hyperactivity	0.89	1.41	6.2	4.5	0.31	0.48	2.1	1.5
Emotional symptoms	0.00	0.00	0.0	0.0	0.01	0.02	0.1	0.1
Conduct problems (scale)	0.31	0.39	2.2	1.2	0.23	0.28	1.6	0.9
Experience of bullying (scale)	0.07	0.13	0.5	0.4	0.06	0.1	0.4	0.3
Pro-social behaviours	-0.11	0.05	-0.7	0.2	-0.05	0.02	-0.3	0.1
Peer problems (scale)	-0.17	-0.2	-1.2	-0.6	-0.12	-0.14	-0.8	-0.5
Teacher-child relations (scale)	-0.01	0.00	0.0	0.0	-0.02	-0.02	-0.2	-0.1
Leisure activities (scale)	0.23	0.68	1.6	2.2	0.07	0.21	0.5	0.7
<b>VIII. Missing flags</b>	<b>-0.49</b>	<b>-0.41</b>	<b>-3.4%</b>	<b>-1.3%</b>	<b>-0.25</b>	<b>-0.26</b>	<b>-1.7%</b>	<b>-0.8%</b>

The main body of the table shows how the numbers in the 'All measured factors' row can be broken down into different components. The first line shows that differences in parental education between deprived and more advantaged children can account for about 3 (8) percentile points of the

Key Stage 2 gap between bottom and middle (top) SEP quintile children, or about 21% (25%) of the total gap. This is the contribution of parental education over and above any influence through the mediating proximal influences we discuss further on. When we consider progress between 7

and 11, we see that parental education accounts for increases in the gaps of 1.3 and 4.0 percentile points between the lowest SEP and the two higher SEP groups, or 9% and 13% respectively of the total gaps. Parental education differences between socio-economic groups are clearly the most important factor in explaining the gaps in attainment at 7, and the gaps in progress during the primary school years. So it represents about a third of the educational inequality at age 11, net of prior attainment at age 7. The huge importance of parental education is particularly notable because, as noted previously, these numbers relate only to the direct effects of parental education, or the part that is not explained by any of the other (more proximal) variables in the model.

All other demographic influences contribute relatively little to the attainment gaps, although the younger age of mothers of disadvantaged children, and larger family size, have some explanatory power. So, lone parenthood, numbers of siblings, ethnicity, etc contribute little to attainment gaps in terms of total attainment at age 11 or progress since 7. Schools do matter, especially the school value-added measure between ages 7 and 11 for child progress, as might be expected. Our measures of parental attitudes, behaviours and beliefs explain in the region of a fifth of the gradients in the levels model, and around 8% when prior ability is held constant. Interestingly, children's views and behaviours independently explain a similar proportion to parents', in both the levels and value-added models, with their importance more than halved when Key Stage 1 is controlled. Unfortunately, without earlier measures of the key variables, we cannot distinguish whether success at Key Stage 1 leads to improvements in children's and parents' outlooks, or whether persistence in these factors means that causation runs the other way from attitudes and behaviours to achievement at 7. Differences in pre-school environments can account for little of the gaps. Hence parental attitudes and beliefs have an influence that is not far away from that of parental education (not transmitted by proximal factors) in driving the social gradient in attainment at age 11 and progress since age 7. Taken together, parent and child attitudes, beliefs and behaviours and school quality, explain over half of the attainment gaps between the most and least affluent children at 11 in the levels model. They explain around one-fifth in the value-added model,

or about half of the 13 percentile point gap that emerges between 7 and 11.

The method used here allows us to be more specific about the key drivers. Differences in maternal aspirations for university alone account for 19% of the top-bottom gap in levels terms and 8% in the value-added model. Assuming causality, equalising aspirations for university across socio-economic groups would narrow the Key Stage 2 gap by 6.0 percentile points or, if we assume that performance at 7 is fixed and determined by other factors, by 2.5 percentile points. Mother's locus of control and attitudes to education, play a smaller but not trivial supporting role in making parental attitudes and beliefs such important predictors of child outcomes.

The four dimensions of child behavioural problems – anti-social, hyperactivity, emotional symptoms and conduct problems – together account for 9.6% (6.6%) of the middle- (top-) bottom gap in level scores and 4.3% (2.9%) of the middle- (top-) bottom gap in value-added scores, a substantial amount relative to the other child-level variables. Of these, the greater emotional symptoms of disadvantaged children play little explanatory role, whereas hyperactivity and conduct problems are the most important. Similarly, it is the beliefs that school results and hobbies and interests are important in life, that drive the beneficial effects of the educational values of higher-income children, and a more internal locus of control, rather than a strong belief in their own ability, that drives the explanatory power of child self-concept.

Altogether, just the four factors of: maternal aspirations for university; hyperactivity; child locus of control; and the belief that school results are important in life, can account for 32% and 28% of the overall top-bottom and middle-bottom SEP gaps respectively. When prior ability is held constant, the equivalent proportions are 13% and 11% of the overall gaps, or around a third of the 5 and 13 percentile point gaps that open up between 7 and 11 among children from different socio-economic groups who started the period with the same Key Stage 1 performance.

## 5. Conclusions

It has long been known that children from more deprived backgrounds, achieve less well in terms of education attainment, than their more affluent

peers. More recently it has been shown that these attainment gaps start early in life and continue to widen through childhood (see Feinstein 2003 and 2004). Furthermore the extent of these gradients is not constant through time and has recently been diminishing (Gregg and Macmillan 2010). This research, with its two companion studies (Dearden et al 2010; Chowdry et al 2010) aims to show the extent of these gradients and give a sense of when in childhood they emerge for the most recent cohort data available. The main thrust of this study, is to explore the contribution that parent and child attitudes, beliefs and aspirations, make toward these gaps, and their emergence in the primary school years.

A number of features stand out from the study as worthy of note. First, around one third of the large educational gaps that are apparent at age 11 emerge through the primary school years. A sizable portion of these attainment gaps are associated with other aspects of families that differ by social background, especially parental education and also school quality. Parental education, which is obviously related to variations in economic circumstances within families, offers an important and distinct source of education attainment gradients. Hence as we consider the contribution of parental and child aspirations, beliefs and behaviours in driving the socio-economic gradients educational attainment, it is also important to note their role in driving gradients by parental education as well by SEP.

Parental aspirations and attitudes to education vary particularly strongly with socio-economic position. 81% of the richest mothers say they hope their 9-year-old will go to university, compared with only 37% of the poorest mothers. There are also large differences according to whether the mother found school valuable for themselves. Children's attitudes and behaviours in primary school vary in the degree to which they are socially graded. Poor children tend to view themselves as scholastically less able, are less likely to believe school results are important in life, and exhibit higher levels of hyperactivity, conduct problems and peer problems. However, their levels of school enjoyment and cooperative behaviour differ little from those of more affluent children. The attitudes, beliefs, and behaviours of parents and children have a major role to play in accounting for the gaps that remain unexplained. Together they explain

around half of the social gradient in educational attainment at age 11 when prior ability is not controlled, and around one-sixth when prior ability is held constant. Given that about 60% of the gaps in the latter model are explained solely by differences in outcomes at age 7, the figure of one-sixth of the total translates into around half of the increase in the gradient over the period. Hence these factors act as major transmission mechanisms for how both social disadvantage itself, and parental education, impact on educational attainment. The adverse attitudes to education of disadvantaged mothers are one of the single most important factors associated with the children's deficits at age 11. In particular, 9-year-olds whose mothers hope they will go to university, will score 2.5 percentile points higher at age 11 than the child of the parent with the lowest aspirations, given the same prior attainment and parental education etc. This factor alone explains a fifth of the test score gap between the richest and poorest children at 11, for given attainment at age 7. Likewise greater behavioural problems of disadvantaged children are the second key factor in accounting for their poorer educational outcomes. We find evidence that children with high levels of anti-social behaviours, hyperactivity and conduct problems at the ages of 8 to 9, scored lower at Key Stage 1, but even taking this into account, such behaviours appear to interfere with the learning process between 7 and 11. Other types of behaviour problems do not appear to play the same role. Yet whilst parental aspirations and behaviour problems are of key importance, a whole range of adverse attitudes, behaviours and beliefs contribute to the educational deficits of low income children. Other factors we identify as important are the lack of a sense of personal efficacy (both of mothers and their children), and the view that school results are not important in life.

The fact that parental aspirations and attitudes to education for their children and the child's own attitudes and behaviours are important to educational attainment in the primary school years, has a number of important policy implications. First, attempts to raise school results and progression through to university for reasonably able children from poorer families needs to start before secondary school starts. Furthermore, parents expectations and attitudes need to be shifted, not just those of the children. Poorer parental attitudes

do appear to reflect in part their own experiences of education, which were often negative, but they may also be substantially out of date in the educational opportunities open to middle ability children these days, with a third of children attending university. The pre-school environment is influential on attainment at 7 years, but perhaps

unsurprisingly, it plays little role in learning development through primary school. This suggests that our measures are not just reflecting an underlying marker of good parenting that drives attainment throughout childhood, as early markers of parenting predict early outcomes but not later progress.

## Acknowledgements

The authors gratefully acknowledge funding from the Joseph Rowntree Foundation (JRF), the Leverhulme Trust and the ESRC Large Grant RES-060-23-0011. We would also like to thank Lindsey Macmillan for research assistance to all the members of our advisory groups. All errors remain the responsibility of the authors.

## References

- Chowdry H, Crawford C and Goodman A. (2010) The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16. *Longitudinal and Life Course Studies*, 2, (59-76).
- Currie J and Stabile M. (2006) Child mental health and human capital accumulation: The case of ADHD. *Journal of Health Economics*, 25, 1094-1118.
- Dearden L, Sibieta L and Sylva K. (2010) The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study. *Longitudinal and Life Course Studies*, 2, 19-40.
- Feinstein L. (2003) Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 70, 73-98.
- Feinstein L. (2004) Mobility in Pupils' Cognitive Attainment During School Life. *Oxford Review of Economic Policy*, 20, No. 2: Education.
- Goodman R. (1997) The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Goodman A, Gregg P and Washbrook L. (2010) Children's educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood. *Longitudinal and Life Course Studies*, 2, 1-18.
- 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, 259-280.
- Harter S. (1985) *Manual for the self-perception profile for children*. University of Denver: Denver, CO.
- McLeod J and Kaiser K. (2004) Childhood Emotional and Behavioral Problems in Educational Attainment. *American Sociological Review*, 69, 636-658.
- NICHD Early Child Care Research Network and Duncan G. (2003). Modeling the Impacts of Child Care Quality on Children's Pre-school Cognitive Development *Child Development*, 74, 1454-1475.
- Wigfield A and Eccles JS. (2000) Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25, 68-81.

## Endnotes

<sup>i</sup> Avon was a former county area containing around 0.5 million residents in Bristol and the surrounding area in the West of England. It is now divided into four separate Local Authorities.

<sup>ii</sup> For information on ALSPAC, see <http://www.bris.ac.uk/alspac>.

<sup>iii</sup> The children in the ALSPAC cohort were age 3 to 4 between 1994 and 1996, before the introduction of guaranteed free half-day places.

# The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16

**Haroon Chowdry**

Institute for Fiscal Studies

[h.chowdry@ifs.org.uk](mailto:h.chowdry@ifs.org.uk)

**Claire Crawford**

Institute for Fiscal Studies and Institute of Education, University of London

**Alissa Goodman**

Institute for Fiscal Studies

(Received July 2010    Revised November 2010)

## Abstract

*It is well known that children growing up in poor families leave school with considerably lower qualifications than children from better off backgrounds. Using a simple decomposition analysis, we show that around two fifths of the socio-economic gap in attainment at age 16 can be accounted for by attainment at age 11, suggesting that circumstances and investments made considerably earlier in the child's life explain a sizable proportion of the gap in test scores between young people from rich and poor families. However, we also find that differences in the attitudes and behaviours of young people and their parents during the teenage years play a key role in explaining the rich-poor gap in GCSE<sup>1</sup> attainment: together, they explain a further quarter of the gap at age 16, and two fifths of the small increase in this gap between ages 11 and 16. On this basis, our results suggest that while the notion that "skills beget skills" implies that the most effective policies in terms of raising the attainment of young people from poor families are likely to be those enacted before children reach secondary school, policies that aim to reduce differences in attitudes and behaviours between the poorest children and those from better-off backgrounds during the teenage years may also make a significant contribution towards lowering the gap in achievement between young people from the richest and poorest families at age 16.*

JEL codes: I20, I32

**Key words:** socio-economic status, inequality, educational attainment, attitudes and behaviours

## 1. Introduction

Children growing up in poor families tend to emerge from school with considerably lower qualifications than children from better off backgrounds. As shown by the other papers in this Special Issue, these gaps are evident from an early age – even before starting school – and tend to widen throughout childhood. This paper complements the others in this issue, by seeking to explain socio-economic differences in attainment at age 16 – the point at which compulsory schooling ends and formal qualifications are typically first

obtained – as well as the small increase in the socio-economic attainment gap between ages 11 and 16.

There is a large literature from many countries which shows that family income and schooling attainment are very strongly correlated (see Shavit and Blossfeld 1993, for a review of 13 countries). Such differences in educational attainment are both an issue of policy concern in their own right, and are also critical for explaining the persistence of disadvantage across generations, an issue of particular concern in the UK where the degree of

inter-generational income mobility has been shown to be low by international standards (see Blanden et al 2005).

It is now widely accepted within the economics literature and elsewhere that an immediate lack of income, or access to other financial resources ('credit constraints') during the teenage years, is not largely responsible for the socio-economic gap in formal educational attainment, or in choices about whether to stay on in post-compulsory schooling or go to university (see Cameron and Taber 2004, Cameron and Heckman 1998 and 2001, and Carneiro and Heckman 2002, for the US; and Dearden et al 2004, and Chowdry et al 2008, for evidence from the UK). Much recent work on this topic instead points to the importance of parental behaviours and decisions in the very earliest years of a child's life, as potential explanations for the socio-economic gaps in educational attainment (see Cunha and Heckman 2007 for a review). Dearden, Sibieta and Sylva in this Special Issue, focus on just these issues for a cohort of children born in the UK in 2000–01.

However, work focusing on the early years alone does not allow us to examine why the gap between rich and poor children persists so strongly throughout childhood, and indeed widens with progression through the schooling system (see Feinstein 2003, and Goodman et al 2009, both for the UK). Neither is it informative about what policy interventions might be effective in raising the attainment of young people from poor backgrounds, once they have moved beyond early childhood.

The focus of this paper is on the extent to which differences between young people from rich and poor families in a range of parental and child attitudes and behaviours – such as educational aspirations, educational interactions in the home, family relationships, ability beliefs, and risky behaviours – during the *teenage* years might be important reasons why children from rich families outperform children from poor families at secondary school, and indeed why the gap between rich and poor continues to widen throughout secondary school.

In doing so, we follow a tradition pioneered by Sewell and Shah (1968) in the sociology literature, examining the role of a number of social-psychological factors in explaining the strong correlation between poverty and educational

attainment. It also complements a growing economics literature which emphasises the importance of the development of social skills and positive behaviours both for cognitive development and for longer-term labour market and social outcomes (see Bowles et al 2001, Heckman et al 2006, and Carneiro et al 2007). This literature increasingly emphasises that if policymakers wish to intervene during adolescence, policies that aim to improve young people's social skills and behaviours are likely to be more effective (and indeed more cost-effective) than interventions that directly seek to improve cognitive development (Cunha and Heckman 2007; Cunha, Heckman and Schennach 2010).

Our work also aims to inform a policy debate in the UK which has increasingly pointed towards improving parents' and young people's aspirations and other attitudes and behaviours, as a means of raising attainment at school among disadvantaged children (see Gutman and Akerman 2008, for a review).

Our work is based on data from the first three waves of the Longitudinal Study of Young People in England (LYSPE), a new study following a cohort of approximately 16,000 young people in English secondary schools from the ages of 13 and 14 onwards. This data is the first nationally representative survey in many years to follow a contemporary group of teenagers in England through secondary school. The survey was designed to allow an in-depth study of the experiences, attitudes, aspirations and motivations of a large group of today's teenagers and their families, and has also been linked to administrative records of national achievement test results, taken by study members at the ages of 11, 14 and 16. It therefore provides us with a unique opportunity to examine the factors associated with the gap in attainment, between pupils from rich and poor families, in secondary schools today.

Using this data, we set out the extent to which young people from rich and poor backgrounds differ in terms of their educational attainment at age 16 – and the change in their attainment between ages 11 and 16 – and use a simple 'decomposition' analysis to illustrate the extent to which these gaps can be explained by differences in other 'distal' factors (such as parental education and family structure) and a wide range of 'proximal' factors (or 'transmission mechanisms'), focusing



mainly on parents' and young people's attitudes and behaviours around age 14. We also consider the quality and composition of the secondary schools they attend.

We find that around *two fifths* of the very substantial gap in educational attainment between young people from rich and poor backgrounds at the end of secondary school (age 16), is accounted for by their attainment at the start of secondary school (age 11). This suggests that while early investments to improve the attainment of children from poorer backgrounds may be more cost-effective and/or more productive than later investments, there is still significant scope for intervention after children have started secondary school.

We also find that our observed measures of attitudes and behaviours of young people and their parents, seem to play a key role in explaining the socio-economic gap in attainment at age 16. We find that differences in these factors are able to explain just over a *quarter* of the socio-economic gap in attainment at age 16, and *two fifths* of the small increase in the rich-poor attainment gap between ages 11 and 16. While our work cannot shed light on the extent to which these transmission mechanisms might be responsive to public policy interventions, if they can be changed, then our results suggest that policies which reduce differences in attitudes and behaviours between richer and poorer children during the teenage years, may contribute towards lowering the gap in achievement age 16.

As with virtually all other work in this area, however, we must emphasise that this is not a *causal* analysis: we cannot be sure that there is no *unobserved heterogeneity* (unobserved factors which might be correlated with both the attitudes and behaviours we observe and with educational attainment) or *reverse causation* (that educational attainment might affect attitudes and behaviours rather than the other way round) which might plausibly account for some or all of the statistical associations we uncover. Whilst we acknowledge the shortcomings of our work in this regard, at the very least our findings can point to areas in which policy might be potentially effective, and where further investigation of a more experimental nature could be usefully deployed.<sup>2</sup>

This paper now proceeds as follows: Section 2 documents the inequalities in educational

attainment between teenagers from different socio-economic backgrounds that we seek to explain. Section 3 describes in further detail our data and methods. Section 4 highlights the attitudes and behaviours that are associated with higher GCSE attainment. Section 5 discusses the extent to which these attitudes and behaviours differ by socio-economic status, and Section 6 quantifies the contribution of these factors to the gap in educational attainment at age 16 between young people from rich and poor backgrounds. Section 7 discusses the extent to which our results suggest that there is an 'aspirations deficit', and Section 8 concludes.

## 2. Socio-economic inequalities in educational attainment at age 16

The degree of socio-economic inequality in educational attainment is highlighted in Figure 1, which is based on data from the LSYPE. The left hand panel shows the average percentile rank in the national achievement (Key Stage) test score distribution of young people in our sample at ages 11, 14 and 16, by quintile of parental socio-economic position (SEP). We discuss the data, our sample and our measure of SEP in more detail in Section 3.

This panel shows that, by age 11, there are already significant differences in test scores among children from different socio-economic backgrounds, with a typical gap of around 7 percentiles between each SEP quintile, and with the average scores of children from the most advantaged backgrounds 31 percentiles higher than those of children from the most disadvantaged backgrounds. These differences are more pronounced at age 14 (with a rich-poor attainment gap of 36 percentiles), before narrowing slightly by age 16 (to 33 percentiles) when the test scores represent GCSE results, the first formal academic qualifications taken in English schools.<sup>3</sup>

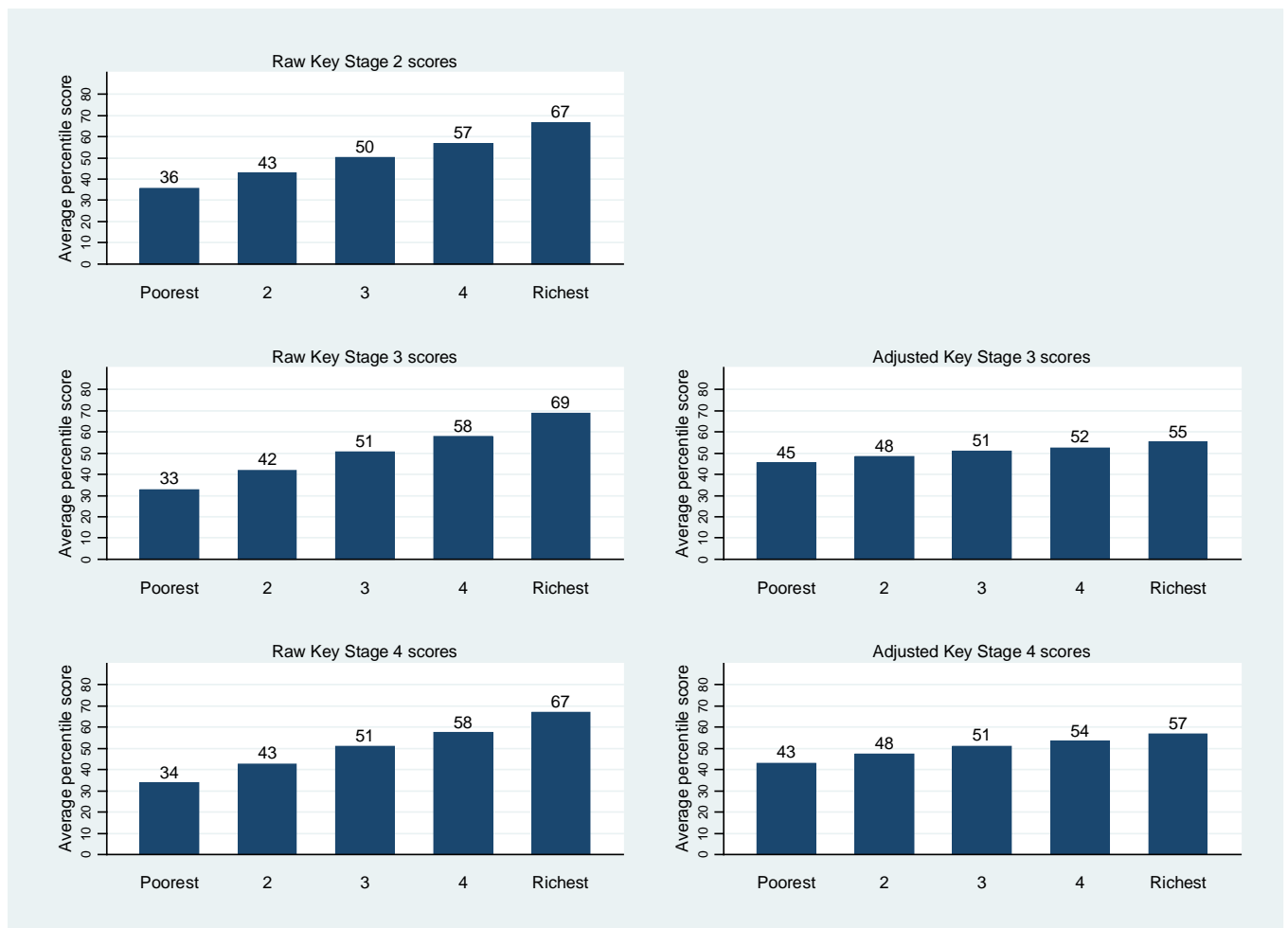
Table 1 shows how these differences in test scores translate into differences in the proportion of children reaching the Government's target (expected) level at each stage. For example, just one in five (21.4%) young people in the poorest SEP quintile attain five good GCSEs including English and maths (a common benchmark of attainment at age 16), compared to three-quarters (74.3%) of young people from the richest SEP quintile, a gap of 52.9 percentage points.

The right hand panel of Figure 1 shows how these socio-economic gaps change once prior attainment at age 11 has been taken into account. It does so by estimating an ‘adjusted’ gap, showing what the average percentile score by SEP quintile would be if all children had scored the same in achievement tests at age 11. This figure shows that the ‘adjusted’ percentile scores are much more equally distributed than the ‘raw’ ones, highlighting that a large fraction of the inequality in test scores observed at ages 14 and 16 is already reflected in

differences that are apparent by the end of primary school.<sup>4</sup> Indeed, we find that 56% of the gap in test scores at age 16 can be accounted for by differences in attainment that are apparent by the end of primary school.<sup>5</sup>

In our work that follows, we examine the extent to which these differences in educational attainment between young people from rich and poor families can be accounted for by a broad range of transmission mechanisms, which we describe in detail in the next section.

Figure 1. Test scores at ages 11, 14 and 16, by parental SEP quintile



**Note:** the right hand panel presents an ‘adjusted’ gap, showing the average percentile score by SEP quintile, assuming all children scored the same at age 11. Such estimates are derived by predicting each individual’s Key Stage 3 or 4 percentile in the situation where all pupils scored equally (i.e. at percentile 50.5) at Key Stage 2, based on a ‘value-added’ regression of the following form:  $KS_{it} = \alpha + \lambda SEP_i + \theta KS_{i11} + \epsilon_{it}$ ,  $t = 14, 16$ .

**Table 1. Proportion reaching national expected level, by SEP quintile**

	Average outcome by SEP quintile				
	Poorest	2	Middle	4	Richest
<b>Key Stage 2 (age 11)</b>					
% reaching expected level	64.3%	75.5%	84.2%	87.8%	94.3%
<b>Key Stage 3 (age 14)</b>					
% reaching expected level	51.9%	66.1%	77.4%	84.7%	92.7%
<b>Key Stage 4 (age 16)</b>					
% attaining 5+GCSEs A* - C	33.2%	46.4%	59.3%	70.6%	84.0%
% attaining 5+GCSEs A*-C including English and Maths	21.4%	33.6%	46.4%	57.9%	74.3%

**Notes:** authors’ calculations using Key Stage test scores from the National Pupil Database for the LSYPE cohort. Our sample includes all individuals for whom we observe Key Stage 2, 3 and 4 test scores.

### 3. Data and methodology

This paper is based on data from the Longitudinal Study of Young People in England (LSYPE).<sup>6</sup> The LSYPE is a longitudinal survey (administered at the school level) following around 16,000 young people in England who were aged 13 or 14 in the academic year 2003–04, and hence were born between September 1989 and August 1990. Interviews with both the young person and their main parent were carried out annually; additionally, school characteristics and Key Stage test results at ages 11, 14 and 16 have been matched in to the sample from administrative records held by the Department for Education. The full Wave 1 sample contains 15,770 individuals. We use the 13,343 young people with valid Key Stage 2, 3 and 4 results for our analysis. This implies, amongst other things, that we keep only state

school pupils in our sample, and that our sample is of slightly lower socio-economic position than if we had not imposed such restrictions.<sup>7</sup> Our analysis is based on data from Waves 1, 2 and 3 (ages 14, 15 and 16), before young people left compulsory education.

Underlying our analysis is a model linking a young person’s socio-economic status and other aspects of their family background – including the secondary school they attend – to educational attainment at age 16, via a set of potential ‘transmission mechanisms’, including parent and child attitudes and behaviours (see Goodman, Gregg and Washbrook in this Special Issue for a more detailed account of this model). This model is estimated as per equation 1 below:

$$KS_{is16} = \alpha + \beta_1 SEP_{is} + \beta_2 FAM_{is} + \beta_3 SCH_s + \beta_4 PAR_{is} + \beta_5 YP_{is} + e_{is} \tag{1}$$

where  $KS_{16}$  represents attainment at age 16 for individual  $i$  in school  $s$ ,  $SEP$  represents quintiles of our index of socio-economic position,  $FAM$  is a vector of demographic and family background characteristics,  $PAR$  is a vector of parental attitudes and behaviours,  $YP$  is a vector of the young person’s attitudes and behaviours, and  $e$  is an individual

error term. We describe each of these groups of factors in more detail below.

When we consider the change in attainment between ages 11 and 16, we add an additional control for attainment at age 11 to equation 1, as follows:

$$KS_{is16} = \alpha + \beta_1 SEP_{is} + \beta_2 FAM_{is} + \beta_3 SCH_s + \beta_4 PAR_{is} + \beta_5 YP_{is} + \beta_6 KS_{11is} + e_{is} \tag{2}$$

Equations 1 and 2 are used to assess the determinants of educational attainment at age 16, and of academic progress between ages 11 and 16, the results of which are discussed in Section 4. They are also used as the basis for a simple

decomposition analysis of the gap in attainment at age 16 between young people from the top and bottom SEP quintiles (discussed in Section 6). In these decompositions, the contribution of each variable to the overall SEP gap is given by the size of

its conditional correlation with educational attainment (its coefficient in equation 1 or 2 above), multiplied by the extent to which it varies with SEP (the difference between the mean values of the variable in the top and bottom SEP quintiles).<sup>8</sup> See Goodman, Gregg and Washbrook in this Special Issue for a more formal treatment of this approach.

We now move on to discuss the outcomes and covariates in our model.

Our main outcome of interest is the percentile rank of the young person's total point score at age 16. When we consider changes in attainment during secondary school, we additionally control for attainment at age 11 ( $KS_{11}$  in equation 2 above). To do so, we rank children according to their average point score across tests in English, maths and science, and group them into quintiles (fifths) of the sample on the basis of this measure. We include indicators for the top four quintiles in our analysis (with the bottom quintile omitted as the reference category).

Our measure of parental socio-economic position (SEP in the equation above) aims to capture the long-term resources of the household in which the young person lives, and is constructed from: log average equivalised household income across ages 14, 15 and 16; reported experience of financial difficulties at age 14; mother's and father's occupational class at age 14; and housing tenure at age 14. We use polychoric principal-components analysis to combine this information into an index, on the basis of which we can rank individuals from lowest to highest SEP.<sup>9</sup> (Note that the first principal component explains 53% of the variation in these factors.) We group the young people in our sample into quintiles on the basis of this measure, and include indicators for the richest four in our model (such that the lowest SEP quintile is the reference category).

We account for a range of demographic and other family background characteristics ( $FAM$  in equations 1 and 2 above), including gender, ethnicity, month of birth, whether English is an additional language, birth weight, mother's and father's highest educational qualification, mother's and father's employment status, mother's and father's health status, lone parent status, mother's age, and the number of older and younger siblings.

We also account for a range of characteristics of the young person's secondary school ( $SCH$  in equations 1 and 2 above), including school type,

whether the school has a sixth form, whether it is a grammar school, the school's Key Stage 2 to Key Stage 4 value-added score, the average Key Stage 2 scores of the young person's year group, the gender mix of school, school size, percentage of pupils eligible for free school meals, whose first language is not English, who are non-white, and what the young person believes their friends will do at age 16.

The key potential transmission mechanisms between SEP and educational attainment that we consider, are informed by a diverse literature on the determinants of attainment (see Goodman, Gregg and Washbrook in this Special Issue for a discussion), which variously emphasise the importance of parental influences, and the motivations and self-regulation of young people themselves. They are summarised as follows:

**Young people's own attitudes and behaviours** ( $YP$  in the equation above):

- **Aspirations and expectations** for future education;
- **Self-concept:** ability beliefs; the intrinsic (enjoyment) and extrinsic (worth) value placed on education by the young person, and the young person's locus of control;
- **Job/career values:** whether having a job and/or a career is important to the young person;
- **Engagement in risky and positive behaviours:** relating to education (truancy, suspension, and exclusion); anti-social and criminal behaviour (shoplifting, fighting, vandalism, graffiti, trouble with the police); use of substances (alcohol, smoking, and drug use); and positive activities (sport, reading for pleasure, and cultural and religious participation);
- **Experiences of bullying;**
- **Teacher-child relations:** how much the child likes their teacher; and their perception of how they are treated relative to others in the class.

**Parental attitudes and behaviours** ( $PAR$  in equations 1 and 2 above):

- **Parental aspirations and expectations** for the child's future education, and the value placed on education by the parent;
- **Parental involvement in the child's education,** such as helping with homework, discussing school reports and subject choice in Year 10, and involvement in school life;

- **Parental closeness:** frequency of spending time together as a family, including sharing family meals and going out; frequency of conflict in the home;
- **Educational resources:** availability of material resources relating to education in the home, including provision of private tuition (in both school and non-school subjects), and access to a computer and the internet.

See Chowdry et al (2009) for full details of how these measures are constructed.

#### 4. What influences GCSE attainment?

In this section, we discuss the results from two simple multivariate regression models (based on equations 1 and 2 in Section 3) which examine the

correlates of GCSE attainment, without and with controls for attainment at age 11 respectively.

Tables 2 and 3 present coefficients on our measures of parental and young people's attitudes and behaviours respectively. (All other coefficients can be found in Supplementary Material published with this paper.) Column 1 presents the results of our "levels" analysis (i.e. without controls for prior attainment) and Column 2 presents the results of our "value-added" analysis (including controls for attainment at age 11). Each coefficient represents the average change in Key Stage 4 percentile rank associated with a unit change in the variable in question.

Column 1 of Tables 2 and 3 suggests that, conditional on family background and school characteristics, a range of parental and young person attitudes and behaviours seem to be strongly positively associated with attainment at age 16.

**Table 2. Influences on GCSE attainment: parental attitudes and material resources**

	Age 16 attainment (percentiles)	Value-added between age 11 and age 16 (percentiles)
Parent education value (scale)	0.099 [0.42]	0.387 [1.84]
Parent wants young person (YP) to stay in FTE at 16	2.089 [1.42]	0.818 [0.58]
Parent wants young person to learn a trade or go into training or an apprenticeship at 16	-0.907 [0.60]	-0.641 [0.45]
Parent has other aspirations for young person at 16	5.177* [2.21]	2.789 [1.34]
Parent thinks YP very/fairly likely to go to university	9.861** [16.60]	4.270** [7.87]
Parent-child education interactions (scale)	0.544 [1.92]	0.222 [0.89]
Family-child interactions (scale)	1.241** [3.42]	1.671** [5.00]
Parental involvement in school activities (scale)	0.511 [0.95]	0.68 [1.38]
Young person has private tuition	1.219* [2.47]	1.402** [3.16]
Young person has computer at home	3.053** [3.94]	2.821** [4.03]
Young person has internet access at home	2.432** [4.07]	1.575** [2.92]

**Notes:** table contains selected coefficients from two OLS regression models of the determinants of Key Stage 4 (age 16) percentile score. Missing dummies are included, where appropriate, but their coefficient estimates are not shown. Both models also include controls for demographic and family background characteristics, secondary school characteristics, and young person's attitudes and behaviours. The second model (shown on the right) additionally controls for attainment at Key Stage 2 (age 11). Coefficient estimates for young person's attitudes and behaviours can be found in Table 3. All other coefficient estimates can be found in our online appendix. Standard errors are robust, corrected for clustering at the school level and shown in parentheses. \*\* indicates significance at the 1% level; \* indicates significance at the 5% level.

**Table 3. Influences on GCSE attainment: young people's attitudes and behaviours**

	Age 16 attainment (percentiles)	Value-added between age 11 and age 16 (percentiles)
<b>Young person's attitudes and behaviours</b>		
Ability beliefs (scale)	8.559** [21.04]	3.633** [9.67]
Enjoyment of school (intrinsic value scale)	-2.498** [7.23]	-1.033** [3.33]
Usefulness of school (extrinsic value scale)	2.758** [7.87]	2.410** [7.57]
Locus of control (scale)	3.629** [8.21]	2.570** [6.62]
Wants to stay on in full-time education (FTE) at 16	2.811** [2.68]	1.186 [1.27]
Wants to leave FTE at 16 but return later	-0.171 [0.10]	-0.641 [0.44]
Wants to learn a trade/go into training	1.371 [1.14]	-0.283 [0.26]
Other intentions at 16	0.433 [0.22]	0.601 [0.33]
Likely to apply to HE, and likely to get in	3.779** [4.66]	2.283** [3.19]
Likely to apply to HE, but not likely to get in	2.075* [2.34]	1.549* [1.98]
Not very likely to apply to HE, but likely would get in	3.660** [3.81]	1.427 [1.63]
Not very likely to apply to HE, and not likely to get in	1.653* [2.03]	1.342 [1.89]
Job aspirations (scale)	-0.056 [0.18]	0.314 [1.16]
Experience of bullying (scale)	-4.353** [13.00]	-2.467** [8.10]
Education behavioural difficulties (scale)	-3.201** [6.71]	-2.634** [7.48]
Anti-social behaviour (scale)	-1.901** [4.51]	-1.994** [5.50]
Smokes cigarettes frequently	-6.557** [6.14]	-6.408** [6.08]
Drinks alcohol frequently	-0.31 [0.38]	-1.074 [1.45]
Has smoked cannabis	-0.481 [0.64]	-2.119** [2.91]
Teacher-child relations (scale)	2.169** [4.95]	3.050** [7.32]
Plays sport weekly	0.434 [0.80]	0.324 [0.65]
Reads every week	1.920** [4.17]	0.898* [2.14]
Plays a musical instrument	2.682** [5.55]	1.132** [2.64]
Engages in other positive activities	1.070* [2.48]	0.462 [1.21]
<b>Observations</b>	<b>13,343</b>	<b>13,343</b>
R-squared	0.54	0.63

**Notes:** table contains selected coefficients from two simple OLS regression models of the determinants of Key Stage 4 (age 16) percentile score. Missing dummies are included, where appropriate, but their coefficient estimates are not shown. Both models also include controls for demographic and family background characteristics, secondary school characteristics, parental attitudes and material resources. The second model (shown on the right) additionally controls for attainment at Key Stage 2 (age 11). Coefficient estimates for parental attitudes and material resources can be found in Table 2. All other coefficient estimates can be found in the Supplementary Material. Standard errors are robust, corrected for clustering at the school level and shown in parentheses. \*\* indicates significance at the 1% level; \* indicates significance at the 5% level.

Aspirations and expectations for future education seem to be of particular importance: for example, young people whose parents think that they are very or fairly likely to go to university score, on average, nearly 10 percentiles higher in their GCSEs than young people whose parents do not hold such views. Similarly, young people who think it likely that they will apply to, and get into, higher education, score on average nearly 4 percentiles higher at age 16, than young people who think it not at all likely that they will apply to university.

A young person's beliefs in their own ability are also strongly positively correlated with their attainment at age 16. For example, a 1 standard deviation increase in our ability beliefs scale (which comprises measures of how good young people think they are at maths, English, science and ICT, as well as school work overall at age 14) is associated with an 8.6 percentile increase in GCSE scores. By contrast, engagement in a range of risky behaviours is strongly negatively associated with attainment at age 16. For example, young people who smoke regularly (at least six cigarettes per week at age 14) score 6.6 percentiles lower, on average, than young people who do not. Similarly, a 1 standard deviation increase in our educational behavioural difficulties scale (which encompasses measures of truancy, suspension and expulsion from school) is associated with a 3.2 percentile fall in attainment at age 16.

Young people whose parents are able to provide them with material resources for educational purposes at home also tend to score more highly at age 16 than those whose parents are not. For example, young people with access to both a computer and the internet at home, score, on average, 5.5 percentiles higher than those who do not. The provision of private tuition is also associated with a small increase in attainment (of around 1.2 percentiles).

Column 2 of Tables 2 and 3 shows how these relationships change once we add controls for attainment at age 11 to our model. If the coefficients in Column 2 are smaller than those in Column 1, then we can interpret this difference as the extent to which the effects of these attitudes and behaviours (and their unobserved correlates) have already been crystallised in test scores at the end of primary school, with any significant differences remaining in Column 2 attributed to the

effect they have during the secondary school years (between ages 11 and 16).

For example, young people whose parents think that they are fairly or very likely to go to university, score just over 4 percentiles higher at age 16 than those who do not, even after controlling for attainment at age 11. This compares with a performance advantage of nearly 10 percentiles when we did not account for attainment at age 11. This suggests that parents who think their child is likely to go to university at age 14, will probably have developed this view earlier in the child's life, such that it may already have had some effect on the young person's attainment by the end of primary school.

By contrast, it is interesting to note that the magnitude of the effects of engagement in risky behaviours on attainment do not appear to be affected by the inclusion of controls for attainment at age 11. For example, young people who smoke regularly still score 6.5 percentiles lower, on average, than those who do not, and a 1 standard deviation increase in our criminal behaviour scale (which comprises measures of involvement in graffiti, vandalism, shoplifting and fighting) is associated with a 1.9 percentile reduction in GCSE scores, both before and after controlling for attainment at age 11. Given how unlikely it is for primary school children to have engaged in these types of behaviour, this is a plausible result, suggesting that the effects of these behaviours on attainment occur solely during the secondary school years.

In summary, we find that, even after controlling for a wide range of family background and school characteristics, as well as attainment at age 11, young people are more likely to do well in their GCSEs if:

#### The young people:

- Have a greater belief in their own ability at school;
- Find school worthwhile;
- Have a more external locus of control (i.e. believe that their actions have consequences);
- Think it is likely that they will apply to, and get into, higher education;
- Do not experience bullying;
- Avoid risky behaviours such as smoking, taking cannabis, anti-social behaviour and truancy;
- Have a good relationship with their teachers;
- Engage in positive activities, such as reading and playing a musical instrument.

**Their parents:**

- Think it is likely that the young person will go on to higher education;
- Spend time sharing family meals and outings, and quarrel relatively infrequently;
- Devote resources towards education, such as private tuition, computer and internet access.

whether these factors also differ by socio-economic status, and thus whether we might expect differences in such characteristics to help explain why young people from poor families score so much lower in their GCSEs, than young people from rich families.

Tables 4 and 5 present differences in parental attitudes and material resources, and young people’s attitudes and behaviours, by socio-economic position (SEP) quintile. (SEP differences in family background and school characteristics can be found in the Supplementary Material.) These tables show that there are large and significant differences between young people from the richest and poorest fifths of our sample, in terms of the majority of these characteristics, including many of those that we found to be strongly associated with attainment at age 16 in the previous section.

**5. Socio-economic differences in attitudes and behaviours**

We have shown (in Section 2) that young people from rich and poor families differ in terms of how well they perform in national exams. We have also established (in Section 4) that, even after controlling for a wide range of family background and school characteristics, as well as attainment at age 11, a variety of attitudes and behaviours are still significantly associated with attainment at age 16. In this section, we move on to document

**Table 4. Parental attitudes and material resources, by SEP quintile**

	Poorest SEP quintile (Q1)	2 <sup>nd</sup> SEP quintile (Q2)	Middle SEP quintile (Q3)	4 <sup>th</sup> SEP quintile (Q4)	Richest SEP quintile (Q5)	Q5-Q1
Parent education value (scale, SDs)	0.106	-0.024	-0.024	-0.030	-0.014	-0.120**
Parent wants YP to stay in FTE at 16	75.8%	75.8%	76.8%	84.0%	91.0%	15.2ppts**
Parent wants YP to learn a trade or go into training or an apprenticeship at 16	19.1%	20.9%	20.2%	13.7%	7.2%	- 11.9ppts**
Parent has other aspirations for YP at 16	1.2%	1.3%	1.6%	1.6%	1.5%	0.3ppts
Parent thinks YP very/fairly likely to go to university	53.4%	52.3%	57.4%	66.3%	80.7%	27.4ppts**
Parent-child education interactions (scale, SDs)	-0.312	-0.076	0.077	0.120	0.191	0.503**
Family-child interactions (scale, SDs)	-0.063	-0.022	-0.003	0.021	0.069	0.132**
Parental involvement in school activities (scale, SDs)	-0.079	-0.037	0.001	0.026	0.088	0.167**
Young person has private tuition	10.3%	17.9%	25.3%	33.6%	45.6%	35.4ppts**
Young person has computer at home	71.4%	86.8%	94.4%	96.5%	99.4%	28.0ppts**
Young person has internet access at home	45.5%	67.9%	82.9%	90.0%	96.7%	51.1ppts**

**Notes:** \*\* indicates significance at the 1% level; \* at the 5% level.



Table 5. Young people's attitudes and behaviours, by SEP quintile

	Poorest SEP quintile (Q1)	2 <sup>nd</sup> SEP quintile (Q2)	Middle SEP quintile (Q3)	4 <sup>th</sup> SEP quintile (Q4)	Richest SEP quintile (Q5)	Q5-Q1
Ability beliefs (scale, SDs)	-0.091	-0.063	-0.018	0.028	0.125	0.217**
Enjoyment of school (scale, SDs)	-0.073	-0.057	0.003	0.023	0.068	0.141**
Usefulness of school (scale, SDs)	-0.090	-0.058	0.022	0.062	0.145	0.235**
Locus of control (scale, SDs)	-0.098	-0.051	0.010	0.013	0.070	0.168**
Wants to stay on in full-time education (FTE) at 16	78.7%	80.1%	82.9%	88.2%	93.0%	14.4ppts**
Wants to leave FTE at 16 but return later	2.3%	2.0%	2.2%	1.7%	1.3%	-0.9ppts*
Wants to learn a trade/go into training	8.6%	9.3%	8.9%	5.5%	3.4%	-5.3ppts**
Wants to enter FT work at 16, w1	8.6%	6.6%	4.6%	3.4%	1.5%	-7.1ppts**
Other intentions at 16	1.8%	1.8%	1.3%	1.1%	0.7%	-1.0ppts**
Likely to apply to HE, and likely to get in	49.2%	49.9%	56.9%	63.2%	76.8%	27.6ppts**
Likely to apply but not likely to get in	9.6%	10.7%	7.1%	8.3%	7.0%	-2.6ppts*
Not very likely to apply to HE, but likely would get in	6.4%	8.6%	7.8%	8.3%	5.2%	-1.1ppts
Not very likely to apply to HE, and not likely to get in	15.5%	14.9%	15.2%	9.8%	6.3%	-9.3ppts**
Job aspirations (scale, SDs)	-0.041	-0.001	0.016	0.039	-0.003	0.037
Experience of bullying (scale, SDs)	0.103	0.039	-0.011	-0.037	-0.081	-0.184**
Education behaviour difficulties (scale)	0.146	0.039	-0.043	-0.062	-0.119	-0.265**
Anti-social behaviour (scale, SDs)	0.149	0.035	-0.002	-0.043	-0.119	-0.268**
Smokes cigarettes frequently	6.2%	5.2%	3.5%	2.6%	1.3%	-4.9ppts**
Drinks alcohol frequently	5.3%	7.0%	7.5%	9.0%	7.8%	2.5ppts**
Has smoked cannabis	10.3%	9.7%	8.8%	9.1%	8.3%	-2.0ppts
Teacher-child relations (scale)	-0.089	-0.064	-0.024	0.021	0.095	0.184**
Plays sport weekly	75.8%	77.7%	81.4%	80.9%	85.8%	9.9ppts**
Reads every week	69.7%	71.0%	75.2%	76.6%	81.4%	11.7ppts**
Plays a musical instrument	12.4%	17.5%	20.8%	24.9%	35.4%	23.1ppts**
Engages in other positive activities	53.2%	56.3%	59.4%	64.1%	69.6%	16.5ppts**

Notes: \*\* indicates significance at the 1% level; \* at the 5% level.

Table 4 shows that:

- **Education aspirations and expectations:** richer parents tend to have higher aspirations and expectations for their children's education than poorer parents. For example, four out of five parents in the top SEP quintile think that their child is likely to apply to university, compared to just over half of parents in the bottom SEP quintile at age 14.
- **Family interactions:** parents in the top SEP quintile are more likely to help their children with their homework (education interactions scale), more likely to get involved in school activities, and more likely to share family meals

or argue less frequently with their children (family-child interactions scale) than parents in the bottom SEP quintile.

- **Computer and internet at home:** almost all young people from the richest families have access to a computer and the internet at home, compared to just over 70 per cent of young people from the poorest families with access to a computer, and under half with access to the internet.

Table 5 shows that:

- **Intrinsic/extrinsic value of schooling and locus of control:** young people from poorer families are less likely to enjoy school, less likely to find

school valuable, and less likely to believe that their own actions make a difference (have an ‘external locus of control’) than young people from richer families.

- **Education aspirations and expectations:** young people from richer families tend to have higher educational aspirations and expectations than young people from poorer families, with nearly four-fifths of teenagers in the top SEP quintile thinking it is likely that they will apply to university (and get in), compared to less than half of teenagers in the bottom SEP quintile, a gap of almost 30 percentage points.
- **Risky behaviours and positive activities:** young people from poorer families are more likely to engage in a range of risky behaviours (such as smoking, taking cannabis, playing truant and other anti-social activities) at age 14 than young people from richer families, while they are less likely to engage in positive activities such as playing sports, reading for pleasure, and playing a musical instrument.
- **Experiences of bullying:** Young people from poorer backgrounds are also more likely to experience frequent bullying at age 14 than young people from richer backgrounds.
- To summarise, this section has shown that there are substantial differences between young people from rich and poor families in terms of their attitudes to education, and their propensity to engage in a range of risky behaviours as teenagers. In the next section, we consider whether these differences can help to explain the socio-economic gaps in educational attainment that we highlighted in Section 2.

## 6. Can differences in attitudes and behaviours help to explain the socio-economic gap in educational attainment at age 16?

Section 2 documented the very large gaps in educational attainment between young people from rich and poor families. In this section, we try to explain why these differences arise. Of particular interest to us is the importance of attitudes and behaviours of young people and their parents during the teenage years, which Section 4 showed to be strongly associated with GCSE attainment and Section 5 showed to differ markedly by socio-economic background.

We use a simple decomposition analysis to investigate the extent to which attitudes and behaviours during the teenage years play an important role in explaining why children from poor families end up with worse GCSE results than children from rich families.

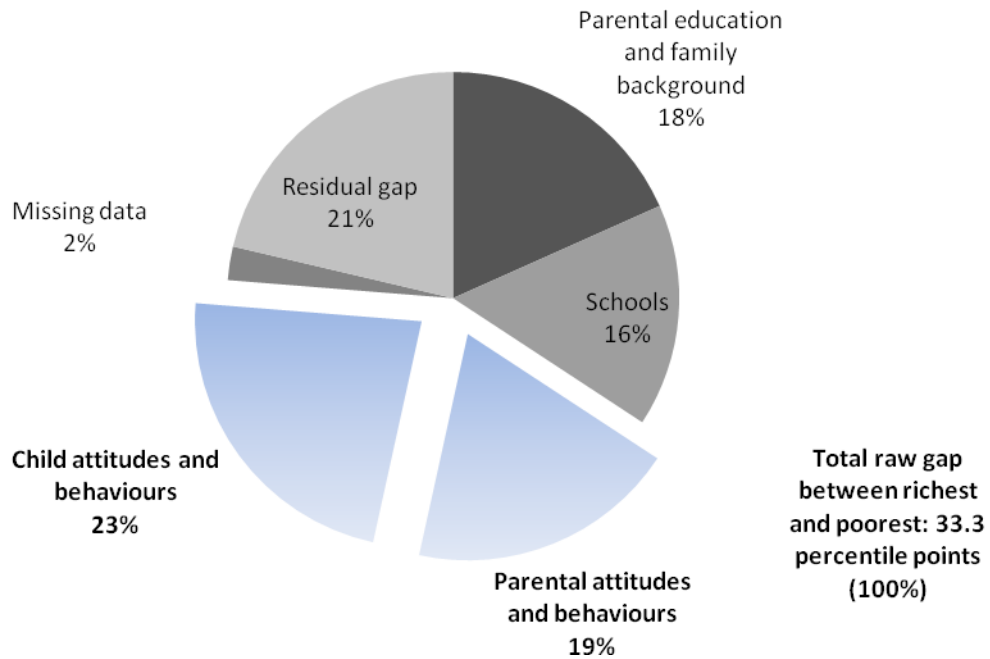
We decompose the very large gap in educational attainment at age 16 (33.3 percentile points), between young people from the top and bottom SEP quintiles, into the contribution made by each characteristic in our model. As set out in Section 3, these relative contributions are calculated by multiplying the difference between the proportions of rich and poor children with each characteristic, by their coefficient estimates from a regression model including all characteristics simultaneously. We do this separately without (as per equation 1) and with (as per equation 2) controls for attainment at age 11.

Figure 2 presents the results of our “levels” decomposition (without controls for attainment at age 11). It shows that our observed measures of parental and young people’s attitudes and behaviours, together account for just over 40% of the gap in attainment at age 16 between young people from the richest and poorest families. More detailed analysis (not shown) suggests that differences in attitudes and expectations towards further and higher education (of both parents and children) are responsible for nearly two fifths (16%) of this contribution, with a further fifth (8%) arising from differences in the provision of material resources for educational purposes and one seventh (6%) arising from differences in the ability beliefs of young people from rich and poor families.

Figure 2 also suggests that differences in the secondary schools attended by young people from rich and poor backgrounds explain a sizeable proportion (16%) of the difference in GCSE test scores. A more detailed breakdown of this contribution suggests that it is differences in the average Key Stage 2 scores of the young person’s year group, and the school’s Key Stage 2 to Key Stage 4 value-added score, which are driving this effect.

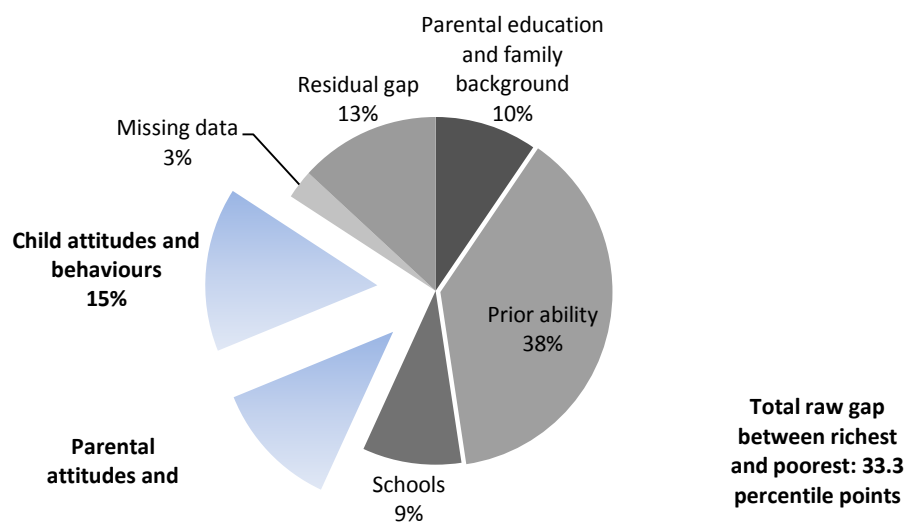
However, the remaining direct contributions of differences in demographic and other family background characteristics – as well as the sizeable ‘residual’ gap (which can be regarded as the direct effect of SEP on attainment) – suggests that our observed measures of attitudes and behaviours are not capturing all of the mechanisms through which differences in socio-economic background give rise to differences in educational attainment.

**Figure 2. Explaining the socio-economic gap in attainment at age 16 (without controls for attainment at age 11): decomposition analysis**



**Notes:** the relative contribution of each set of factors is calculated by multiplying the difference in the proportions of rich and poor with each characteristic (shown in the Supplementary Material) by the coefficient estimates from a regression model, including all characteristics simultaneously (shown in Tables 2, 3 and in the Supplementary Material).

**Figure 3. Explaining the socio-economic gap in attainment at age 16 (with controls for attainment at age 11): decomposition analysis**



See notes to Figure 2.

Figure 3 examines the extent to which these conclusions still hold, once we control for prior attainment at age 11. The inclusion of this measure switches the interpretation of our results to contributions to *academic progress during secondary school* (between ages 11 and 16). In this way, it allows us to capture any observed and unobserved differences, between children from rich and poor backgrounds, that have already had an effect on attainment by the end of primary school. As a result, we might expect the magnitude of the residual gap – as well as the contributions of other observed factors in our model – to fall, once we include controls for prior attainment.

This is indeed what we see. Figure 3 shows that differences in attainment at age 11 explain nearly 40% of the gap in GCSE scores between young people from rich and poor families.<sup>10</sup> Of the remainder, just over *two fifths* (27% of the overall gap) is accounted for by our observed measures of parental and young person's attitudes and behaviours. Just under *one third* (19% of the overall gap) is accounted for by the direct effects of family background and secondary school characteristics, leaving around *one fifth* (13% of the overall gap) unexplained.

As expected, the contribution of each of these sets of characteristics is smaller here than in Figure 2, when we did not include controls for prior attainment. There are at least two possible explanations for this finding: 1) at least some of our observed measures of attitudes and behaviours are correlated with earlier measures of similar concepts, which in turn affect primary school attainment; 2) we are now taking into account unobserved characteristics that are correlated both with our observed measures of attitudes and behaviours and attainment. In reality, both mechanisms may be present.

What can we conclude from these results? While the notion that “skills beget skills” (e.g. Cunha and Heckman 2007) suggests that it remains important to invest as early as possible in a child's life in order to reap the greatest benefits of later investments, the fact that only 40% of the gap in attainment at age 16 can be explained by what has happened up to the end of primary school suggests that, even during secondary schooling, it is not too late to intervene to try to close the socio-economic gap. And while our results certainly cannot be regarded as causal, it is interesting to note that a

sizeable proportion of the gap in progress between ages 11 and 16 seems to be explained by our observed measures of parental and child attitudes and behaviours.

The extent to which attitudes and behaviours may truly represent a route through which the socio-economic gap in attainment at age 16 can be reduced, hinges crucially on the extent to which attitudes and behaviours during the secondary school years are malleable and responsive to public policy interventions. While we cannot address this issue directly in our work, others (e.g. Cunha and Heckman 2007) have suggested that non-cognitive skills (including attitudes and behaviours) are considerably less malleable at later than earlier ages, meaning that it would be considerably more expensive to achieve the same degree of change in the teenage years as it would in the pre-school or primary school years. Nonetheless, our interpretation of these results is one of hope for policymakers seeking to reduce the gap in GCSE qualifications between young people from rich and poor families: secondary school is not too late, and it appears that attitudes and behaviours might be one possible route through which such gaps can be reduced.

## 7. Is there an ‘aspirations deficit’?

In the last section, we concluded that there may be a sizeable role for attitudes and behaviours to play in reducing the gap in GCSE attainment between young people from rich and poor families. However, in this section, we sound a note of caution in drawing such a conclusion.

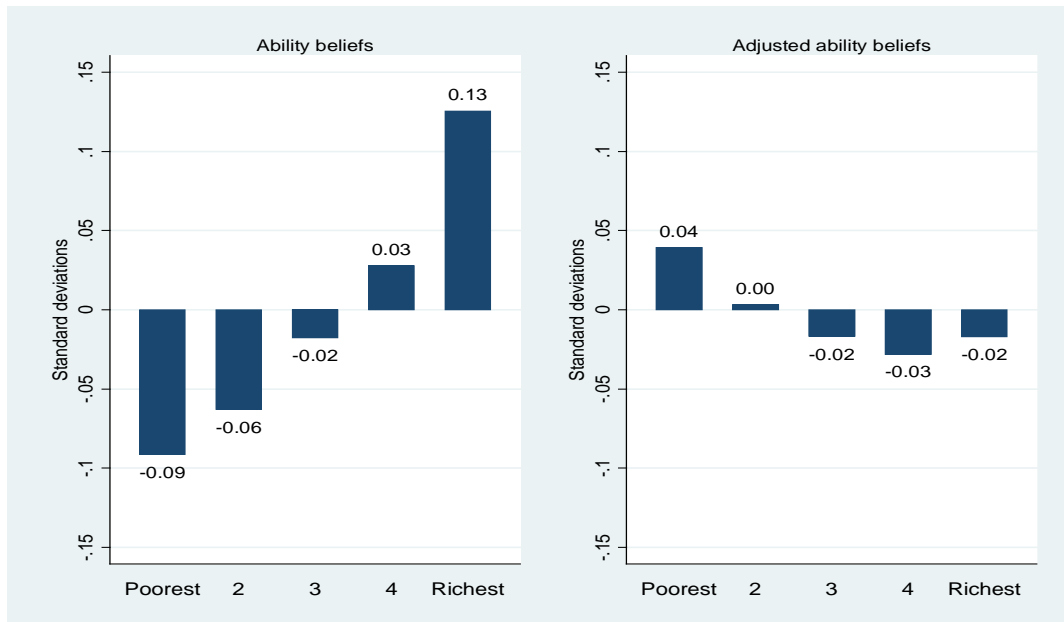
We saw in Sections 4 to 6 that ability beliefs may help to explain the socio-economic gap in attainment at age 16 (as they are both highly correlated with GCSE scores and strongly socially graded). The socio-economic gradient in our ability beliefs scale is shown in the left hand panel of Figure 2. On this evidence, one might be tempted to conclude that if young people from poorer backgrounds can be encouraged to have more confidence in their own ability, then the socio-economic gap in test scores at age 16 may be somewhat reduced.

However, this is not necessarily true. The right hand panel of Figure 4 illustrates what happens to the socio-economic gradient in ability beliefs if we take account of test scores at age 11. It shows that poor children do not necessarily under-estimate

how well they do at school, because once we take into account their earlier performance, they are typically *more* likely to think that they are good at school than young people from richer

backgrounds.<sup>11</sup> This is consistent with a story in which young people compare themselves to peers from similar backgrounds.

**Figure 4. Young person’s ability beliefs, by SEP quintile (age 14)**

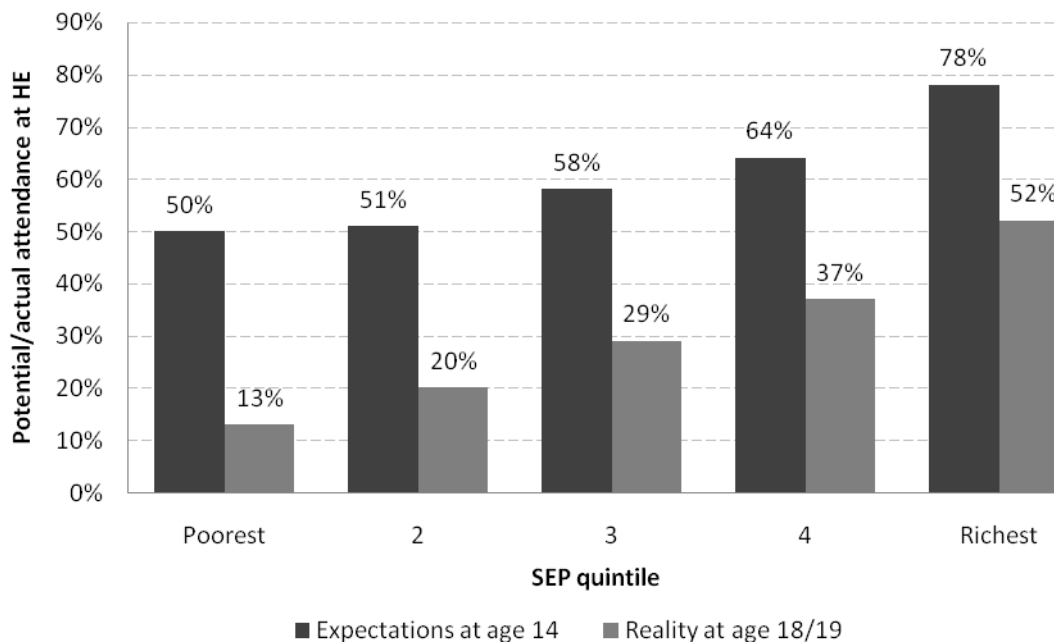


**Note:** the right hand panel presents average ability beliefs by SEP quintile, assuming all children had the same attainment at age 11. These estimates are derived in the same way as those in Figure 1.

There is also a steep socio-economic gradient in young people’s expectations for higher education: we saw in Section 4 that there is a gap of almost 30 percentage points between the proportion of young people from the richest and poorest families who think that they will apply to university and get in. On this evidence, it is tempting to conclude that if we can encourage more young people from poorer backgrounds to aspire to go to university, then we may be able to reduce the socio-economic gap in test scores at age 16.

However, aspirations for higher education are *high across the board*: many more young people, from *all* socio-economic backgrounds, think that they will apply to and get into university than are

likely to do so in practice. This is borne out by comparing HE expectations amongst the LSYPE cohort at age 14 with administrative data on actual HE participation by age 19 for a slightly older cohort. For example, while almost half (49%) of young people from the poorest fifth of the LSYPE sample report that they are likely to go to university, only one in eight (13%) of the poorest fifth among the slightly older cohort actually did so. Similarly, almost four fifths (78%) of young people from the richest quintile of the LSYPE think that they are likely to go to university, compared with just over half (52%) of the older cohort who actually did go.

**Figure 5. Comparing HE expectations at age 14 with HE participation at age 18/19**

**Notes:** we do not observe actual HE participation among the LSYPE cohort yet; the comparison instead use figures on HE participation derived from linked administrative data combining individuals' school, further and higher education records for two cohorts who sat their GCSEs in 2001–02 and 2002–03. This means that they are slightly older than the LSYPE cohort, who sat their GCSEs in 2005–06. It should also be noted that the deprivation quintiles are also defined in a slightly different way in the two datasets.

These results highlight a potentially less than straightforward role for attitudes and behaviours in helping to close the socio-economic attainment gap.

## 8. Conclusions

It is well known that children growing up in poor families tend to emerge from school with considerably lower qualifications than children from better off backgrounds. This paper has examined the determinants of the socio-economic gap in GCSE results using a simple decomposition analysis. Of particular interest has been the role of attitudes and behaviours.

Our work shows that around *two fifths* of the gap in educational attainment at age 16 between young people from rich and poor families can be accounted for by attainment at age 11. This suggests that circumstances and investments made considerably earlier in the child's life are an important driver of the socio-economic gap in test scores at the end of compulsory schooling.

However, we also find a potentially significant role for our observed measures of attitudes and behaviours of young people and their parents: together, they explain a further *quarter* of the socio-economic gap in GCSE attainment, and *two fifths* of the small increase in the rich-poor attainment gap between ages 11 and 16.

We must interpret these findings with caution, however, for at least two reasons.

First, as with virtually all work in this area, we must emphasise that this is not a *causal* analysis: we cannot be sure that there is no *unobserved heterogeneity* or *reverse causation* which might plausibly account for some or all of the statistical associations we uncover. However, whilst we acknowledge the potential shortcomings of our work in this regard, the richness of the LSYPE data, coupled with the results of Crawford, Goodman and Joyce in this Special Issue, suggest that our findings regarding the relative importance of our observed measures of attitudes and behaviours, are unlikely to be seriously biased by the omission of detailed

measures of parental ability and social skills (which we might have regarded as the most important potential sources of omitted variables bias).

Second, our work has highlighted some important nuances that should be borne in mind when making policy recommendations on the basis of such results. For example, we find that many more young people think that they will apply to university (and be accepted) than are ultimately likely to do so. This suggests that simply improving HE aspirations and expectations amongst teenagers from poor backgrounds, is unlikely to eliminate the large socio-economic gap in HE participation that exists in the UK. Similarly, while we find substantial socio-economic differences in ability beliefs, this does not necessarily suggest that young people from poor families under-estimate how well they do at school;

indeed, once we account for prior attainment at age 11, teenagers in the lowest SEP group are actually *more* likely to think they are good at school than young people from the highest SEP group.

Even with these caveats in mind, however, our results still suggest that, while the notion that “skills beget skills” implies that the most effective policies in terms of raising the attainment of young people from poor families are likely to be those enacted before children reach secondary school, policies that aim to reduce differences in attitudes and behaviours between the poorest children and those from better-off backgrounds during the teenage years may also make a significant contribution towards lowering the gap in achievement between young people from the richest and poorest families at age 16.

## Acknowledgements

The authors gratefully acknowledge funding from the Joseph Rowntree Foundation (JRF) and the former Department for Children, Schools and Families (DCSF, now the Department for Education). In particular, we would like to thank Michael Greer and Andrew Ledger of DCSF, and Helen Barnard and Chris Goulden of JRF. We would also like to thank all the members of our advisory groups. All errors remain the responsibility of the authors.

## References

- Blanden J, Gregg P and Machin S. (2005) *Intergenerational Mobility in Europe and North America: a report supported by the Sutton Trust*. Centre for Economic Performance, London School of Economics, London.
- Bowles S, Gintis H and Osborne M. (2001) The Determinants of Earnings: A Behavioral Approach. *Journal of Economic Literature*, 39, 1137–1176.
- Cameron S and Heckman J. (1998) Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males. *Journal of Political Economy*, 106, 262-333.
- Cameron S and Heckman J. (2001) The Dynamics of Educational Attainment for Black, Hispanic and White Males. *Journal of Political Economy*, 109, 455-499.
- Cameron S and Taber C. (2004) Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112, 132-182.
- Carneiro P, Crawford C and Goodman A. (2007) The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes. Centre for Economic Performance Discussion Paper No. 92.
- Carneiro P and Heckman J. (2002) The Evidence on Credit Constraints in Post-Secondary Schooling. *Economic Journal*, 112, 705–734.
- Chowdry H, Crawford C, Dearden L, Goodman A and Vignoles A. (2010) Widening Participation in Higher Education: Analysis using Linked Administrative Data, Institute of Fiscal Studies Working Paper W10/04.
- Chowdry H, Crawford C and Goodman A. (2009) Drivers and barriers to educational success: evidence from the Longitudinal Study of Young People in England. Department of Children, Schools and Families Research Report No. RR102.
- Cunha F and Heckman J. (2007) The technology of skill formation. *American Economic Review*, 97, 31-47.
- Cunha F, Heckman J and Schennach S. (2010) Estimating the technology of cognitive and non-cognitive skills formation. *Econometrica*, 78, 883-931.
- Dearden L, McGranahan L and Sianesi B. (2004) The role of credit constraints in educational choices: evidence from the NCDS and BCS, CEE Working Paper No. 48.
- Feinstein L. (2003) Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 70, 73-97.

- Goodman A and Gregg P. (2010) Poorer children's educational attainment: how important are attitudes and behaviours? Report for the Joseph Rowntree Foundation  
<http://www.jrf.org.uk/publications/educational-attainment-poor-children>
- Goodman A, Sibieta L and Washbrook L. (2009) Inequalities in education outcomes among children aged 3 to 16. Report for the National Equality Panel,  
<http://www.equalities.gov.uk/pdf/Inequalities%20in%20education%20outcomes%20among%20children.pdf>
- Gutman L and Akerman R. (2008) Determinants of aspirations. Centre for Research on the Wider Benefits of Learning Research Report No. 27.
- Gutman L and L. Feinstein L. (2008) Children's Well-Being in Primary School: Pupil and School Effects. Centre for Research on the Wider Benefits of Learning Research Report No. 25.
- Heckman J, Stixrud J and Urzua S. (2006) The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24, 411–482.
- Sewell W and Shah V. (1968) Social class, parental encouragement and educational aspirations. *American Journal of Sociology*, 73, 559-572.
- Shavit Y and Blossfeld H. eds. (1993) *Persistent Inequality: changing educational attainment in thirteen countries*. Westview Press, Colorado.
- Thomas S, Sammons P, Mortimore P and Smees R. (1997) Differential secondary school effectiveness: examining the size, extent and consistency of school and departmental effects on GCSE outcomes for different groups of students over three years. *British Educational Research Journal*, 23, 451-469.

---

## Endnotes

- <sup>1</sup> General Certificate of Secondary Education. This is the standard academic qualification taken in England, Wales and Northern Ireland at the end of secondary education (usually at the age of 16).
- <sup>2</sup> Successful identification strategies are extremely hard to come by in this area, since attitudes and behaviours are not randomly allocated across individuals, and experimental variation in psychological and behavioural phenomena is generally extremely rare or non-existent. In other work (Chowdry et al 2009) we unsuccessfully attempted to use policy interventions that might plausibly be thought to introduce exogenous variation in specific child attitudes and behaviours to achieve more rigorous identification.
- <sup>3</sup> Note that these results are based on total Key Stage 4 points scored (including marks awarded for qualifications that are regarded as equivalent to GCSEs). The gap between age 14 and age 16 still narrows if we do not include GCSE equivalents in our total point score, but if we use a capped point score (representing the student's eight best exam results), the gap between 14 and 16 remains roughly constant.
- <sup>4</sup> The fact that the adjusted gap in Key Stage 3 scores (10 percentiles) is smaller than the adjusted gap in Key Stage 4 scores (14 percentiles) simply suggests that Key Stage 2 scores are more strongly correlated with attainment at age 14 than at age 16, which seems plausible.
- <sup>5</sup> This figure is obtained from a decomposition analysis similar to that described in Section 6, but with age 11 attainment as the only control.
- <sup>6</sup> See [www.esds.ac.uk/longitudinal/access/lstype/L5545.asp](http://www.esds.ac.uk/longitudinal/access/lstype/L5545.asp) for more information on the LSYPE.
- <sup>7</sup> Individuals with missing test scores tend to be from either very low SEP backgrounds (such as those who were not entered for the tests) or very high SEP backgrounds (such as those in private schools), with the former slightly outweighing the latter in this case.
- <sup>8</sup> Note that the 'contribution' of a variable to the SEP gap says nothing about statistical significance: it depends on the magnitude of estimated coefficients but not the precision with which they are estimated.
- <sup>9</sup> We have also carried out our analysis using family income instead of socio-economic position. This makes little difference to our findings. Results are available from the authors on request.
- <sup>10</sup> Note that this contribution increases to 61% if we also include controls for attainment at age 14 (see results in our Supplementary Material).
- <sup>11</sup> The right hand panel of Figure 2 presents average ability beliefs by SEP quintile, assuming all children had the same attainment at age 11. These estimates are derived in the same way as those in Figure 1.



# Explaining the socio-economic gradient in child outcomes: the inter-generational transmission of cognitive skills

**Claire Crawford**

Institute for Fiscal Studies and Institute of Education, University of London

[claire\\_c@ifs.org.uk](mailto:claire_c@ifs.org.uk)

**Alissa Goodman**

Institute for Fiscal Studies

**Robert Joyce**

Institute for Fiscal Studies

(Received July 2010 Revised November 2010)

## Abstract

*Papers in this Special Issue and elsewhere consistently find a strong relationship between children's cognitive abilities and their parents' socio-economic position (SEP). Most studies seeking to explain the paths through which SEP affects cognitive skills suffer from a potentially serious omitted variables problem, as they are unable to account for an important determinant of children's cognitive abilities, namely parental cognitive ability. A range of econometric strategies have been employed to overcome this issue, but in this paper, we adopt the very simple (but rarely available) route of using data that includes a range of parental characteristics measured during the parents' childhood, such as parental cognitive ability and social skills. In line with previous work on the intergenerational transmission of cognitive skills, we find that parental cognitive ability is a significant predictor of children's cognitive ability; moreover, it explains one sixth of the socio-economic gap in those skills, even after controlling for a rich set of demographic, attitudinal and behavioural factors. Despite the importance of parental cognitive ability in explaining children's cognitive ability, however, the additional parental characteristics we examine here do not alter our impression of the relative importance of other factors in explaining the socio-economic gap in cognitive skills. This is reassuring for studies that are unable to control for such characteristics.*

JEL codes: I20, I32

**Key words:** cognitive skills, intergenerational transmission, socio-economic gap

## 1. Background and motivation

Papers in this Special Issue and elsewhere have consistently found a strong relationship between a child's cognitive skills and his or her parents' socio-economic position (henceforth SEP). Previous research on this topic has sought to understand both: i) whether it is parents' SEP which influences children's cognitive test scores, or other correlated factors (Dahl and Lochner 2008; Brooks-Gunn and Duncan 1997; Mayer 1997), and ii), the

transmission mechanisms through which parents' SEP might influence test scores (taking i) as given) (Guo and Harris 2000; Davis-Kean 2005; Jean Yeung et al 2008, plus papers in this issue). Parents are likely to have many characteristics that are correlated both with SEP and with their children's cognitive outcomes. Many of these characteristics are rarely observed, such as social skills, other personality traits and (at least in UK data) parental

cognitive ability. This presents a potentially serious omitted variables problem with considerable policy implications. By omitting important variables from models relating parents' SEP to child outcomes, researchers risk both overstating the importance of parental SEP in determining children's cognitive outcomes, and falsely pinpointing specific transmission mechanisms between them. Such mistakes could lead to misguided policy recommendations, both in terms of how much raising the living standards of children growing up in poverty is likely to improve their cognitive development (and hence, through various channels, their own living standards as adults); and indeed whether specific interventions, such as those that target particular hypothesised transmission mechanisms, are likely to be effective.

Such latent factors may play a large role in explaining why children from poorer families have worse cognitive outcomes than children from richer families: simply put, more able parents are likely both to be able to command higher incomes in the labour market, and might also raise more cognitively developed children for other reasons, for example through better parenting skills, a greater preference for educational investments in their children, or due to some direct, more genetic link in cognitive ability between parents and their offspring. Given the growing literature on the importance of an individual's non-cognitive skills for a wide range of economic and social outcomes (see, for example, Blanden et al 2007; Heckman et al 2006; Carneiro et al 2007), a similar line of argument applies to the omission of these types of parental characteristics in such analyses.

There are a significant number of studies employing a range of econometric strategies to tackle the problem of omitted variables in estimating the relationship between parental income and a number of child outcomes, among which measures of children's educational attainment feature heavily and measures of cognitive development feature occasionally. However, there are few papers, to our knowledge, that attempt to identify the potential transmission mechanisms through which income may affect these outcomes in a similarly robust framework.

Typical among the approaches taken in the studies that do attempt to address the issue of correlated unobservable characteristics are those that adopt some sort of experimental or quasi-

experimental approach (Duflo 2000; Mayer 2007; Duncan Morris and Rodrigues 2006); those that make use of instrumental variables (IV) techniques (Shea 2000; Dahl and Lochner 2008); and those that use sibling differences or other fixed effects-type methodologies (Blau 1999; Duncan et al 1998; Levy and Duncan 2000; Ermisch et al 2002). While these approaches share the advantage that they can, under certain assumptions, identify the causal effects of specific observed variables in the absence of any data on unobserved characteristics, they also suffer from obvious drawbacks. First, the identifying assumptions are often problematic, such as exogeneity of the instrument in IV approaches and strict exogeneity in fixed-effects studies (which rules out 'feedback' effects within families in the case of sibling difference models). Second, even if the identifying assumptions are valid, we learn nothing about the key sources of the omitted variables bias: we only consistently estimate the causal effects of things we observe, but the things we do not observe may be at least as important (in terms of explaining the outcome of interest) and thus of immense policy relevance.

Few papers follow our approach, which is to take the very simple (but rarely available) route of utilising data in which some of these variables are observable. Here, we make use of a dataset – the British Cohort Study (BCS) – that provides detailed information on cognitive ability, SEP, family background characteristics, health, social skills and attitudes and behaviours across two generations within the same families. We therefore observe the cognitive ability of (one of) the parents of each child in our sample, as well as other parental characteristics that are often unobservable to the researcher looking at the determinants of children's cognitive skills. The contribution of this paper is not to identify the causal effect of SEP on children's cognitive abilities, but rather to investigate how this additional information changes our understanding of what matters for cognitive development.

Like the other papers in this series, we emphasise that our analysis is not causal: we cannot rule out endogeneity due to omitted variables or reverse causation. Instead, we aim to identify possible transmission mechanisms between SEP and cognitive skills that can rarely be investigated in studies of this kind, thus highlighting areas where future experimental research aimed at identifying causal effects may be best targeted. Crucially, we

also aim to supplement the other papers in this series by assessing whether information about additional parental characteristics, such as cognitive ability, might change their conclusions about which transmission mechanisms matter. To do so, we compare the results obtained from regressions of a child's cognitive ability on a wide range of explanatory variables with and without including often unobserved parental characteristics in the model. We also use these estimates to 'decompose' the SEP gap in cognitive skills into the contributions of various groups of factors, again comparing the results obtained with and without accounting for the parental characteristics that are often unobservable to the researcher.

By doing this we also contribute to a wider literature on the intergenerational transmission of cognitive abilities, in which the correlation of cognitive ability between parents and children is well documented (for recent evidence, see Anger and Heineck 2010; Black et al 2009; Björklund et al 2009). Genetic inheritance is one obvious potential explanation for the observed correlation in ability across generations, but the extent of its explanatory role is contested: this is a complex literature spanning a number of disciplines, and one in which the simplistic dichotomy between 'nature vs nurture' in explaining child development is increasingly recognised as overly simplistic and obsolete (see Cunha and Heckman 2007). Decomposing environmental from genetic effects is complicated by the likelihood that the environments people choose (or parents choose on their behalf) are ones which complement their genetic endowments (Dickens and Flynn 2001; Loehlin 2002; Rowe and Rodgers 2002; Dickens and Flynn 2002).

Complementing this literature, we find a very strong correlation between cognitive ability across generations, as well as a strong relationship between family socio-economic status and cognitive ability in both generations. Our decompositions suggest that often unobserved information on parental cognitive ability, social skills and attitudes to education, explains almost one quarter of the gap in cognitive test scores between high and low SEP children. This mostly reflects the importance of parental cognitive ability, which accounts for 16 percent of the gap after controlling for a wide range of mechanisms through which cognitive ability may be passed across

generations (for example, differences in the home learning environment that parents provide for their children), and 50 percent of the gap if we do not include such factors. As the importance of parental ability remains substantial after controlling for so many of the mechanisms through which it may plausibly affect child's cognitive ability, one possibility – though not the only one – is that genetic transmission of cognitive skills is significant. Reassuringly for previous studies, however, we find that accounting for these additional parental characteristics does *not* hugely alter our impression of which factors observed during the *current* generation are important for cognitive skills (attitudes towards education and social skills remain the most important predictors).

This paper now proceeds as follows: Section 2 provides a brief description of the BCS data and some summary statistics. Section 3 describes the omitted variable problem that studies of children's cognitive skills typically face when key parental characteristics are unobserved, and illustrates the likely extent of this problem using some simple analysis of the BCS data, which includes many of these characteristics. Sections 4 and 5 present our main analysis, starting with our regression results (in Section 4), which are then used as the basis for a decomposition analysis of the SEP gap in children's cognitive skills (in Section 5). Section 6 concludes.

## 2. Data

The British Cohort Study (BCS) sampled all people born in Great Britain in a particular week in April 1970 and has surveyed them at various points throughout their lives.<sup>1</sup> We have data from seven waves: at the cohort members' birth, and at ages 5, 10, 16, 26, 29 and 34. Importantly for our purposes, the age 34 wave (in 2004) randomly selected half of the cohort members (henceforth CMs) who lived with their natural or adopted children for an additional battery of questions about those children.<sup>2</sup> This means that we have rich measures of cognitive ability, social skills, attitudes and behaviours and family background characteristics across two generations within the same family. Note, however, that while the CMs are all the same age, the children of the CMs were all interviewed at the same time and are therefore of different ages. This is an important issue to which we return later.

Our main outcome of interest is the cohort member's child's cognitive ability, as measured in

2004 using the British Ability Scales (BAS).<sup>3</sup> Children aged 3 to 5 were tested on vocabulary and early number concepts; those aged 6 and over were tested on word reading, spelling and number skills.

We would like to age-normalise these test scores to take account of the fact that cognitive performance tends to increase with age. This is typically achieved by using the residuals from a regression of test scores on age. The problem with this procedure in our data is that, as explained, the age of the child is collinear with the age of one of their parents (the cohort member) at the child's birth, which is itself correlated with a wide range of possible determinants of cognitive skills (such as SEP and parental ability). Thus, a straightforward age-normalisation would effectively be 'partially normalising' with respect to these possible determinants of cognitive skills (because they are correlated with age). Clearly this could be undesirable, since it is precisely the impact of these variables on cognitive skills that we want to investigate. To avoid this problem whilst normalising for the actual 'effect' of age as best we can, we run the following two-step procedure when defining our outcome variable:

- 1) Estimate the equation: cognitive test score =  $\text{age}'\alpha + X'\beta + u$
- 2) Estimate the 'residual': cognitive test score –  $\text{age}'\alpha$  by replacing  $\alpha$  with its first stage estimate.

In the above, 'age' is a matrix of dummy variables capturing each possible age of the child in months, 'X' is a matrix of all other variables that we use in our full regression specification including a constant (see sections 4 and 5),  $\alpha$  and  $\beta$  are coefficient vectors, and  $u$  is a vector of error terms. The 'residual' (estimated in step 2) is our age-normalised score. This differs from a more standard age-normalisation, because we control for  $X$  in the first stage. Having carried out the age-normalisation, we arrange the normalised scores into percentile ranks for our analysis. Note that we have repeated our analysis with test scores age-normalised in the standard way, and our results are extremely similar. In addition, as we emphasise later, our main conclusions also hold if we disaggregate the sample of children by age.

Our aim is to explain variation in this age-normalised measure of children's cognitive ability using not only information from the current generation (such as SEP and other family background characteristics, plus attitudes and

behaviours of both the parents and their children<sup>4</sup>), but also information that is often unobserved from the CM's own childhood. The source of this often unobserved information is the first four waves of the BCS (from birth to age 16). Specifically, we make use of cognitive ability test scores at ages 5 and 10, social skills measures at ages 5 and 10 and a range of other information on attitudes and behaviours at ages 10 and 16.

At age 5, the cognitive ability measure includes tests on vocabulary, copying designs, human figure drawing, and profile recognition. At age 10, it includes tests from the British Ability Scales (BAS), as well as additional tests of reading, vocabulary, writing, spelling, maths, copying, sentence formation and sequence recognition. At each age, we average across standardised scores from each test and split the sample of cohort members into cognitive ability quintiles on the basis of this average.

The social skills measures are based on mother reports of the Rutter behaviour scale<sup>5</sup> at ages 5 and 10 and the Conners behaviour scale at age 10 (see Conners 1969). Other attitudinal information from the parents' childhood that we control for includes: measures of self-perceived ability, self-esteem and self-concept at ages 10 and 16; and attitudes at age 16 towards school, the prospect of post-compulsory education, and the importance of earning lots of money in the future.<sup>6</sup>

We are also interested in the extent to which these factors can help to explain the gap in cognitive ability scores between children from rich and poor families. As described in Goodman, Gregg and Washbrook 2010, we follow the standard approach of using polychoric principal components analysis to construct an index of socio-economic position, using information on permanent family income,<sup>7</sup> housing tenure, parental occupational class and an indicator of financial difficulty (the first principal component explains 39% of the variation in these variables). Where there is missing or insufficient information on one or more of these dimensions (but not all of them), we use multiple imputation techniques to help construct the SEP index. The SEP gaps in each component of the SEP index are in Table A2 in our Supplementary Files.

Table 1 presents some summary statistics for various subsamples of the BCS cohort, from the original childhood cohort (column 1) to our estimation sample of the CMs' children in 2004

(column 5). It shows that almost 50% of the original sample of CMs had dropped out of the survey by 2004 (compare columns 1 and 2). This attrition was non-random: the CMs who dropped out before 2004 were relatively more likely to be from low SEP backgrounds and of low cognitive ability, and slightly more likely to be male.

Since the CMs are the parents of the children we study, the parents are all the same age.<sup>8</sup> Some of them will go on to have children, but had not done so by age 34 when surveyed. Therefore, those CMs with children in our sample are relatively likely to have characteristics that are not only correlated with having children (which follows trivially), but correlated with having children at a *younger* age: in our data, such characteristics include being female, less cognitively able, less well educated, and in a lower SEP group (compare columns 2 and 3). This aspect of the sample selection offsets some, but not all, of the over-sampling of high-SEP and high cognitive ability CMs caused by attrition (although clearly we cannot be sure that these two aspects of the sample selection act to 'offset' each other in terms of characteristics that we do *not* observe).

Column 4 additionally restricts our sample to the CMs who were included in the additional child surveys, and whose children completed the

cognitive ability tests. We restrict our attention to 3 to 16 year-old children, since they were all assessed using the British Ability Scale, and the developmental scores for under-3s are not comparable. This leaves us with 3,416 children from 2,059 families as our main estimation sample (column 5). This age restriction again excludes children whose parents were older when they were born, with the same implications for the gender, SEP, cognitive ability and education balance of the parents of the children in our estimation sample.

### 3. Parental ability and the omitted variable problem

The other papers in this Special Issue describe how children growing up in poorer families have worse cognitive outcomes than children growing up in richer families, and seek to explain this association by considering factors *observed in the current generation* (such as the home-learning environment and parents' attitudes to education) which are correlated with both SEP and cognitive outcomes. The BCS allows us to extend this analysis back a generation by taking account of much more detailed information about one of the children's parents.

Table 1. BCS sample descriptive

Variable	Statistic	1	2	3	4	5
		Original cohort <sup>a</sup>	In 2004 wave	With children	With children with cognitive test scores	With 3-16 year-old children with cognitive test scores
CM's SEP quintile during childhood	1 (poorest), %	20	16	18	17	18
	5 (richest), %	20	24	22	22	19
CM's SEP quintile, age 34 <sup>b</sup>	1 (poorest), %	-	-	-	17	20
	5 (richest), %	-	-	-	22	18
CM's cognitive ability quintile, age 10	1 (lowest), %	20	16	16	14	16
	5 (highest), %	20	23	21	22	19
CM is male	%	52	48	41	38	34
Birthweight (kg)	Mean	3.27	3.31	3.31	3.33	3.31
CM stayed in FT education after 16	%	-	54	49	51	47
CM has (at least) degree level qualification	%	-	36	31	31	26
CM is lone parent	%	-	-	11	11	13
Mean age of CM's children	Mean	-	-	5.8	5.5	7
<b>Number of CMs</b>		<b>18,332</b>	<b>9,507</b>	<b>5,875</b>	<b>2,573</b>	<b>2,059</b>
<b>Number of CMs' children</b>		-	-	<b>11,083</b>	<b>4,515</b>	<b>3,416</b>

a) The 'original' cohort includes 1,409 CMs who were not surveyed in the first wave (at birth) but were surveyed later in childhood (at ages 5, 10 or 16). It excludes 273 CMs who were surveyed as children but did not provide sufficient information for us to construct their SEP index.

b) Note that the age 34 SEP distribution from which we define SEP quintiles is the distribution of the CMs' children, since they are our estimation sample (not the CMs themselves). There are thus more CMs in the top quintile than in the bottom quintile in column 4, since those in the top quintile have fewer children, on average.

In particular, we have data on the performance of the CMs in cognitive tests when they were children. The potential significance of this additional information is large. Children’s cognitive skills, parents’ cognitive skills and SEP could all be causally related<sup>9</sup> and are certainly correlated: more able parents are likely to command higher incomes in the labour market, and might also raise more cognitively developed children for other reasons, for example through better parenting skills. To the extent that the effects of parental ability on cognitive test scores are not perfectly captured by other factors observed in the current generation (for example, the quality of the home learning environment that they provide for their children), analysis of the link between children’s cognitive skills and SEP has an omitted variable problem to contend with when parental ability is unobserved.<sup>10</sup> Here, as background to our analysis, we document the extent of the association between parental

cognitive ability, child cognitive skills and SEP in our sample.

Table 2 shows the extent of the intergenerational links in cognitive ability; or, to put it another way, the raw correlation between our outcome of interest (children’s cognitive ability) and an explanatory variable that is not usually observable in most UK datasets (parental cognitive ability). We split both children and their parents into ability quintiles, creating 25 possible combinations of child and parental ability.<sup>11</sup> The Table shows, for given parental ability quintiles, the percentage of children in each ability quintile (so the numbers in each row sum to 100, subject to rounding). There is clear persistence of high and low cognitive abilities across generations: for example, of those children whose parent was in the lowest cognitive ability quintile at age 10, 29% are in the lowest quintile themselves, with just 14% in the highest quintile.<sup>12</sup>

**Table 2. Cognitive ability across generations**

% from each parental ability quintile	Child’s cognitive ability quintile				
	Low	2	3	4	High
↓ Parent’s cognitive ability quintile, age 10					
Low	29	23	19	14	14
2	25	20	19	17	19
3	16	21	22	22	19
4	15	19	19	23	24
High	12	18	20	23	27

**Table 3. The association between parental ability and socio-economic position**

% from each parental ability quintile	SEP quintile				
	Poor	2	3	4	Rich
↓ Parent’s cognitive ability quintile, age 10					
Low	40	23	18	15	4
2	25	21	22	23	10
3	18	23	21	20	17
4	17	21	23	19	20
High	13	18	16	26	27

Furthermore, Table 3 shows that children growing up in high SEP families are far more likely to have cognitively able parents. This pattern is particularly striking at the bottom of the parental ability distribution: among children whose parents are in the bottom cognitive ability quintile, ten

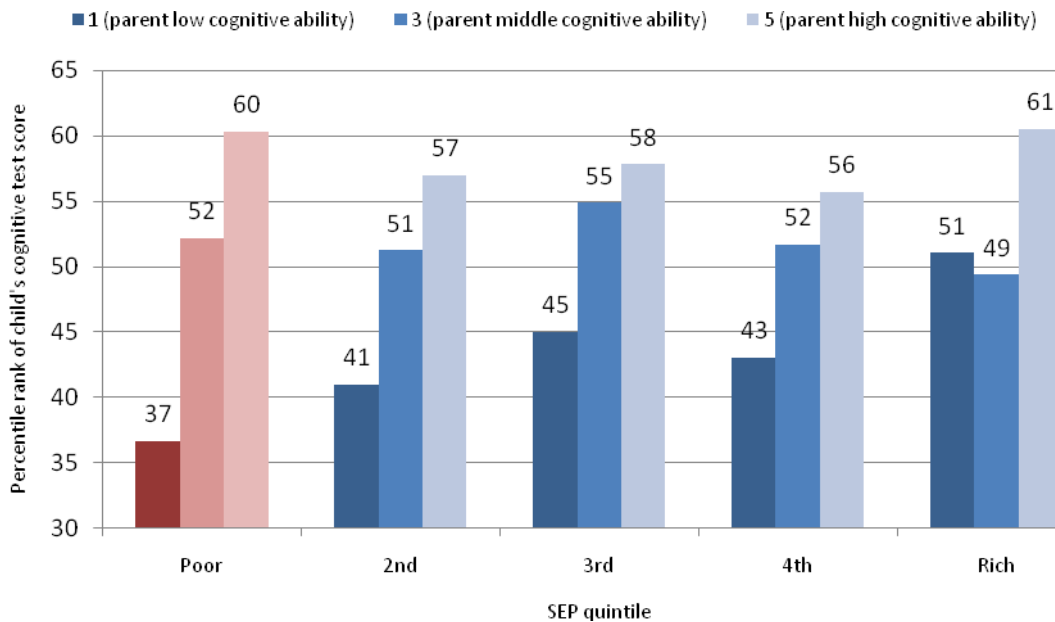
times as many are in the poorest SEP quintile (40%) as in the richest (4%). This suggests that, if parental ability is an important determinant of children’s abilities, observing SEP but not parental ability could lead to erroneous conclusions about the

determinants of cognitive skills (in the absence of a valid identification strategy).

Figure 1 allows us to disentangle the correlations documented in Tables 2 and 3, by showing children’s cognitive test scores in sub-groups defined by both parental ability and SEP. The gradients *across* SEP quintiles suggest that children whose parents are of low ability tend to do significantly better in cognitive tests if they are in a high SEP group (an average percentile rank of 51 in

cognitive test scores for those in the top SEP quintile, compared to 37 in the bottom SEP quintile); however, there is no discernible SEP gradient for those whose parents are of middle and high ability. This could be because we only observe the cognitive ability of one parent. Low ability CMs who are in a high SEP group may be relatively likely to have high ability partners (which is unobservable in our data), and that may explain the better cognitive ability of their children.

**Figure 1. Children’s cognitive test scores by SEP and parental ability**



**Notes:** Authors’ calculation using British Cohort Study. Our sample includes all children of the cohort members of the BCS aged between 3 and 16 in 2004 for whom we observe a BAS score. Parental ability refers to cognitive ability as measured when the parents were 10 years old, using BAS as well as additional tests of reading, vocabulary, writing, spelling, maths, copying, sentence formation and sequence recognition.

The clear upward gradient *within* SEP quintiles in Figure 1 shows that the correlation between parental abilities and children’s abilities remains very strong after conditioning on SEP. Unless this conditional correlation is entirely explained by other observables (such as parental education), this implies that observing SEP but not parental ability would result in an over-estimate of the importance of SEP for cognitive skills (given the association between parental ability and SEP shown in Table 2). In the next section we investigate whether the association between SEP and/or parental ability and child cognitive skills holds once we control for a large set of additional covariates.

#### 4. Adding more parental characteristics to the model: what difference does it make?

As explained, the fact that (unlike most studies of children’s cognitive skills in the UK) we can account for detailed characteristics of one of the children’s parents is of great significance. There is clear potential for these characteristics to be correlated both with our outcome of interest (the cognitive ability of children) and with typically observable explanatory variables (such as SEP and parental education). Therefore, when such characteristics are unobserved, they are likely to confound analyses that try to explain differences in



cognitive skills amongst children. This has three particular consequences of interest here:

- i. The role of parents’ SEP in ‘explaining’ children’s cognitive outcomes may be overstated: the correlation between the two may be due to the fact that high ability parents raise high ability children, and would do so irrespective of their socio-economic status.
- ii. When looking for possible transmission mechanisms, the extent to which the SEP gap in children’s cognitive skills remains ‘unexplained’ by observable characteristics may also be overstated: controlling for parental ability and other information about the parent might allow us to ‘explain away’ more of the raw SEP gap. (This ‘unexplained’ component is sometimes interpreted as the ‘direct effect’ of SEP on cognitive ability.)
- iii. The role of other factors in explaining the raw SEP gap may also be mis-stated: for example, the role of parental ability (or other parental characteristics), when it is not observed, may get attributed to observed factors with which it is correlated (e.g. parental education, attitudes, the home learning environment, and so on).

What difference does the additional information we have about the children’s parents make to the kind of analysis of cognitive skills presented in the companion papers in this Special Issue? Here, we compare the results we get from the BCS data with and without using the information on parental

cognitive ability, social skills and other characteristics that is not typically observed.

To do this we run two series of regressions, with children’s cognitive test scores (age-normalised and in percentile ranks) as the dependent variable. In the first of these (shown in the upper panel of Table 4), we obtain the ‘raw’ SEP gap in children’s test scores (shown in column 1) and sequentially add to the model parental education (column 4), demographics and other family background (column 5), and attitudes, educational aspirations, behaviours and the home-learning environment (column 6). These regressions are similar to those run in the other papers in this volume, which use information from a single generation only. The second series of regressions (shown in the lower panel of Table 4) differs only in that additional information from the parent’s childhoods is also added to the model, *before the other covariates* (in columns 2 and 3). The difference between the two series of regressions is thus that the first one does not include columns 2 and 3: detailed information from the parent’s childhood is never added to the model.

Column Regressors added to the model:

- 1) Current SEP quintile.
- 2) Parental cognitive ability.
- 3) Parental attitudes and social skills.
- 4) Parental education.
- 5) Demographics and family background.
- 6) Attitudes, educational aspirations, behaviours and the home-learning environment.

**Table 4. Children’s cognitive test scores: regression results<sup>13</sup>**

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
<b>Using information from current generation only</b>						
Current SEP quintile = 2	3.4*	-	-	3.2*	2.4	0.8
Current SEP quintile = 3	4.0**	-	-	3.8**	3.5*	1.3
Current SEP quintile = 4	6.4***	-	-	6.1***	5.1**	2.8
<b>Current SEP quintile = 5 (richest)</b>	8.7***	-	-	7.3***	6.6***	3.4
Age mother left FT education: 17-18 (relative to 16 or younger)	-	-	-	2.4*	3.8***	2.2*
Age mother left FT education: 19-22 (relative to 16 or younger)	-	-	-	4.2**	6.4***	4.3***
Age mother left FT education: 23-25 (relative to 16 or younger)	-	-	-	3.6	6.2**	4.4
Age mother left FT education: 26 + (relative to 16 or younger)	-	-	-	9.0***	9.1***	3.6
Father is self-employed (relative to full-time employee)	-	-	-	-	-2.4	-2.9**
Age of mother at birth of child: 20-24 (relative to under 20)	-	-	-	-	1.4	5.4**
Age of mother at birth of child: 25-29 (relative to under 20)	-	-	-	-	-14.2***	5.5*
Age of mother at birth of child: 30-34 (relative to under 20)	-	-	-	-	-18.3***	6.6**
1 older siblings	-	-	-	-	-6.3***	-3.8***
2 older siblings	-	-	-	-	-9.1***	-6.2***
3 or more older siblings	-	-	-	-	-13.2***	-7.3**
Twin	-	-	-	-	-13.9***	-14.1***

(Table 4 cont'd)

Breastfed for 6 months or more (relative to not breastfed)	-	-	-	-	-	4.4***
Attends or attended playgroup	-	-	-	-	-	2.4**
Parent thinks child likely to go to university or college	-	-	-	-	-	13.2***
SDQ score (standardised)	-	-	-	-	-	5.2***
Parental discipline score (standardised)	-	-	-	-	-	2.8**
Parent-child conflict score (standardised, high score means more conflict)	-	-	-	-	-	1.9**
Home-learning environment score for under-6s (standardised)	-	-	-	-	-	2.9**
Harter perceived competence score (standardised, 10-16s only)	-	-	-	-	-	-3.7***
CM speaks to teacher about child at least once per term	-	-	-	-	-	-3.4**
Child goes to private school	-	-	-	-	-	7.3**
Child can use a computer at home for homework (6-16-s only)	-	-	-	-	-	4.9**
Child reads for enjoyment several times a week (6-16s only)	-	-	-	-	-	9.6***
Child wants to continue studying post-16 (10-16s only)	-	-	-	-	-	8.5***
Child has smoked (10-16s only)	-	-	-	-	-	-6.5**
Child (at least) sometimes drinks alcohol (10-16s only)	-	-	-	-	-	5.6**

**Including information about parent's childhood (including cognitive ability)**

Current SEP quintile = 2	3.4*	1.6	1.5	1.6	1.3	0.2
Current SEP quintile = 3	4.0**	2.4	2.2	2.4	2.5	0.7
Current SEP quintile = 4	6.4***	4.3**	3.9***	4.2**	3.6*	2.1
<b>Current SEP quintile = 5 (richest)</b>	<b>8.7***</b>	<b>4.8**</b>	<b>3.8**</b>	<b>3.7*</b>	<b>3.5</b>	<b>1.5</b>
Age mother left FT education: 17-18 (relative to 16 or younger)	-	-	-	1.4	2.9**	1.8
Age mother left FT education: 19-22 (relative to 16 or younger)	-	-	-	2.2	4.5**	3.2*
Age mother left FT education: 23-25 (relative to 16 or younger)	-	-	-	0.9	3.7	3.1
Age mother left FT education: 26 + (relative to 16 or younger)	-	-	-	7.4***	7.2***	2.8
Father is self-employed (relative to full-time employee)	-	-	-	-	-2.5*	-3.2**
Age of mother at birth of child: 20-24 (relative to under 20)	-	-	-	-	0.3	4.6*
Age of mother at birth of child: 25-29 (relative to under 20)	-	-	-	-	-15.7***	5.0
Age of mother at birth of child: 30-34 (relative to under 20)	-	-	-	-	-20.1***	5.8*
1 older sibling	-	-	-	-	-5.7***	-3.2***
2 older siblings	-	-	-	-	-7.8***	-5.0***
3 or more older siblings	-	-	-	-	-11.3***	-5.7
Twin	-	-	-	-	-13.5***	-13.8***
Breastfed for 6 months or more (relative to not breastfed)	-	-	-	-	-	3.1**
Attends or attended playgroup	-	-	-	-	-	2.4**
Parent thinks child likely to go to university or college	-	-	-	-	-	13.4***
SDQ score (standardised)	-	-	-	-	-	4.7***
Parental discipline score (standardised)	-	-	-	-	-	2.6**
Parent-child conflict score (standardised, high score means more conflict)	-	-	-	-	-	1.9**
Parent-child closeness score (standardised)	-	-	-	-	-	-0.7
Home-learning environment score for under-6s (standardised)	-	-	-	-	-	2.7*
Harter perceived competence score (standardised, 10-16s only)	-	-	-	-	-	-3.4***
CM speaks to teacher about child at least once per term	-	-	-	-	-	-3.6***
Child goes to private school	-	-	-	-	-	8.3**
Child can use a computer at home for homework (6-16-s only)	-	-	-	-	-	4.3**
Child reads for enjoyment several times a week (6-16s only)	-	-	-	-	-	10.0***
Child wants to continue studying post-16 (10-16s only)	-	-	-	-	-	8.3***
Child has smoked (10-16s only)	-	-	-	-	-	-6.9**
Child (at least) sometimes drinks alcohol (10-16s only)	-	-	-	-	-	5.6**
CM thought they were good at maths, age 16	-	-	4.0**	4.4**	4.5**	4.2**
CM's quintile in cognitive test scores, age 5 = 2	-	3.2	3.1	2.8	3.7*	3.5**
CM's quintile in cognitive test scores, age 5 = 3	-	7.3***	6.8***	6.5***	7.3***	6.1***
CM's quintile in cognitive test scores, age 5 = 4	-	2.3	2.0	1.4	2.6	1.7
<b>CM's quintile in cognitive test scores, age 5 = 5 (brightest)</b>	<b>-</b>	<b>7.2***</b>	<b>6.6***</b>	<b>6.2***</b>	<b>7.0***</b>	<b>4.8***</b>
CM's quintile in cognitive test scores, age 10 = 2	-	3.9*	3.5*	3.3	2.7	3.3*
CM's quintile in cognitive test scores, age 10 = 3	-	7.9***	7.3***	7.1***	6.4***	6.1***
CM's quintile in cognitive test scores, age 10 = 4	-	10.3***	9.2***	9.1***	9.0***	9.2***
<b>CM's quintile in cognitive test scores, age 10 = 5 (brightest)</b>	<b>-</b>	<b>12.7***</b>	<b>10.8***</b>	<b>10.3***</b>	<b>10.8***</b>	<b>10.6***</b>

**Notes to Table 4:** Authors' calculations using the British Cohort Study (BCS). Our sample includes all children of the cohort members of the BCS aged between 3 and 16 in 2004 for whom we observe a BAS score. Results reported are from a series of OLS regressions, robust to within-family clustering. One, two and three asterisks indicate statistical significance at the 10%, 5% and 1% level respectively. The covariates not listed here but that are included in the model (see Supplementary files for all coefficient estimates) are: father's education; the other behavioural and attitudinal information about the cohort member (i.e. one of the parents) as a child described in Section 2; social skills as measured by the Strengths and Difficulties Questionnaire; height and weight of the mother at birth; gender; ethnicity; whether a lone parent family; current health status of cohort member ('poor' to 'excellent'); whether cohort member currently has a long-term limiting illness; employment status of both the mother and father; age of mother at child's birth; number of older and younger siblings; whether a twin; whether attended nursery or playgroup; the home-learning environment (see Supplementary Files for details); whether child thinks good marks in education are important; whether they and their parent (the cohort member) want the child to continue in post-compulsory education; whether they and their parent think it is likely that the child will continue in post-compulsory education; whether they read for enjoyment; whether they have been suspended from school, smoked, broken the law, used cannabis, stolen from a store, bullied other children, been bullied, and been to youth clubs, scouts/guides or sports clubs/lessons at least once a week; whether school is private or single sex; whether parent (cohort member) thinks teaching at the school is 'very good'; whether child likes most of their teachers; levels of parental discipline, parent-child closeness and parent-child conflict (see Supplementary Files for details); whether the parent (cohort member) feels they spend enough time with their children; whether the family eats together every day; number of activities done as a family in last month; whether breastfed and how long for.

In Table 4 we present the coefficients for SEP and parental cognitive ability, as well as those coefficients that were statistically significant at the 5% level or above in the first series of regressions (without controlling for parental ability, social skills and attitudes) so that we can see whether they remain significant after controlling for the additional information that we observe about the parents.<sup>14</sup> It is interesting to note that parental cognitive ability is the only factor observed during the parent's childhood which is statistically significant at the 1% level; conditional on other characteristics, social skills and attitudes and behaviours of the parents as children are not statistically significant (though they may have an effect that is mediated through the social skills, attitudes or other characteristics of their children).

Column 1 in Table 4 shows that children in the top SEP quintile are 9 percentiles higher up the cognitive test score distribution than those in the bottom SEP quintile, on average. This is the 'raw' SEP gap. This 'raw' gap is noticeably smaller than that present in the datasets analysed in the companion papers in this Special Issue. This may be because our sample is more homogeneous than those examined in the other papers, for at least two reasons: 1) all children in our estimation sample have at least one parent who was aged 31 or younger at the time of their birth; 2) the BCS is, by some distance, the oldest of the cohort studies considered, thus there has been considerably more time for attrition to occur since the first wave of analysis in 1970.

Columns 4-6 of the upper panel show how the residual SEP gap is reduced by controlling for the

kinds of factors observed in the other papers in this Special Issue: parental education, demographics and other family background characteristics, and attitudes, educational aspirations, behaviours and the home-learning environment. Once we account for these, the residual SEP gap in cognitive skills is not statistically significant. This does not necessarily diminish the importance of SEP for cognitive development, as many of the factors we control for are plausible transmission mechanisms between SEP and cognitive skills.

If instead we control first for parental ability (column 2 of the lower panel), the residual SEP gap is reduced by 3.9 percentiles; controlling additionally for attitudes and social skills when the parents were children (column 3) reduces the residual gap by a further 1.0 percentile.<sup>15</sup> This has the interesting implication that more than half of the raw SEP gap in children's cognitive skills, expressed in percentile ranks, can be predicted by one parent's cognitive ability and social skills, observed by the time that parent was 16 years old. Column 6 shows that, after adding all the controls, the residual gap between the top and bottom SEP quintiles is more than twice as large when we do not account for parental ability, attitudes and social skills (although this gap is not statistically significant in any case).

It is striking that, in contrast to SEP, the apparent importance of parental cognitive ability is relatively insensitive to the number of factors we control for (as shown by the coefficients on parental ability at the bottom of Table 4). We cannot rule out the possibility that there are unobserved covariates which would change this

story if included in our model. However, given the very rich set of information we do observe, the results are at least suggestive of the fact that parental ability plays some kind of role in determining children's cognitive skills.

Reassuringly though, almost all of the statistically significant predictors of cognitive skills in the upper panel of Table 4 remain significant, and similar in magnitude, in the lower panel. This is (tentative) evidence that an inability to observe parental cognitive ability and other detailed information about the parents (as in the companion papers in this Special Issue, for example) may not grossly distort conclusions about which observable factors from the current generation are the most important predictors of cognitive skills. These factors include family structure (in particular, having older siblings and being a twin, strongly predict worse cognitive skills), the child's social skills and attitudes of both the parent and the child towards education.

To summarise, Table 4 suggests that, when *unable* to observe the detailed information about the children's parents that datasets like the BCS provide:

- i) the (conditional) correlation between SEP and cognitive outcomes overstates the causal impact of SEP on cognitive outcomes;
- ii) the 'unexplained' SEP gap is too large (although not statistically significant in our case), and an important part of the explanation for the raw SEP gap is therefore missed; but
- iii) an assessment of which *observed* factors are the most important predictors of cognitive skills appears fairly robust to this omitted variable problem.

We investigate this further in the next section by quantifying the role of different factors in

'explaining' the SEP gap using decomposition analysis, both with and without accounting for the information we have from the BCS about the parents as children.

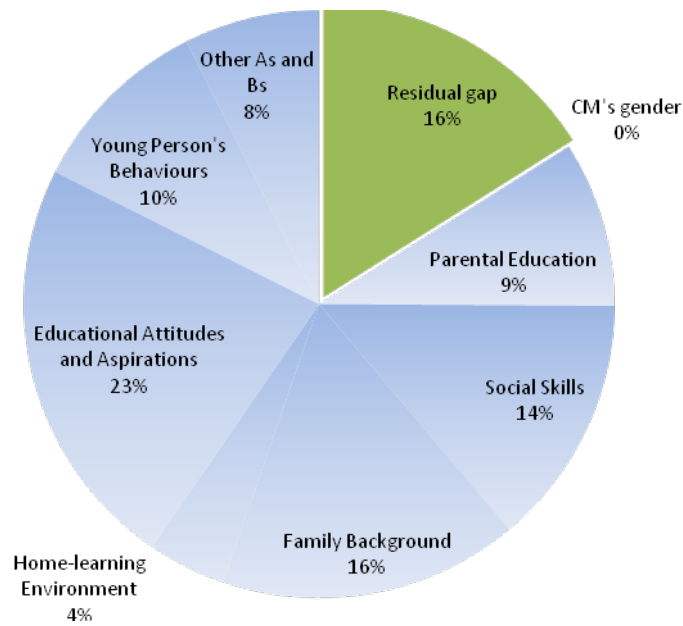
## 5. Decomposing the SEP gap in children's cognitive test scores

In this section, we compare in more detail the results from our final specification (that shown in column 6 in Table 4) obtained with and without controlling for parental ability, attitudes and social skills (as observed when the parents were children). To do this, we present results from two very simple decomposition analyses in Figure 2. The decompositions show the fraction of the 'raw' cognitive gap between the top and bottom SEP quintiles that is attributable to other factors that we can observe. In the upper decomposition, these 'other factors' do not include parental ability, attitudes and social skills, while in the lower composition, these additional factors are included.

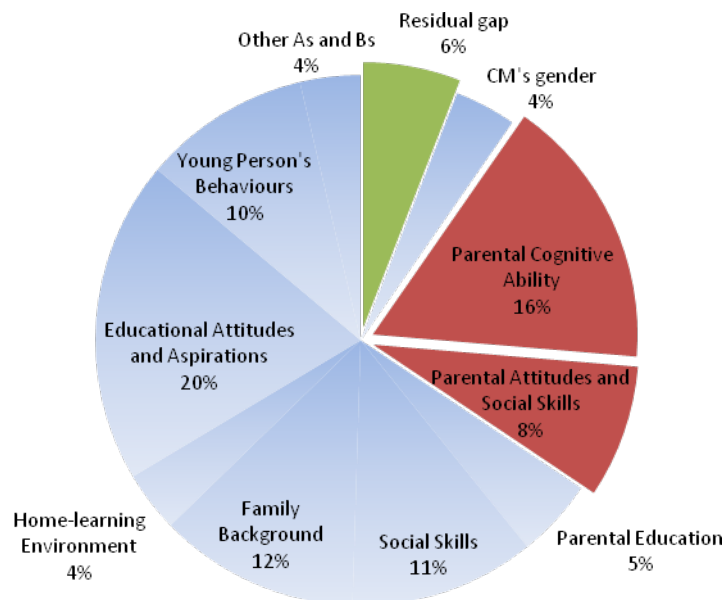
The contribution of each variable to the SEP gap is given by its coefficient estimate from the full regression specification multiplied by the extent to which it varies with SEP (the difference between the mean values of the variable in the top and bottom SEP quintiles).<sup>16</sup> See Goodman, Gregg and Washbrook 2010 for more details. The variables have been grouped thematically for the purposes of the pie charts in Figure 2. The full list of variables in each group, along with coefficient estimates, conditional means and contributions to the raw SEP gap for the lower decomposition (including characteristics from the parents' own childhood) can be found in our Supplementary Files.

Figure 2. The SEP gap in cognitive test scores: decomposition analysis

Excluding parental cognitive ability, attitudes and social skills



Including parental cognitive ability, attitudes and social skills



**Notes:** Authors' calculations using the British Cohort Study (BCS). Our sample includes all children of the cohort members of the BCS aged between 3 and 16 in 2004 for whom we observe a BAS score. The contribution of each variable to the SEP gap is given by the size of its conditional correlation with cognitive test scores (its coefficient estimate from the full regression specification) multiplied by the extent to which it varies with SEP (the difference between the mean values of the variable in the top and bottom SEP quintiles). Parental cognitive ability and parental attitudes and social skills were measured during the parent's childhood. Statistical significance of the coefficient estimates is not taken into account.

Note that we exclude ‘missing data’ as a category in the decompositions, so the contribution of each group of variables in driving the SEP gap is presented as a percentage of the total contribution of all *non-missing* variables. This is because some information about the children of the BCS is observed only for children of certain ages (for example, self-reported attitudes and behaviours are observed only for those aged 10 or over), and is therefore subject to a lot of missing data when we group all children together in the analysis. This highlights that the decision to group children of all ages together is not trivial, in principle; however, the main conclusions we reach here are robust to disaggregating our sample of children by age.

The upper pie in Figure 2 shows that, of the variables from the current generation, the most important in accounting for the SEP gap in children’s cognitive test scores are educational attitudes and aspirations (23%), family background (16%) and social skills (14%). After accounting for everything that we observe in the current generation, 16% of the SEP gap remains unexplained.

The lower pie shows the importance of the additional information we have about the children’s parents. These factors (shown in red) account for nearly one quarter of the raw SEP gap. Parental cognitive ability is by far the most important of these (16%). Most noticeably, this reduces the unexplained component of the SEP gap to just 6%. However, it does not change the relative importance of the factors observed during the current generation – educational attitudes and aspirations, family background and social skills are still the most important of these – which, if our results have external validity, is reassuring for previous studies unable to account for parental ability.

Note that the *indirect* effects of parental cognitive ability may be very large. For example, more able parents may stay in education longer and have more positive attitudes towards the education of their children. The significant contribution of parental cognitive ability in Figure 2 is net of any indirect effects via the many covariates that we control for, such as the parents’ educational outcomes and attitudes. In a similar decomposition based on a model including only SEP and parental cognitive ability, parental cognitive ability accounts for 50% of the gap in cognitive test scores between children from the richest and poorest SEP

backgrounds. It thus seems likely that the true contribution of parental ability in explaining the SEP gap in cognitive skills, accounting for both direct and indirect effects, lies somewhere in the range 16-50%.

In addition, we could conjecture that higher parental ability increases children’s cognitive skills by making the time spent on children’s cognitive development more productive (e.g. reading with one’s children may be more productive if the parent is a good reader). However, the importance of parental cognitive ability does not seem to be largely driven by a complementarity with observed aspects of parenting, in as much as most interaction terms between parental cognitive ability and environmental factors are not statistically significant when added to the model (not shown).<sup>17</sup> Thus, although our findings are based on a simple linear regression model in which endogeneity cannot be ruled out, we interpret the large contribution of parental cognitive ability in Figure 2 – which remains even after controlling for a very wide range of plausible transmission mechanisms between parental and child cognitive ability – as being suggestive of the possibility of some kind of genetic link between the cognitive skills of parents and their children, although we acknowledge that the interactions between genetics and the environment are necessarily complex.

Finally, since we only have detailed information about one of the parents (the cohort member) of the children in our sample, assortative mating (the tendency for people to have partners with characteristics similar to themselves) may obscure one reason why cleverer parents raise cleverer children – namely, that they have cleverer partners. Since we observe many *family*-level environmental factors (such as the home-learning environment), we reduce the scope for us to confound the importance of parental ability with the importance of the role played by the CM’s partner in providing a home environment conducive to cognitive development. However, we are not able to fully account for the likelihood that high cognitive ability of the CM is associated with high cognitive ability of the CM’s partner (we may partially account for it because we observe the education levels of *both* parents, which should proxy cognitive ability). We are therefore likely to be picking up the combined effect of some kind of genetic transmission from both parents.

## 6. Conclusion

Studies of children's cognitive abilities typically relate them to the cognitive abilities of their parents or to socio-economic and attitudinal factors (all of which tend to correlate). In this paper, we have made use of a dataset which unusually allows us to do both simultaneously. This has enabled us to consider the intergenerational transmission of cognitive abilities and the rich-poor gap in cognitive abilities as a coherent whole. In particular, it has allowed us to deal directly with a potentially serious omitted variables problem that could confound many studies of the relationship between childhood poverty and children's outcomes (as well as studies that seek to identify potential transmission mechanisms between the two).

In line with other studies, we have found that parental cognitive ability is a very important predictor of children's cognitive skills (the most important, along with educational attitudes and aspirations). Because parents' cognitive ability and socio-economic position are in turn also very strongly related, studies which examine the relationship between parental SEP and children's ability, but do not control for parental ability – or in some other way attempt to account for these factors – will suffer from a serious omitted variables problem. Indeed, our estimates suggest that parental cognitive ability accounts for 16 percent of the gap in cognitive test scores between children from rich and poor families after controlling for a wide range of mechanisms through which ability

may be transmitted across generations (for example, differences in the home learning environment that parents provide for their children), and 50 percent of the gap if we do not include such factors.

This finding raises important questions about the relevance of genetic inheritance in accounting for the gap in cognitive test scores between children from rich and poor families. This is clearly a controversial and complex topic. Our results do not suggest the strong complementarities between parental ability and other observed aspects of parenting that we had expected to find, and are tentatively suggestive of an important – albeit likely complex – role for genetic inheritance.

Interestingly, however, the inclusion of parental ability within our model does not substantively alter our impression of which of the other determinants of cognitive skills are important: for example, family structure, the child's social skills, and attitudes of both the child and the parent towards education are all important predictors of children's ability, even after taking parental ability into account. This is reassuring for studies such as the others in this volume, which are not able to control for parental ability, but find an important role for characteristics such as parental attitudes to education, the home learning environment, and children's behaviours and social skills as potential transmission mechanisms between parental SEP and children's cognitive outcomes.

## Acknowledgements

This work was carried out as part of a larger body of work funded by the Joseph Rowntree Foundation (JRF) through its programme of research on education and poverty. We gratefully acknowledge the support of JRF, and the generous and productive guidance of Helen Barnard and Chris Goulden. We are also thankful to members of the advisory group: Tim Crosier, Naomi Eisenstadt, Leon Feinstein, Zoe Ferguson, Tony Foot, Leslie Gutman, Lisa Harker, Andrew Ledger, Mark Newman, Ivan Turok, Anna Vignoles, Jane Waldfogel, and Stephen Witt. All errors remain the responsibility of the authors.

## References

- Anger S and Heineck G. (2010) Do Smart Parents Raise Smart Children? The Intergenerational Transmission of Cognitive Abilities. *Journal of Population Economics*, 23, 1255-1282.
- Björklund A, Hederöf Eriksson K and Jäntti M. (2010) IQ and family background: are associations strong or weak? *Berkeley Electronic Journal of Economic Analysis and Policy* 10.
- Black S, Devereux P and Salvanes K. (2009) Like father, like son? A note on the intergenerational transmission of IQ scores. *Economic Letters*, 105, 138–140.
- Blanden J, Gregg P and Macmillan L. (2007) Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education, *Economic Journal*, 117, C43-C60.
- Blau D. (1999) The Effect of Income on Child Development, *Review of Economics and Statistics*, 81, 261-276.

- Carneiro P, Crawford C and Goodman A. (2007) The impact of early cognitive and non-cognitive skills on later outcomes. *CEE Discussion Paper No. 92*.
- Conners CK. (1969) A teacher rating scale for use in drug studies. *American Journal of Psychiatry*, 126, 884-888.
- Cunha F and Heckman J. (2007) The technology of skill formation. *American Economic Review*, 97, 31-47.
- Dahl G and Lochner L. (2008) The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *NBER Working Paper No. 14599*.
- Davis-Kean P. (2005) The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19, 294-304.
- Dickens W and Flynn J. (2001) Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, 108, 346-369.
- Dickens W and Flynn J. (2002) The IQ paradox: Still resolved. *Psychological Review*, 109, 1-25.
- Duflo E. (2000) Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program. *American Economic Review*, 90, 393-398.
- Duncan G, Morris P and Rodrigues C. (2006) Does Money Really Matter? Estimating Impacts of Family Income on Young Children's Achievement with Data from Random-Assignment Experiments. [https://www.earlyadolescence.org/files/IOM/Does%20Money%20Really%20Matter\\_Restatrev\\_FINA\\_L\\_1\\_23\\_2006.pdf](https://www.earlyadolescence.org/files/IOM/Does%20Money%20Really%20Matter_Restatrev_FINA_L_1_23_2006.pdf)
- Duncan G, Young W, Brooks-Gunn J and Smith J. (1998) How much does childhood poverty affect the life chances of children? *American Sociological Review*, 63, 406-423.
- Duncan G and Brooks-Gunn J. (1997) *Consequences of Growing Up Poor*. Russell Sage Foundation, New York.
- Ermisch J, Francesconi M and Pevalin D. (2002) Childhood parental behaviour and young people's outcomes. *ISER Working Paper 2002-12*, University of Essex.
- Goodman A, Gregg P and Washbrook L. (2010) Children's educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood. *Longitudinal and Life Course Studies*, 2, 1-18.
- Guo G and Harris K. (2000) The mechanisms mediating the effects of poverty on children's intellectual development. *Demography*, 37, 431-447.
- Heckman J, Stixrud J and Ursua S. (2006) The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24, 411-482.
- Jean Yeung W, Linver M and Brooks-Gunn J. (2008) How money matters for young children's development: parental investment and family processes. *Child Development*, 73, 1861-1879.
- Levy D and Duncan G. (2000) Using Sibling Samples to Assess the Effect of Childhood Family Income on Completed Schooling. *JCPR Working Papers, Northwestern University/University of Chicago Joint Center for Poverty Research*.
- Loehlin J. (2002). The IQ paradox: Resolved? Still an open question. *Psychological Review*, 109, 754-758.
- Mayer S. (1997) *What Money Can't Buy: Family Income and Children's Life Chances*. Harvard University Press, Cambridge, MA.
- Rowe D and Rodgers J. (2002) Expanding variance and the case of historical changes in IQ means: A critique of Dickens and Flynn (2001). *Psychological Review*, 109, 759-763.
- Sacerdote B. (2002) The Nature and Nurture of Economic Outcomes. *American Economic Review*, 92, 344-348.
- Sacerdote B. (2007) How Large Are the Effects from Changes in Family Environment? A Study of Korean American Adoptees. *Quarterly Journal of Economics*, 122, 119-157.
- Shea J. (2000) Does parents' money matter? *Journal of Public Economics*, 77, 155-184.

## Endnotes

<sup>1</sup> Originally called the British Births Survey (BBS), the birth survey covered the whole of the United Kingdom, but those from Northern Ireland were dropped for subsequent sweeps.

<sup>2</sup> Children aged 10 or over were given an additional questionnaire of their own to complete.

<sup>3</sup> The British Ability Scales comprise a mixture of measures of educational attainment, including number skills, spelling and reading, and measures of cognitive abilities that are the outcome of interactions between a child's innate capabilities and his or her experiences, both at home and at school. For more details, see: [http://www.gla-assessment.co.uk/health\\_and\\_psychology/resources/british\\_ability\\_scales/british\\_ability\\_scales.asp?css=1](http://www.gla-assessment.co.uk/health_and_psychology/resources/british_ability_scales/british_ability_scales.asp?css=1).

<sup>4</sup> See <http://www.ifs.org.uk/publications/5269> for a full list of variables included in our analysis.

<sup>5</sup> See [www.cls.ioe.ac.uk/core/documents/download.asp?id=932&log\\_stat=1](http://www.cls.ioe.ac.uk/core/documents/download.asp?id=932&log_stat=1) for more details.



<sup>6</sup> There is also a myriad of other information about the parents' childhood that we could have included in our analysis (including SEP). However, the inclusion of these other factors did not materially affect our results, so for the sake of brevity, we focus our attention on cognitive ability, social skills and some selected attitudes to education and future careers.

<sup>7</sup> This means we take account of income in the age 29 wave as well as the age 34 wave.

<sup>8</sup> Precisely, all children have at least one parent who is 34.

<sup>9</sup> Indeed, the only directions of causality we could rule out *a priori* would be parental ability being affected by either of the other two (since we observe parental ability during the parent's childhood).

<sup>10</sup> Of course, parental ability may not be the only latent factor that can lead to omitted variable bias. The same could be true of other characteristics of the parents which we also (unusually) observe in the BCS and are able to account for in this paper, such as their attitudes or social skills. For example, parents that are more motivated may end up in higher paying jobs and may also provide greater stimulation for their children (leading to higher cognitive test scores).

<sup>11</sup> Note that SEP and parental cognitive ability are measured continuously but we have categorised them into quintiles here. This means, for example, that children in the same parental ability quintile who are in very different SEP quintiles are relatively likely to be far apart *within* the parental ability quintile (or vice versa). An association between parental ability and cognitive skills would then be obscured by our discrete categorisation, and would be attributed to SEP instead (or vice versa). Therefore, the precise gradients in Figure 1 depend upon how crudely we disaggregate the SEP/parental ability distributions, which is arbitrary. Our qualitative conclusions are robust to this choice, however: continuous measures of SEP and parental ability are both statistically significant predictors of cognitive ability when controlled for simultaneously in an OLS regression, but the p-value for parental ability is much lower than that for SEP; and the tendency for SEP to be more strongly associated with cognitive skills where parents are low ability is also statistically significant (confirmed by adding an interaction between SEP and an indicator variable equal to one if the parent is below the median of the ability distribution).

<sup>12</sup> The association between children's and parents' ability varies by gender, as found by Anger and Heineck (2010) using German data. It is strongest where the parent and child are the same sex (a correlation of 0.23 for fathers and sons and 0.22 for mothers and daughters), and lowest of all for fathers and daughters, with a correlation of 0.13. These correlations are quite low compared to others in the literature. For fathers and sons, Bjorklund et al 2010, obtain a correlation of 0.35 for Sweden, which closely matches estimates from Black et al 2009, for Norway. The Anger and Heineck study, looking at parents and children of all genders, found a correlation of 0.5. Possible explanations for these different correlation coefficients include: different measures of cognitive ability; data from different countries and different cohorts (for example, the parents in our sample were born nearly 40 years later than those used by Black et al); and non-random sample selection in our study and (to varying extents) the other studies.

<sup>13</sup> We have replicated this regression with SEP quintiles replaced by income quintiles, to test the sensitivity of our results to our measure of material wellbeing. The coefficients on income are notably smaller than those on SEP (the raw gap between the top and bottom quintiles is 6.2 percentile ranks, rather than 8.7) – this could reflect measurement error in income - but the raw gaps are still statistically significant at the 1% level. Coefficient estimates on other variables are virtually unaffected.

<sup>14</sup> All coefficient estimates from column 6 of the lower panel of Table 4 can be found in our Supplementary Files: a full set of results is available from the authors on request.

<sup>15</sup> As explained, the variables that we loosely refer to as parental characteristics are more precisely the characteristics of *one* of the children's parents (the cohort member). Exceptions are education and employment status, which we observe for both parents.

<sup>16</sup> Note that the 'contribution' of a variable to the SEP gap says nothing about statistical significance: it depends on the magnitude of estimated coefficients but not the precision with which they are estimated.

<sup>17</sup> Specifically, we created an indicator variable that equals one if the parent was in the top half of the cognitive ability distribution at age 10, and interacted this with variables capturing: parental discipline, parent-child closeness, parent-child conflict, a composite home learning environment score, the number of activities done together as a family per week, whether the parent reads to the child daily, whether the parent regularly attends parents' evenings, whether the parent feels they spend enough time with their children, whether the parent helps with homework and whether the parent makes sure homework is completed. These interactions are jointly insignificant at the 10% level.

# NEWS, EVENTS AND RESOURCES

## News and Resources

### The Society for Longitudinal and Life Course Studies

In September 2009 a new international learned society was established by the think tank Longview - the Society for Longitudinal and Lifecourse Studies, to which Longview has handed over responsibility for this journal.

The Society is multi-disciplinary in nature, bringing together members across the health, behavioural and social sciences who have a shared interest in longitudinal and life course research. In September 2010 the Society held an extremely successful inaugural conference at Clare College, Cambridge. The conference provided the attendees with an opportunity to meet and to pool and exchange ideas. The Society is now planning its second annual conference, which will be hosted by the University of Bielefeld, Germany on 26-28 September 2011.

The Society currently has nearly 200 members worldwide, and is actively seeking to enlist more. Membership of the Society offers opportunities to:

- make collaborative contacts throughout the global longitudinal research community
- contribute papers and symposia to the Society's annual conference, and attend the conference for a reduced fee
- make nominations, be nominated and vote in elections for the future President and Executive Committee\* of the Society who are also the Society's trustees
- participate in online policy debates and forums
- access capacity building initiatives, including our annual summer school, online master classes and methodological and longitudinal practice workshops
- automatic free registration to receive our peer-reviewed international journal, Longitudinal and Life Course Studies

You may like to consider individual membership or, if your colleagues also have an interest in longitudinal and life course research, corporate membership. There is also a reduced membership fee available for students.

For full information on the Society's activities and for the online application form to join, please visit the Society website at [www.slls.org.uk](http://www.slls.org.uk)



\*The Society's Executive is: **Heather Joshi**, President (Institute of Education, London), **Tom Jupp**, Hon Treasurer (Institute of Education, London), **Elizabeth Webb**, Hon Secretary (Imperial College, London); Committee Members: **David Blane** (Imperial College London), **John Bynner** (Longview), **Elizabeth Cooksey** (Ohio State University), **Walter Heinz** (University of Bremen), **Amanda Sacker** (University of Essex), **Ingrid Schoon** (Institute of Education, London), **Tom Schuller** (Longview).

### Understanding Society, the UK Household Longitudinal Study



*Understanding Society* is the successor household panel study to the British Household Panel Survey (BHPS) which ran for eighteen years from 1991 to 2008. *Understanding Society*, funded by the Economic and Social Research Council (ESRC) and a consortium of government departments, is designed to provide high quality longitudinal data to answer research and policy needs over the coming decades and represents a significant investment in the social sciences in the UK. Managed by the same team responsible for the BHPS at the Institute for Social and Economic Research (ISER), University of Essex, the study is being conducted to the highest standards of best practice within survey methodology and the methodology of conducting longitudinal surveys. Fieldwork is conducted by The National Centre for Social Research (NatCen).

The future of this prestigious and important study was secured with the refunding of *Understanding Society* in 2010 so that data collection up to and including wave 5 is now in place. In February 2011, a volume of first findings using data from the first year of the study will be published, promising just a taste of the rich and varied analysis that is possible using *Understanding Society* data. A Special Issue of the LLCS journal

# NEWS, EVENTS AND RESOURCES

devoted to longitudinal analysis using *Understanding Society* and the BHPS is planned within the next twelve months, marking the start of what we are sure will be a wealth of interesting and important substantive research using the data.

*Understanding Society* went into the field in January 2009 and began the third wave of data collection in January 2011. The key features of the design reflect its scientific rationale and are intended to generate major innovations in scientific research. The study is distinctive in having a:

- **Large sample size** of 40,000 households to allow more fine-grained analysis
- **Household focus** with data collected on all members aged 10 and over
- **Annual interview** enabling analysis of short-term and long-term dynamics of change
- **Full age range sample** complementing age-focused cohort studies in the UK and providing a unique look at behaviours and transitions in mid-life and throughout the life course
- **Innovation Panel** for methodological research and survey development
- **Multi-topic design** to meet a wide range of disciplinary and inter-disciplinary research needs
- **Ethnic minority research** agenda and inclusion of a boost sample of ethnic minority groups to support research on ethnic diversity and commonality
- **Collection of biomarkers and health indicators** as a resource for research at the interface of social and biomedical sciences
- **Data linkage to administrative records and geo-coded data** to provide significant new research opportunities where the survey data can be used in combination with administrative data
- **Inclusion of the BHPS sample** within *Understanding Society* from wave 2 to enable continuing longitudinal analysis of the BHPS sample alongside the new samples

The large sample size offers new opportunities to research sub-groups that may be too small for separate analysis on other studies, the UK-wide sample affords new avenues for comparative country, regional and geographic research, and the multi-topic design encourages inter-disciplinary approaches. The Innovation Panel is proving an invaluable resource for methodological research

pushing the boundaries of knowledge within longitudinal survey methodology. An annual competition for experiments to be carried on the Innovation Panel has been established and is advertised on the [Understanding Society](#) website.

Two areas of emphasis are support for research on ethnicity and identity, and health and bio-measure collection.

## Ethnic Minority Research

There has been no dedicated national survey of Britain's main ethnic minority groups since 1994, and there has never been a panel survey in which substantial numbers of minority group members have been followed from year to year. Ethnicity has therefore been a core element of *Understanding Society* from the start. Ethnic diversity and disadvantage, and issues of adaptation, opportunity and inclusion, are of wide general interest and are highly salient for research and policy purposes. The ethnic minority boost sample provides a sufficiently large sample of each of five key ethnic groups to allow these groups to be analysed separately (and compared with each other). The boost is designed to achieve an additional 1,000 individual interviews from members of Indian, Pakistani, Bangladeshi, Caribbean and Black African groups. All those of mixed background are included as are other minority groups including Chinese, Turkish, other Asian and Middle Eastern.

## Bio-measure collection

*Understanding Society* is a bio-social survey providing data that will support biomedical and social science research. The addition of bio-measures permits the examination of objective biological, anthropometric and functional measures within a large sample that spans many ages and which can be studied in a household context. The study provides information about social and economic factors that influence health status and the trajectory of health outcomes. The data will enable analysts to assess exposure to and antecedent factors of people's current health status, give a better understanding of disease mechanisms such as gene-environment interactions, allow an assessment of household and socio-economic effects on health, as well as analyses of outcomes using direct assessments. The collection of bio-measures began during wave 2 (2010/11) and will continue throughout wave 3 (2011/12). An estimated 25,000 to 30,000 adults will have these measures.

# NEWS, EVENTS AND RESOURCES

Two types of bio-measures are being collected. The first are direct measures including height and weight, waist circumference, bio-electrical impedance, grip strength, blood pressure and pulse rate and lung function. Secondly, biological samples including whole blood, saliva, and dried blood spots. In addition to the bio-measures, wave 3 is collecting cognitive ability/functioning measures for the whole sample. Combined with the questionnaire data, these data will provide a unique resource for inter-disciplinary research across the social and medical sciences.

## Data availability and further information

[Data](#) from the first year of wave 1 of *Understanding Society* is now available. An interim set of data (14,000 of the planned 40,000 households) has been deposited with the [Economic and Social Data Service](#). The full release of wave 1 data is scheduled for October 2011.

A twice-a-year *Understanding Society* e-bulletin giving updates on data release, latest research findings, training events, conferences and workshops that would be of interest to academics undertaking longitudinal research is available. Email [cgarr@essex.ac.uk](mailto:cgarr@essex.ac.uk) if you would like to receive this.

**Click [here](#) for further information about the design and content of the study.**

## Longview activities

Readers may or may not know that this journal originated from discussions within Longview, a think-tank which exists to promote longitudinal and lifecourse studies. Here I give a brief account of the issues and activities in which Longview engages. As a small organisation, we have been very keen to focus our efforts on where we can add value in a way that other organisations cannot.

- 1. Launching new initiatives.** Longview has been behind the launch of two new initiatives in the field of longitudinal and lifecourse studies: this journal, and the Society for Longitudinal and Lifecourse Studies, closely associated with it. The SLLS held its inaugural international conference in September 2010 in Cambridge UK; this was a very successful event, with around 160 participants from all over the world, and some very lively exchanges. The conference abstracts are collected in the Supplement to this LLCS issue.
- 2. Making the political case.** We have been concerned to make the case for the value

of longitudinal studies (LS), publicly and privately. This is particularly important at a time of fiscal constraint in most countries. In the UK, where such constraints are particularly severe, we had a meeting with David Willetts, Minister for Science, to press the case. We argued especially for stronger mechanisms for making effective use of existing studies. A paper on the economic arguments for LS prepared for that meeting is available [here](#). We also put the case for strengthening international links on LS. It would be interesting to know of equivalent exchanges in other countries.

- 3. Visualisation.** A strand of work which has aroused considerable interest is that of the visualisation of data. The goal here is to develop ways of dynamically representing, in accessible and interesting formats, data and interpretations which track populations and individuals over time. My sense, confirmed by a wide range of people with whom I have spoken on this, is that better and more imaginative visualisation techniques could play an enormously valuable role in improving communication at many different levels. This would include communication between researchers from different disciplines and methodologies. The Gapminder approach of Hans Roslin in Sweden is a fine example, though clearly we would need to develop a broader approach in order to cope with a wider range of research. Ideally the visualisation would be of results and interpretations as well as data.
- 4. Solving the Problem of Attrition.** This three year project funded by the UK Economic and Social Research Council is in pursuit of the Longview aim of linking findings of methodological research directly to the improvement of survey practice. It is a two stage project, involving secondary analysis of response data from a number of major UK longitudinal surveys, followed by a field experiment to test hypotheses about the causes of attrition, based on the first stage findings. Interviewer experience, quality and continuity from one wave to the next were targeted. The project is now in a dissemination and reporting phase, to be completed before the next (April 2011) Issue of the journal, which as it happens is a Special Issue devoted to "Attrition

# NEWS, EVENTS AND RESOURCES

---

in US and UK Health and Ageing longitudinal studies". Watch this space and the Longview website!

5. **Events.** As part of our promotion of LS, Longview is mounting a number of events in the UK; we would be most interested in hearing about similar events elsewhere.

- Firstly, we are planning a series of policy-relevant seminars designed to engage policy-makers fully, as distinct from transmitting results to them. The initial two or three will be on aspects of ageing, one in collaboration with the Personal Social Services Research Unit at the London School of Economics
- Secondly, we are collaborating with the British Library and the UK Academy of Social Sciences on a series of events to be held at the British Library in spring and summer 2011. The series is designed in tandem with the BL's exhibition on the Census, and contains events which will appeal to the public and to intending or more experienced researchers
- Thirdly, Longview will host the annual Neville Butler Memorial Lecture in May 2011. We are delighted that this year it will be given by the Rt Hon David Willetts, UK Minister of Science. He is a politician well-known for his genuine commitment to using evidence in policy, and will be speaking on this topic

Longview's research networks are all strongly linked internationally. The field of longitudinal and life course studies has been enormously enriched by comparative studies, and by the exchange of results and approaches across boundaries. Longview

itself as a body is perhaps rather unusual. For that reason we would welcome information from the Journal's readers on equivalent bodies in other countries, with which contact could be established. For more information (or to read my blog) go to [www.longviewuk.com](http://www.longviewuk.com).

**Tom Schuller**, Director  
[tomschuller@longviewuk.com](mailto:tomschuller@longviewuk.com)

## Québec Longitudinal Study of Child Development – publication

The Institut de la statistique du Québec published in December 2010, the 'fascicle' entitled *Diverse and Changing Family Structure During Early Childhood*. This [publication](#) is based on data from the Québec Longitudinal Study of Child Development (QLSCD 1998-2010). The fascicle covers family transitions experienced by the children and certain aspects of the family environment in which they are growing up, such as the atmosphere between the parents and living or custody arrangements if the parents are separated or divorced. The conclusion suggests further directions for intervention.

The Québec Longitudinal Study of Child Development (QLSCD) is being conducted on a representative sample of children born in Québec in the late 1990s. The main goal of this study is to identify the factors that, during early childhood, contribute to social adjustment and success in school.

Printed versions can be requested from the Centre d'information et de documentation of the Institut de la statistique du Québec by sending an email to [cid@stat.gouv.qc.ca](mailto:cid@stat.gouv.qc.ca).

---

## Events 2011

---

**SRCD** - (Society for Research in Child Development) Biennial Meeting **March 31<sup>st</sup> – April 2<sup>nd</sup>**, Montreal, Canada. <http://www.srcd.org>

**SLLS** - (Society for Longitudinal and Life Course Studies) Summer School, **July 4<sup>th</sup> – 8<sup>th</sup>**, University of Antwerp, Belgium, particularly for research post-graduates and post-doctoral fellows. Other sponsors include the European Association of Population Studies and CELLO (Centrum voor Longitudinaal en Levensloop Onderzoek - Research Centre for Longitudinal and Life Course Studies), University of Antwerp. Registration at [www.ua.ac.be/cello/summerschool](http://www.ua.ac.be/cello/summerschool).

**US/BHPS** – (*Understanding Society*/ British Household Panel Study) conference **June 30<sup>th</sup>-July 1<sup>st</sup>**, University of Essex, UK. A call for papers is now open - **deadline for abstracts is 15 March**. Details at <http://www.iser.essex.ac.uk/understanding-society-bhps-conference-2011>.