WIDER BENEFITS OF LEARNING RESEARCH REPORT NO.23

What role for the three Rs? Progress and attainment during primary school

Kathryn Duckworth

Centre for Research
on the Wider
Benefits of Learning

WHAT ROLE FOR THE THREE RS?

PROGRESS AND ATTAINMENT DURING PRIMARY SCHOOL

Kathryn Duckworth

SEPTEMBER 2007

Published by: Centre for Research on the Wider Benefits of Learning
Institute of Education, 20 Bedford Way, London WC1H 0AL
www.learningbenefits.net

ISBN: 978-0-9552810-3-7

Individual copy price: £10.00

The Department for Children, Schools and Families (DCSF) funded Centre for Research on the Wider Benefits of Learning (WBL) investigates the benefits that learning brings to the individual and to society as a whole. WBL's main objectives are to clarify, model and quantify the outcomes of all forms of intentional learning so as to inform the funding, implementation and practice of educational provision through the life course. The views that are expressed in this work are those of the author and do not necessarily reflect the views of DCSF. All errors and omissions are those of the author.

The Wider Benefits of Learning Research Report Series

The reports in this series are published by the Centre for Research on the Wider Benefits of Learning, Institute of Education, London. Reports published to date are listed below and are available online at **www.learningbenefits.net**.

2002
1. The wider benefits of further education: practitioner views
2. Parental perspectives of family learning
3. Learning, continuity and change in adult life
4. Learning, family formation and dissolution
5. Quantitative estimates of the social benefits of learning, 1: crime
6. Quantitative estimates of the social benefits of learning, 2: health (depression and obesity)

2003
7. Education, equity and social cohesion: a distributional model
8. The contribution of adult learning to health and social capital
9. The macro-social benefits of education, training and skills in comparative perspective

2004
10. A model of the inter-generational transmission of educational success
11. Adult education and attitude change
12. Education, training and the take-up of preventative health care
13. Identity, learning and engagement: a qualitative inquiry using the NCDS

2005
14. Education and youth crime: effects of introducing the Education Maintenance Allowance programme
15. Leisure contexts in adolescence and their effects on adult outcomes (online only)

2006
16. Does education have an impact on mothers' educational attitudes and behaviours?
17. Are those who flourished at school healthier adults? What role for adult education?
18. What is the relationship between child nutrition and school outcomes?
19. Are there effects of mothers' post-16 education on the next generation? Effects on children's development and mothers' parenting
20. Development in the early years: its importance for school performance and adult outcomes

2007
21. Determinants and pathways of progression to Level 2 qualifications: evidence from the NCDS and BHPS
22. Parenting behaviours and children's development from infancy to early childhood: changes, continuities and contributions
23. What role for the three Rs? Progress and attainment during primary school

Executive Summary

Introduction and Background

The acquisition and development of children's numeracy and literacy skills has long been of central concern to education policy, illustrated for example by debates around the teaching of reading and most recently by the announcement of a major review of maths teaching in primary schools (DCSF, 2007).

While a wealth of data shows that children's achievement test scores are strongly related to their prior cognitive functioning and attainment of basic skills in numeracy and literacy, children also demonstrate shifts and fluctuations in the trajectories of their skills and abilities, particularly during middle childhood.

This study builds on earlier work by the Centre for Research on the Wider Benefits of Learning (Feinstein & Duckworth, 2006) looking at the importance of school entry capabilities. Using data from a longitudinal sample of children surveyed from birth, we explore the balance of change and stability in children's school test scores during primary school. We consider how the foundation stones of literacy and numeracy operate to support progress in Key Stage attainment, looking at within- and between-subject effects for literacy, numeracy and science. We examine the importance of progress over the Key Stage 1 (5 to 7 years) period for progression between Key Stage 1 and Key Stage 2 (7 to 11 years), controlling for entry assessments, earlier cognitive development and child-level characteristics as well as a broad range of features of family and social background.

Key Findings

Stability and change in primary school attainment

- Overall there is substantial stability over the Key Stage 2 period, but also mobility, both upward and downward in the distribution. Based on the level of attainment at Key Stage 1 reading, 62.1% of children made the expected two levels of progress in Key Stage 2 English[1], 12% made more than two levels of progress, and 25.8% made lower than the expected levels of progress. In maths, 61.1% advanced two levels, 15% made more than the expected level of progress, and 23.9% failed to progress the expected two levels.

- There is a greater level of stability at the extreme ends of the distribution, more so in maths than in English, and greater mobility in the middle, although that mobility is generally limited in range: pupils of mid-level ability are more likely to remain towards the centre of the distribution than to fall into the bottom, or escape to the top, quartile.

[1] This figure includes pupils advancing all two level combinations between Key Stage 1 and Key Stage 2 across the Key Stage 1 distribution, i.e. working towards level 1 – Level 2; Level 1 – Level 3; Level 2C, 2B, 2A – Level 4; Level 3 – Level 5.

Effects within and between subjects

- Prior attainment of students measured by Key Stage 1 tests is the most important predictor of performance in Key Stage 2 assessments.

- These effects are higher within subject than between subjects.

- These subject effects do not interact to accelerate attainment. Thus, higher Key Stage 1maths scores are related to higher Key Stage 2 English scores, but they do not speed up progress in English from Key Stage 1 to Key Stage 2, and vice versa.

Who benefits? Sub-group differences

- Although overall girls make better progress in English and boys make better progress in maths at Key Stage 2, Key Stage 1 results are equally important as predictors of Key Stage 2 performance for both boys and girls.

- Children from more socially advantaged homes, on average, do better in both Key Stage 1 and 2 assessments.

- However, for children of parents with lower levels of education, doing well in Key Stage 1 tests, particularly in maths, is more important (i.e. more predictive of later attainment) than for other groups.

Methodology

Data are from the Avon Longitudinal Study of Parents and Children (ALSPAC), an extremely rich longitudinal study of children born in the former Avon Health Authority in the early 1990s. To be eligible for the study, mothers had to be resident in Avon while pregnant with an expected date of delivery between 1st April 1991 and 31st December 1992 inclusive. More than 80% of the known births from the geographically defined catchment area were included, resulting in a total cohort of 14,062 live births. This study design therefore contains children in three adjacent academic cohorts: those born 1st April – 31st August 1991, those born 1st September 1991 to 31st August 1992, and those born 1st September 1992 to 31st December 1992.

Data have been collected at numerous time points since birth and come from a variety of sources: the cohort member, their mother, her partner, clinic-based tests, and schools, covering Local Education Authority (LEA) entry assessment scores and information from the teacher. In addition, administrative data from the National Pupil Database (NPD) have been merged with the ALSPAC data providing records of individual achievement in the National Curriculum Key Stage 1 and 2 assessments. These data cover all relevant state schools in the four LEAs covering the Avon area.

Children within the sample have just finished years 11, 10 and 9 respectively and will have sat Key Stage 1 tests in the academic years 1997/98, 1998/99, and 1999/00 and Key Stage 2 in the academic years 2001/02, 2002/02 and 2003/04, providing a fairly contemporary analysis.

We use transition matrices and Ordinary Least Squares regression analysis to explore progress over the Key Stage 1 period and highlight the impact of Key Stage 1 attainment in reading, writing, spelling and maths for subsequent performance at Key Stage 2 English, maths and science.

Findings

Stability and change in primary school attainment

In English overall 62.1% made the expected progress, 12% made better than expected progress and 25.8% made worse than expected progress. In maths 61.1% remained on their expected trajectories, 15% did better, and 23.9% worse than expected. There is thus considerable stability in the figures, with the majority of pupils making the expected progress, but also substantial movement both upwards and downwards.

Table 1 shows the average probability of Key Stage 2 quartile position given pupils' position in the distribution at Key Stage 1. The majority of pupils remain in the same quartile of the distribution. Fifty per cent of bottom quartile scorers in Key Stage 1 still have scores in the bottom quarter of the distribution in Key Stage 2. The corresponding likelihood of "escaping" the bottom quartile is therefore also 50%, with 3% making it into the top quartile. At the other end of the distribution, the probability of "dropping out" of the top quartile between Key Stage 1 and 2 is 36.5%. Pupils of middle ability are thus more mobile than those in the top and bottom quartiles, but are more likely to remain in the centre of the distribution than to fall into the bottom or escape to the top quartiles.

Table 1: Transition matrix: Key Stage 1 factor score to Key Stage 2 factor score

	Key Stage 2 score				
Key Stage 1 score	Bottom quartile	Q2	Q3	Top quartile	**Total %**
Bottom quartile	50	33.5	13.5	3.0	100
Q2	20.7	38.3	30.4	10.6	100
Q3	6.0	25.1	39.9	29.1	100
Top quartile	1.1	8.8	26.6	63.5	100

Effects within and between subjects

Table 2 shows the changes in Key Stage 2 score associated with an increase of one level in Key Stage 1 attainment across subjects. Thus an increase of one level at Key Stage 1 in maths leads, on average, to an increase in the Key Stage 2 score of 8.29 points.

Reflecting the high degree of stability in pupils' trajectories noted above and in line with other research (for example, Melhuish, *et al.,* 2006), we find that performance in

a given subject is highly predictive of performance at Key Stage 2 in that subject. Thus reading, writing and spelling at Key Stage 1 are more important for Key Stage 2 English than maths and, likewise, doing well in maths is the best gauge for subsequent maths performance, more than four times as important as Key Stage 1 writing and over five times more important than the earlier assessment of reading. Literacy and numeracy skills thus seem to become increasingly entrenched in their subject specific domains as children move through primary school.

Table 2: Change in Key Stage 2 score associated with each one level increase in KS1 score (Regression model coefficients)

	Key Stage 2 assessments		
	English	**Maths**	**Science**
Key Stage 1 tests:			
Reading	3.19***	1.32***	1.95***
Writing	3.02***	1.85***	1.01***
Spelling	1.79***	1.78***	-0.68***
Maths	1.58***	8.29***	3.04***
Obs	9142	9123	9152
R^2	0.58	0.60	0.51

*** p <.001

Nevertheless, we do find evidence of smaller, but statistically significant cross-over effects between skill areas, with performance in Key Stage 1 maths showing a relationship with Key Stage 2 English and science results, and performance in Key Stage 1 reading and writing having a relationship with Key Stage 2 attainment in maths and science. Thus an increase in one level at Key Stage 1 in maths results, on average, in an increase of 1.58 points in Key Stage 2 English for example. In terms of the relative strength of the relationships between subjects, some are perhaps what we might intuitively expect, with Key Stage 1 maths performance predicting more strongly to Key Stage 2 attainment in science than in English. However, it is interesting to note that Key Stage 2 science scores are more strongly related to Key Stage 1 reading than writing skills. Spelling is the least predictive of all the Key Stage 1 skills and, in fact has a negative relationship with Key Stage 2 science results: once the effects of reading, writing and maths have been taken account of, doing well in spelling has no further benefit for a broader subject such as science which is influenced more equally by earlier literacy and numeracy skills.

Higher achievement in Key Stage 1 leads to gains in achievement at Key Stage 2. However, these subject effects do not interact to accelerate attainment. Thus, higher Key Stage 1 maths scores are related to higher Key Stage 2 English scores, but they do not speed up **progress** in English from Key Stage 1 to Key Stage 2, and vice versa.

Who gets the benefits? Sub-group differences

For all Key Stage 2 assessments, the prior attainment of pupils is the most important predictor of later performance and adding in controls for child, family and other background characteristics does not substantially change the size of the Key Stage 1

effects. However, our sub-group analyses find significant relationships between background factors and pupils' progression and attainment in primary school. Our results suggest a greater tendency for girls to be "upwardly mobile" in English and boys to be so in maths. However, this does not result in a significant difference in the **predictability** of Key Stage 2 results based on Key Stage 1: Key Stage 1 results are equally predictive of Key Stage 2 success (or failure) for both boys and girls. Rather, it reflects the differential attainments whereby girls have tended to do better in English and literacy-related skills and boys in maths.

Doing well in maths in the early years is more important (i.e. a better predictor of later success) for children from households with low levels of parental education than for children from well-educated households, better performance signalling greater probability of success in English and particularly maths (and conversely, poorer Key Stage 1 performance signalling greater likelihood of failure at Key Stage 2). On average, for a child from a low-education household, an increase of one level in the Key Stage 1 test results in an extra 2 points in Key Stage 2 English and an extra 9.4 points in Key Stage 2 maths. For a child from a well-educated household, the increases at Key Stage 2 for each level increase at Key Stage 1 are 1.2 points in English and 8.1 points in maths respectively.

This finding may appear at first glance to conflict with results from earlier research (Feinstein and Duckworth, 2006) which found that children from households of low socio-economic status who showed good early development at age 3 to 5, had relatively poor attainment at age 10. However, we suggest that these findings reflect important differences between these early development measures and the Key Stage 1 assessments as well as the developmental ages considered. Sensitivity analysis discounted the possibility that it was simply a feature of the metrics used. Rather, we suggest that the difference may lie in the fact that Key Stage 1 measures, unlike the early development measure used in the earlier research, are known both to child and teacher, thus informing their expectations of later performance. Where households lack the ability and educational resources to scaffold a child's progress, the ability of the school to recognise and scaffold potential assumes greater importance.

Conclusions and implications

In line with other research, we find that while Key Stage 1 tests are the best predictor of subsequent attainment at Key Stage 2, there is both stability and change in middle childhood. The stability is important because early achievement matters and is the foundation of subsequent success. But recognising change and discontinuity is also important because if people believe that ability is fixed and innate, then the response to poor educational performance, both by teachers and pupils, will be to disengage as nothing can be done. An appreciation of the balance between these two facets of progression is therefore important.

The results of the current study suggest that the signals provided by key stage tests and emphasis on basic numeracy skills may particularly benefit children from socially disadvantaged backgrounds. We hypothesise that doing well in Key Stage 1 maths is particularly important for these children because the public awareness of the test result and corresponding self-knowledge signals ability that might otherwise go

undetected. We believe that a greater understanding about why early maths carries such weight in predicting later attainment is warranted, particularly for those who might otherwise be at risk of failure. This is likely to require a greater appreciation of children's own interpretations of what key stage assessments mean, how they are valued by children and teachers, as well as a more in depth understanding of how different configurations of school and family contexts relate to stability and change in attainment.

We believe that the results presented here lend support to calls for personalised learning – by which we mean not simply targeting particular "treatments" at particular groups. Rather, we recommend a more holistic approach for children across the full spectrum of achievement, in which assessment practices – both formative and summative – are developed to provide a more complete picture of children's development and are used to support learning as well as measure it. Strategies that enable pupils to see, in frequent and meaningful ways, how they are progressing, will increase engagement and motivation and help develop their appetites for lifelong learning.

Finally, academic attainment is only one facet of educational success and only part of the Government's Every Child Matters agenda. There are other important school outcomes, such as a child's engagement and motivation for learning, relationships with peers and teachers and overall self-concept. These, and the influences upon them, need to be reflected in configurations of pupil abilities so as to understand individual learning trajectories more fully.

Acknowledgements

I am very grateful to comments and inputs from Leon Feinstein and other research staff at the Centre for Research on the Wider Benefits of Learning for their useful comments on this report. I would also like to thank Rachel Barker and colleagues at the Department for Children, Schools and Families (DCSF) for helpful suggestions.

The report has also benefited from comments during presentations at the Institute of Education Doctoral School Conference, (London, June 2007), the International Conference on Cohort Studies (University of Oxford, September 2006) and in the Bedford Group seminar series. Useful discussions were also held following the DCSF seminar on Early Development and Later Trajectories (London, February 2007) and the DCSF Economists' Awayday (Manchester, March 2007).

I would like to thank the DCSF for their financial support on this project.

Finally, I am grateful to all the mothers who took part in the study and to the midwives for their cooperation and help in recruitment. The whole ALSPAC Study Team comprises interviewers, computer technician, laboratory technicians, clerical workers, research scientists, volunteers and managers who continue to make the study possible. The ALSPAC study has been supported by funding from the Medical Research Council, the Wellcome Trust, UK government departments, medical charities and others and is part of the WHO initiated European Longitudinal Study of Pregnancy and Childhood.

Contents

Tables

Figures

1. Introduction

The study of cognitive development, especially in school-age children, has been one of the central focuses of developmental research for the past 50 years and continuity in cognitive attainments is well established (Kowleski-Jones and Duncan, 1999; McCall, Applebaum, and Hogarty, 1973; Wilson, 1983). A wealth of data shows that children's achievement test scores are strongly related to their prior cognitive functioning and attainment of basic skills in maths and literacy such as number and letter recognition. However, although children's academic achievement is largely stable throughout childhood, children do demonstrate both shifts and fluctuations in the trajectories of their skills, abilities and behaviours, particularly during middle childhood (Kowleski-Jones & Duncan, 1998; Pungello *et al.,* 1996).

This study builds on earlier work by the Centre for Research on the Wider Benefits of Learning (Feinstein & Duckworth, 2006) looking at school readiness and the importance of school entry capabilities, and explores the balance that exists between change and stability in children's test scores during primary school. The report seeks to inform the understanding of how changes in measured ability over the Key Stage 1 period predict subsequent educational success at Key Stage 2. Using data from a longitudinal sample of children surveyed from birth, we consider how the different foundation stones of literacy and numeracy operate to support progress and attainment in Key Stage assessment, looking both within and across the domains of literacy, numeracy and science. We examine the importance of progress over the Key Stage 1 period for progression between Key Stage 1 and Key Stage 2, controlling for earlier cognitive development and child-level characteristics measured prior to and at school entry as well as a broad range of features of family and social background.

1.1 Development in middle childhood

Middle childhood roughly spans the period from 5 to 12 years of age and is marked by considerable developmental changes in many dimensions. Consequently it has been distinguished from early childhood and adolescence along a number of dimensions (e.g. Collins *et al.,* 1984), yet has received comparatively little attention relative to other developmental periods (infancy, early -childhood and adolescence) in the literature. Moreover, the period between five and seven years of age forms the first years of schooling for almost every society that provides formal education, suggesting some universal recognition that the capabilities needed for such schooling emerge during this age period.

Changes in capabilities during this period are considerable with children showing skills, task approaches and distinctive patterns of behaviour that differ significantly from other times. Between six and nine years old, children start to gain skills that allow them to reason about increasingly complex problems and situations. For example, between six and seven years old children start to identify logical patterns and relationships in increasingly complicated problems, answer questions about abstract as well as concrete objects, and represent their ideas in organised and systematic ways. These developmental advances equip children with an increasing capacity to follow a line of enquiry, make predictions about likely outcomes and explain meaning in a variety of ways. Later, between ages 10 and 12, children begin

to develop more mature logic and reasoning skills, such as argumentation and hypothesis testing, and are increasingly able to control, monitor and adapt behaviour.

Throughout middle childhood, children become more adept at acquiring new information, consolidating, extending and integrating previously acquired knowledge, and using this information in reasoning and action. Evidence here suggests that during this period of development children "learn to learn", adapting previous strategies to develop new cognitive skills that enable them to think more flexibly and subtly than during the pre-school years. More general, underlying cognitive skills, such as oral language and conceptual ability, are thus likely to become increasingly important for mastering later and more complex tasks. In reading, for example, basic language skills often resurface when difficulty increases and early strategies such as "sounding out" come back into play as individuals make the transition from "learning to read" to "reading to learn" (NICHD Early Child Care Research Network, 2005a; Scarborough, 2001; Snow et al, 1998; Storch & Whitehurst, 2002; Whitehurst & Lonigan, 1998). Similarly, for maths, solid foundational concepts of numbers allow for deeper understanding of more complex mathematical problems and flexible problem solving techniques (Baroody, 2003; Ferrari & Sternberg, 1998; Hiebert & Wearne, 1996).

There is then both stability and change in middle childhood and to best support progression in learning for all children practitioners and policy makers need to have a clear understanding of this balance. Moreover, like those of other developmental periods, the changes that occur in primary school aged children are also marked by individual differences and continue to be influenced by broader features of development and background. The attainment averages that are typically reported, however, often conceal a great deal of this individual variability.

Coupled with the government's recent focus on progression and personalised learning (DfES, 2006; 2007), there is therefore an increasing need to address questions that focus on the extent to which and in what ways fundamental skills such as literacy and numeracy are fixed or become crystallised before children make the important transition to secondary school. Is attainment in the Key Stage assessments relatively stable? Are there differences in patterns of progression amongst the least and the most able? Are there spillovers from one type of skill to another? Do all children show systematic development in a given domain? Do some children benefit more from one skill than another? By understanding the complex patterns of stability and change in attainment during primary school, we will be better placed to address disparities of underachievement and ensure that every child is able to achieve their full potential.

1.1.1 The development of literacy skills

Acquiring literacy is a process that involves many different skills and reading, writing and spelling are based on similar developmental foundations and underpinned by similar cognitive processes (Fitzgerald & Shanahan, 2000). To become accurate and proficient readers, children need to assemble and decode words and their pronunciations, acquire spelling-sound mappings, learn regular rules and their exceptions, and make the leap from singular, staccato word recognition to reading fluency that requires the automatic retrieval of word strings and their related meaning. In the later stages of literacy development, the knowledge underlying these skills becomes increasingly consolidated and correspondingly more flexible. The

development of literacy skills then, is a cumulative process in which skills and abilities are revised, reworked and built on as children mature and foundational knowledge accrues.

Nation and Snowling (2004) highlight the breadth of language skills that contribute to the development of reading above and beyond decoding and word recognition. They followed the development of reading skills in children from the age of 8.5 to 13 years and found that earlier phonological awareness, non-phonological oral language skills, vocabulary knowledge and listening comprehension all contributed uniquely to individual differences in later reading performance, conditional on early reading skills. Leppänen *et al.* (2006) investigated the development of reading and spelling and found that these skills emerge in complementary ways, forming a recursive developmental pattern with positive development in one skill predicting growth and subsequent development in the other, conditional on earlier performance, and vice versa.

The developmental skills and capabilities involved in learning to read and write are dynamic and it therefore stands to reason that middle childhood is marked by both change and stability. A recent study by Phillips *et al.* (2002) demonstrates, for example, the changing and somewhat porous nature of relative reading performance throughout the primary school period. Their study examined the courses of reading achievement in Canadian children from first to sixth grades and found considerable mobility up and down the ability distribution, i.e. children below average in first grade became average in subsequent grades; a significant proportion of average students later being above average; and an almost equal probability of above average readers becoming average as remaining above average.

Reading and writing are also more than just decoding words on a page. Children need to understand what they have read and use that information to comprehend meaning and make inferences about problems being posed, as well as to amass information more generally. Literacy skills are thus implicated in the development of other skills such as those involved in maths and science; pupils who have difficulty reading and understanding the questions being asked will have problems answering those questions, irrespective of whether they have the necessary numeracy or problem-solving skills to do so.

1.1.2 The development of numeracy skills

Like the acquisition of literacy, the development of numeracy and mathematical skills is made up of a number of components which are built on in hierarchical ways over time (Aunola *et al.,* 2004; Entwisle & Alexander, 1990; Gelman & Gallistel, 1978; Nesher, 1986). To understand in maths, means to construct relations among facts, procedures and ideas. Even before they enter formal schooling, children intuitively start to piece together basic mathematical concepts, such as describing absolute size (e.g. “This one is bigger?”; “Who is older?”), part-whole relations (e.g. “I have the most”; “This is my half, that’s your half”), and early counting strategies. Learning basic skills, such as number and shape knowledge, one-to-one correspondence, spatial sense, and comparing and classifying groups, provide the basis for mastering more complex understanding of concepts, such as ordering and patterning, understanding of measurement, as well as additive and multiplicative relations.

Again, as with literacy-related skill development, as these early abilities mature and become increasingly automatic, attentional resources can be devoted to more complex problem-solving. Hiebert & Wearne (1996) investigated the dynamic relationships between instruction, understanding and skill in addition and subtraction tasks. Following 70 children over the first three years of school, the authors found that pupils applied specific understandings of existing mental structures to invent new procedures and adapt old ones. They report evidence of a persistent relationship between understanding and skill over this period and argue that a solid, conceptual understanding is critical in stimulating and guiding the development of new procedural skills.

Furthermore, they find that instruction encouraging pupils to develop, modify and reflect their own methods, as well as make sense of the strategies employed by their peers, facilitated higher levels of understanding and better connections between understanding and attainment. This focus on the relationship between what is taught and what is learned fits well with the recommendations of the 2020 Vision report on personalised learning (DfES, 2007) and the increasing recognition that assessment practices can support learning as well as measure them (e.g. Black & Wiliam, 2003).

Much of the research exploring the development of maths ability has focussed on particular subcomponents of mathematical knowledge, such as arithmetic computations or word-problem solving (Hiebert & Warne, *op cit*; Geary, 1993; Kulak, 1993; Share *et al.,* 1988) or on specific learning disabilities (Geary, 1993; Geary *et al.* 1999; Rourke & Conway, 1997). Consequently, less is known about the more general course of maths performance in non-selected populations. A recent exception to this comes from Aunola *et al.* (2004) who investigated the heterogeneity of developmental trajectories in pupils' overall maths development, i.e. individual differences in performance over time, during the transition to, and through, the first years of primary school. Their results suggest that in a developmentally normal sample of children, maths performance is stable, but shows increasing variance over time. That is, earlier maths performance positively predicted later maths performance, following a cumulative pattern, and individual differences in performance grew larger, fanning out, as children progressed through the early part of primary school (see also Muthén & Khoo, 1998; Williamson, *et al.,* 1991, for comparable findings among children later in middle childhood and in adolescence). Findings here also suggest that growth in maths ability was faster among those who entered school with an already high level of numeracy skills. The authors argue that their results suggest that not only may children gain in the long term from extra investment in maths from the very earliest stages of schooling, but also that those entering school with different levels of skills may further benefit from a different curriculum and/or methods of instruction.

There is also growing interest in the relationship between maths and reading. We noted above the intuitive link between literacy skills and other academic abilities, but several studies have shown that maths and reading skills share some of the same basic cognitive components, such as working memory, auditory and visual perception, and decoding and phonological skills (Bull & Johnston, 1997; Feinstein & Duckworth, 2006; Hecht *et al.,* 2001; Kail & Hall, 1999; Macaruso & Sokol, 1998; Swanson, Cooney & Brock, 1993). Melhuish *et al.* (2006) report high levels of within-subject stability from Key Stage 1 to 2, but also note between-subject effects between literacy and numeracy skills. Lerkannen *et al.* (2005) also use longitudinal data to explore the

prospective relationships between maths performance and reading comprehension in early primary school. Their results indicate that while earlier maths performance predicted subsequent development of reading comprehension, no evidence was found for the reverse pattern.

1.2 Key stage assessment in primary school

The DfES's Five Year Strategy (2004) states that attainment in primary education in England is improving, with many schools delivering high quality education as part of a rich and fulfilling primary experience. The national literacy and numeracy strategies (NLNS) introduced a systematic programme of teaching English and maths in primary schools in order to raise overall standards of literacy and numeracy and include a recommended daily entitlement of time dedicated to these areas of the curriculum. The NLNS lays down the objectives for each key stage of the National Curriculum, structures the core learning for each academic year and strand of the literacy and numeracy frameworks, and provides guidance on the expected ability for pupils at each level on the Key Stage spectrum. For example, in the *Using and Applying Maths* strand, children in Year 2 should be able to "Solve problems involving addition, subtraction, multiplication or division in contexts of numbers, measures, or pounds and pence", in Year 4 "Solve one-step and two-step problems involving numbers, money or measures, including time", and in Year 6 "Solve multi-step problems, and problems involving fractions, decimals and percentages". In English, a child who is at Level 3 should be able to read a range of texts, but a child who is at Level 4 can read between the lines, use inference and give an overview of the whole story.

The government believes that achieving this more advanced level of understanding, Level 4, at the end of primary school is an essential platform for achievement in secondary school and beyond. In the summer of 2002, for example, 70% of pupils who achieved Level 4 at Key Stage 2 went on to get five or more GCSEs at grades A* - C while just 12% achieved five GCSEs at grades A* - C who did not achieve Level 4 in 1997. Accordingly, a nationwide target for 85% of pupils to reach Level 4 at Key Stage 2 by 2008 has been set. However, despite improvements since 1997, there have been difficulties raising the levels of the lowest attaining children: one in four children goes on to secondary school ill-prepared for their secondary education and lacks the secure literacy and numeracy skills that enable them to access the secondary curriculum and make the most of their education. Consequently, improving the standards of reading, writing and maths in primary schools and ensuring progression at each stage of the children's education remains top priority in order that all children are able to achieve their full potential (DCSF press release, July 2007).

The importance of individual level improvement has been further highlighted in two recent DfES documents, *Making Good Progress* (2006) and 2020 Vision: Report of the teaching and learning in 2020 review group (2007). Recognising that not all children learn, progress or attain equally, these documents seek to further the current debate concerning progression, assessment and attainment throughout children's school careers and so develop better ways "to measure, assess, report and stimulate progress in our schools – so that every child develops at the best pace, and no child gets left behind" (p.1, 2006). This study is intended as a contribution to this debate and aims to unpack some of the individual differences, progression disparities and

attainment gaps that can get averaged out in nationally descriptive, target-focussed only data.

1.2.1 Attainment gaps

Gender

There is a considerable literature documenting the gender gap in attainment, which by the end of compulsory schooling, is considerable: girls are approximately 10 percentage points more likely to achieve five or more A* - C grades in GCSE than their male counterparts, a figure that has shown little variation since 1995. The gap in attainment at GCSE is relatively stable across the social class groupings but does vary by ethnic group, with Black Caribbean and Black Other pupils having wider gender gaps than other ethnic groups (DfES, 2005; see also Melhuish, *et al.,* 2006). This imbalance however, is not confined to achievement at GCSE and is evident at most stages in the educational system, typically emerging during primary school and widening as children move to secondary school. National statistics indicate that there is a persistent gender gap in English in favour of girls which is evident from the Foundation Stage through to GCSE. The equivalent gap in maths is smaller but girls are performing slightly better than boys at Foundation Stage, and at Key Stages 1, 3, and 4 (DfES, 2007b).

Recent research by the Effectiveness Pre-school and Primary Education 3 – 11 project (EPPE 3 - 11, Melhuish *et al.,* 2006) finds patterns of attainment at Key Stage 2 consistent with those reported by the DfES (2005) for 2002 to 2005, wherein girls outperform boys in English and related literacy subjects and boys do better in maths. There are no discernible gender differences in science. Similar gender differences have also been reported in research considering progress for the baseline to Key Stage 1 assessments (Strand, 1999). In their evaluation of the literacy hour, Machin & McNally (2003) find some evidence that at age 11, boys received a greater benefit from being exposed to the literacy hour than girls. However, the authors note that the evidence for statistically significant gender gaps is not strong and the effects disappear once KS2 controls are included in the regression suggesting that the literacy hour impact on gender differences operates more at primary level.

Socio-economic background

Despite overall improvement across all groups of pupils, the gap in average attainment continues to be strongly related to family socio-economic background. Feinstein (2003) finds evidence that this social class gradient kicks in significantly before children enter school suggesting that family contexts are particularly important in explaining educational disadvantage. However, the influence of family background is not a one-off phenomenon evident when children enter school, but rather exists throughout pupils' educational careers, with, for example, those eligible for free school meals (FSM) falling behind at every point in the Key Stage assessments (DfES, 2006b).

In part, these gaps in attainment persist simply because they are difficult to remedy; they are a product of complex relationships between individual characteristics, family background factors, as well as deep-seated social and educational difficulties. However, the compounding effect of socio-economic background as children move through compulsory schooling into further and higher education highlights the

importance of identifying those falling behind early on and the need for quality, sustained support.

1.3 The current study

In sum, this report considers how the different foundation stones of literacy and numeracy operate across the Key Stage assessments, looking both within and across the domains of literacy, numeracy and science. Our interest lies in exploring progress and attainment during primary school, conditional on entry assessment and individual characteristics. We therefore test how changes in attainment measured over the Key Stage 1 period predict attainment in Key Stage 2 assessments at the end of primary school, controlling for cognitive skills and other developing capabilities, such as social and communication skills and features of self-regulation, measured prior to and at school entry.

Findings from similar analyses in other longitudinal birth cohort studies highlight the predictive power of early assessments of developing capabilities and suggest important roles for early mastery of mathematical skills. We hope to further unpack these existing findings both through the use of a more recent birth cohort containing extremely rich child, family and school level measures, and by assessing developmental trajectories using current National Curriculum Key Stages. We predict a high degree of stability in the Key Stage assessments and hypothesise, in line with earlier research, that Key Stage 1 maths will be particularly important for predicting Key Stage 2 performance. The specific research questions we address in this research are:

1. How do children progress from Key Stage 1 to Key Stage 2?

2. What Key Stage 1 tests are most important in predicting success across Key Stage 2 assessments, conditional on pupils' ability at entry to school? Are there differences within and across the different domains of skill assessed?

Finally, in line with wider interests concerning the well-documented gender and socio-economic attainment gaps, we pay particular attention to these patterns of change and stability for different sub-groups of children and, within each of these two central research questions, ask:

3. Are there differences based on gender and features of children's family background?

2. Method

2.1 Data

Data are from the Avon Longitudinal Study of Parents and Children (ALSPAC), an extremely rich longitudinal study of children born in the former Avon Health Authority in the early 1990s, currently in years 9, 10 and 11. To be eligible for the study, mothers had to be resident in Avon while pregnant with an expected date of delivery between 1st April 1991 and 31st December 1992 inclusive. More than 80% of the known births from the geographically defined catchment area were included, resulting in a total cohort of 14,062 live births. This study design therefore contains children in three adjacent academic cohorts: those born 1st April – 31st August 1991, those born 1st September 1991 to 31st August 1992, and those born 1st September 1992 to 31st December 1992. Children within the sample have just finished years 11, 10 and 9 respectively. Mothers who were resident in the area but left shortly after enrolment were omitted from further follow-up. However, those who had completed the questionnaire scheduled for the third trimester of pregnancy before leaving the study area have been kept in the study, even if they had not delivered at the time of moving.

Data have been collected at numerous time points since birth, including full cohort information during the pre-school years and come from a variety of sources: the cohort member, their mother, her partner, clinic based tests, and schools, covering Local Education Authority (LEA) entry assessment scores, information from the teacher about the cohort member and the class as a whole. In addition to these rich data, administrative data from the National Pupil Database (NPD) have been merged with the ALSPAC data providing records of individual achievement in the National Curriculum Key Stage 1 and 2 assessments. These data cover all relevant state schools in the four LEAs covering the Avon area; Bristol, South Gloucestershire, North Somerset and Bath, and North East Somerset.

2.2 Measures

2.2.1 Outcome measures

Academic achievement is measured when the children are in year 6 of primary school, aged between 10 and 11 years old, in terms of Key Stage scores in English, maths and science. Children in this sample will thus have taken their key stage 2 tests in the period 2001/02 to 2003/04. The Key Stage 2 (KS2) tests are designed to assess pupils' knowledge and understanding of the KS2 programmes of study and provide a snapshot of pupils' attainment at the end of primary school.

We focus on the raw KS2 test scores rather than the teacher assessments as the latter are only available at the broad National Curriculum level and so provide a more general measure of attainment. However, correlations between the levels awarded by the tests and the teacher assessments are very high and so the results are unlikely to be substantively different. Correlations between assessment levels and teacher assessments are 0.79, 0.83, and 0.73 for English, maths and science respectively and are all statistically significant at $p<.001$. Moreover, there is evidence to suggest that

final teacher assessments are sometimes made after the KS2 tests results are received, creating an incentive for teachers to match their judgements with the level awarded (Black & Wiliam, 2003).

Key Stage 2 English

The English assessment is marked out of 100 and consists of three tests: a reading test, a writing test (made up of a longer task and a shorter task), and a spelling test. The spelling test is aggregated with the writing test. In the reading test, questions cover descriptive criteria as well as more open ended responses. The test includes different texts, such as stories, poems, explanations and interviews. The writing test is designed to assess pupils' individual independent work and marks are awarded for sentence structure, punctuation and text organisation, and compositions and effect. Marks are also awarded for handwriting.

Key Stage 2 Maths

The maths test is marked out of 100 and consists of three separate tests: a calculator paper, a non-calculator paper and a mental arithmetic test. Most of the questions are worth one point, although some are worth two and carry one point for showing appropriate working.

Key Stage 2 Science

The science test is made up of two papers. Pupils may be asked to describe how an investigation could be carried out, what factors need to be controlled and measured, whether an outcome can be predicted and how the results could be presented, as well as to explain the outcome and whether the evidence collected is significant, reliable and valid.

The marks and corresponding level boundaries are shown in Table 1. For further details on the assessing and reporting arrangements for these tests see: http://www.qca.org.uk/eara/documents/KS2_ARA.pdf

Table 1: Key Stage 2 level boundaries

	English	Maths	Science
Level	*Mark range*		
N	0 – 20	0 – 16	0 – 19
2	21 – 23	17 – 19	20 – 22
3	24 – 41	20 – 47	23 – 41
4	42 – 67	48 – 77	42 – 62
5	68 – 100	78 – 100	63 - 80

2.2.2 Key independent measures

The Key Stage 1 (KS1) assessments consist of standardised tests and teacher assessments, administered at 7 years of age. For this sample, the tests would have been conducted in the period 1997/98 to 1999/00. Teacher assessments were only

obtained for the youngest cohort of children in the ALSPAC dataset and so we focus here on the tests in reading, writing, spelling and maths.

The metric of these tests is only available in the basic scale of Key Stage levels and consists of levels 1, 2, 3 and 4+ with grades A, B and C within Level 2. Level 4 + is assessed by means of KS2 materials. However, there were so few children achieving Level 4+ (N=10) that it has been combined with Level 3. In addition, code W ("working towards Level 1") means the child was assessed but did not achieve Level 1. Children who were absent or disapplied from the KS1 tests, for example, due to special educational needs, are excluded from all analyses.

Key Stage 1 Reading

The Key Stage 1 reading assessment is made up of a reading task and a reading comprehension test. The reading task awards levels 1 and 2 only (with grades A to C, A being the highest). Those children achieving Level 2 in the reading task are then entered for the reading comprehension test. The reading comprehension test first awards Level 2 (with grades A to C). Children achieving grade A at Level 2 in both reading assessments are then assessed with the Level 3 reading comprehension test.

For simplicity, we use an overall reading result derived by taking the results of the reading task and adding Level 3/4+ from the reading comprehension test. This does not make any allowance for differing Level 2 grades in the reading task and the reading comprehension test, but does at least provide a reading variable that has a score for all children that were assessed. However, sensitivity checks on these data revealed remarkably linear relations within the dual reading assessments and in their predictive associations with the KS2 assessments. Therefore we do not consider the use of an overall reading score problematic.

The levels awarded reflect the skill and complexity pupils adopt in reading. For example, a pupil at Level 1 should be able to use a range of strategies, including accurate decoding of text to read for meaning, such as reading words on sight, making use of phonic strategies, using sentence and whole text knowledge and showing awareness of punctuation. A child at Level 2 will be able to understand, describe, select or retrieve information, events or ideas from texts and use quotation and reference to text, for example, recalling the main ideas of a story or information and identifying how characters behave. At Level 3, children are able to deduce, infer or interpret information, events or ideas from texts by using knowledge of a story so far to make predictions about what will happen next and clues from what characters do and say to explain their motives.

Key Stage 1 Writing

The writing task awards levels 1 to 3, again with grades A to C at Level 2. Writing skills include composition and effect, for example, being able to communicate meaning on a topic or idea to an outside reader, linking ideas and structuring information in meaningful and appropriate ways; sentence structure, for example, use of grammatically accurate statements and questions, clearly connected ideas, and being able to expand and adapt elements of sentence structure to gain more precise meaning. Punctuation and vocabulary are also assessed as part of the writing

assessment.

Key Stage 1 Spelling

The spelling test depends on the results of the writing task but awards a separate result: those pupils achieving Level 2 or above in the writing task, or whose teacher assessment in writing is above Level 2 or above take the spelling test. Pupils not achieving Level 2 or above may take the spelling test at the teacher's discretion. The spelling test awards levels 1 to 3, with no separate grading within Level 2.

Key Stage 1 Maths

Maths is assessed by means of a task, which awards Level 1, and a test, which awards Level 2, with grades A to C, and Level 3. The majority of the maths assessments focus on number and using and applying maths, for example, understanding of addition, subtraction, multiplication and division and the ability to carry out computations, ordering numbers and place values. The remainder of the assessments focus on understanding of shape, space and measures.

The means and standard deviations of the outcome and key independent measures are presented in Table 2 below. Table 3 reports the means and standard deviations for sub-groups of interest (see section 3.2.2 for further detail on these distinctions).

Table 2: Summary statistics for Key Stage 2 outcomes and Key Stage 1 measures

Variable	Obs	Mean	Std. Dev.	Min	Max
Key Stage 2					
English	9142	59.96	13.96	0	97
Maths	9123	66.66	19.43	0	100
Science	9152	59.97	11.06	14	80
Key Stage 1					
Reading	9858	3.67	1.25	0	5
Writing	9858	2.98	0.98	0	5
Spelling	9858	1.09	0.65	0	2
Maths	9858	3.54	1.18	0	5

Table 3: Means and standard deviations for Key Stage 2 outcomes and Key Stage 1 measures, by sub-group

	Gender		Household education		
	Girl	Boy	High	Medium	Low
Key Stage 2					
English	62.1	57.7	67.1	60.6	55.6
	(13.6)	(14.0)	(12.3)	(13.1)	(13.4)
Maths	64.6	68.9	76.2	67.6	60.4
	(19.3)	(19.3)	(16.4)	(18.2)	(19.7)
Science	59.6	60.4	65.7	60.9	56.2
	(11.3)	(10.8)	(8.8)	(10.1)	(11.2)
Key Stage 1					
Reading	3.8	3.5	4.1	3.7	3.4
	(1.2)	(1.3)	(1.1)	(1.2)	(1.2)
Writing	3.1	2.8	3.4	3.0	2.7
	(1.0)	(0.9)	(1.0)	(1.0)	(0.9)
Spelling	1.1	1.0	1.3	1.1	1.0
	(0.6)	(0.7)	(0.6)	(0.6)	(0.6)
Maths	3.5	3.6	4.0	3.6	3.3
	(1.2)	(1.2)	(1.1)	(1.1)	(1.2)

2.2.3 Key controls

The Entry Assessment data come from teacher-administered tests, developed by Reception teachers in partnership with head teachers, advisors and an educational psychologist, and are not the same as the National foundation stage assessments. There is no exact date in the documentation for when these tests were administered. However, according to the test booklet, teachers should set the tests once the children are generally settled in school and in the class (South Gloucestershire Professional & Curriculum Support Service, 1996). This has been broadly interpreted by ALSPAC administrators as 1st October of that academic year.

The primary purposes of the Entry Assessment were to establish an entry assessment of strengths and needs for pupils from which to plan and against which progress can be measured to the end of KS1. The Entry Assessment is made up of four required areas, each marked on a scale of 2 – 7.

Entry Language

The language assessment focuses on the use and understanding of language and assesses, for example, whether the child can carry out instructions, answer questions about him/herself and ask questions to satisfy needs and establish information, contribute to discussions, maintain dialogue with an adult and re-tell a story.

Entry Reading

The reading assessment covers whether the child can attentively listen to a story, look at books alone, talk about the pictures in a book, recognise names and words as well as notice and remark on visual details in words, draw inferences from stories and suggest more than one possible ending for a story.

Entry Writing

The writing assessment recognises that children come into the classroom with different levels of writing experience, with some pupils being able to write full sentences and others not able to hold a pencil. With this in mind, the assessment covers a wide range of skills from whether the child can use pencils/crayons/paint to make patterns on paper, and make letter-like shapes without adult direction, to writing their own names, explaining what the writing says, beginning to use invented spellings and writing a story.

Entry Maths

Like the writing assessment, the maths assessment covers a wide range of skills which include understanding pairs, matching colours, arranging items in specific orders, counting, recognising written numerals, and pointing to the first, middle and last items in a row of five objects.

2.2.4 Covariates

Covariates were added into the regression models in stepwise sets under the following headings.

Distal factors

Parents' education: This was based on mother-reported mother and father/partner highest level of educational qualifications coded on a scale from 0 to 4: CSE/lower; less than Level 2 academic and vocational qualifications; O-level/GCSE/Level 2 vocational qualifications; A-levels/Level 3 vocational qualifications; University degree and higher.

Other distal variables controlled for include socio-economic status, mother-reported income per week, employment status when the child was 18 months old, parents' age at the time of the child's birth, marital status, and family size and structure.

Maternal characteristics and proximal process

Characteristics of the mother included self-reported locus of control, her score on the Edinburgh Post Natal Depression scales, details of her social networks and social support and reports of her own school experience. Mother/partner-child interactions were measured using the summed responses to questions about how frequently they sing and read to/with their children, play with them and try to teach them colours, letters and nursery rhymes etc. Questions regarding the number of books and toys in the home were also asked.

Child development

Five areas of early development were used to control for emerging IQ and developing skills and capabilities:

1. *Communication: assessed at 6, 18, and 38 months*

Assessment of emerging vocabulary, including word combination scores, general communication & intelligibility scores.

2. *Social & personal development: assessed at 6, 18, 30 and 42 months*

Items were adapted from the Denver Developmental Screening Test and include age appropriate items such as playing pat-a-cake, helping with household chores, eating with a spoon and fork, washing and drying hands, and dressing. Items were scaled 0 = not started yet; 1 = only done 1 or 2 times; 2 = yes, often. Responses were summed to give an overall score.

3. *Fine & gross motor development: assessed at 6, 18, 30 and 42 months*

Again adapted from the Denver Developmental Screening Test, the fine motor scale includes items such as holding a rattle, ability to focus on and pick up small objects, and using a pencil. Gross motor development covers skills such as kicking and throwing a ball, jumping, and climbing stairs. Items were scaled 0 = not started yet; 1 = only done 1 or 2 times; 2 = yes, often. Responses were summed to give an overall score.

4. *Temperament: assessed at 6 and 42 months*

At 6 months, mothers completed the 88 item Carey Infant Temperament scale, which assesses how her infant behaves in different situations such as feeding, playing, being changed, sleeping, etc. The scale yields scores for activity, rhythmicity, approach, adaptability, intensity, mood, persistence, distractibility and threshold.

At 42 months, mothers completed the Revised Rutter Parent Scale for Pre-school Children. This scale is an extension of the Rutter behaviour scale and provides different behaviour scores covering emotional difficulties, conduct difficulties, hyperactive behaviour, prosocial behaviour and other behavioural difficulties.

5. *Physical development: assessed at 6, 18, 30 and 42 months*

Mothers were asked to report any general growth and development worries that they had, such as problems with speech, eyesight, weight, height or behaviour. Items were scaled 0 = no, not worried, 1 = yes, worried. Responses were summed to give a total score.

2.3 Analysis plan

The focus of this report is to explore progress and attainment during primary school and understand the balance that exists between change and stability in children's

school test scores. Therefore to examine how children progress from Key Stage 1 to 2 as well as what skills predict success at the end of primary school we use two main analytic methods.

2.3.1 How do children progress from Key Stage 1 to Key Stage 2?

Following a discussion of the descriptive statistics and outlining the broad picture of Key Stage attainment in these data, the first step in our analyses is a consideration of how children progress between KS1 and KS2. We focus on transition matrices to investigate level and quartile change and stability, also referred to as quartile continuities and discontinuities. Transition matrices report the conditional probabilities of being at a certain level or quartile at KS2, given relative position in the distribution at KS1. They therefore enable a focus on mobility in Key Stage attainment, not only with respect to being on or off target, but also in terms of who is making the expected levels of progress and who is exceeding or falling short of the two level improvement expectations. We also use this information to examine the likelihood of persistence in the top and bottom quartiles by gender and social status indicators of family background such as education and socio-economic group.

2.3.2 What Key Stage 1 tests are most important in predicting success at Key Stage 2?

The second stage of these analyses uses multiple regression analysis to examine the importance of progress over the KS1 period for attainment at KS2. We want to estimate the contribution that attainment in each KS1 test has for attainment in English, maths and science at KS2, conditional on their ability at entry to school and other background characteristics. In other words, we are interested in the benefit that pupils gain over the KS1 period, i.e. the change between ability assessed between entry assessments and KS1, for subsequent educational success, above and beyond earlier development and innate capabilities. In particular, we are interested to know what the balance is between literacy and numeracy related skills in relation to the different subject domains examined at the end of primary school.

As in the pre-school version of this study, the key difficulty in estimating these effects is that of differentiating between effects of innate characteristics and of features of development. Clearly, they are both highly related, but our interest here lies in the contribution made by the latter, holding the former constant so that we can get a guide to the issue of how children are moving through primary school and how formal experiences of school and the National Curriculum assessments influence this development.

For each KS2 outcome we estimate the same set of regression models to test the importance of the four KS1 tests of reading, writing, spelling and maths, given entry assessments, baseline developmental controls between 6 and 42 months, family context measures and school fixed effects.

The resulting general form of the equation is as follows:

$$(1)\ OUTCOME_{ij} = a_1 + ß_1READ_i + ß_2WRITE_i + ß_3SPELL_i + ß_4MATHS_i + e_{ij}$$

where $OUTCOME_{ij}$ is one of three j KS2 outcomes (English, maths and science) of child i; $READ_i$, $WRITE_i$, $SPELL_i$ AND $MATHS_i$ are the four KS1 tests; a_1 is a constant and e_{ij} is a stochastic error term. Our interest is in estimating $ß_1$, $ß_2$, $ß_3$ and $ß_4$, which under our estimation assumptions can be interpreted as the effects of KS1 performance on subsequent KS2 attainment and so provide clues about which features of these early capabilities might be most powerful in supporting positive educational trajectories for children.

One of the difficulties in estimating such effects using standard OLS regression is ensuring that we have taken into account in our estimation model the possibility that there may be other underlying factors that influence this development and so affect the outcomes under consideration. This will arise since characteristics of the child or her family such as IQ, earlier development and/or personality will be correlated both with children's test scores in KS1 and their later attainment.

To counter this and obtain more robust estimates of $ß_1$, $ß_2$, $ß_3$ and $ß_4$, we estimate a form of equation (1) that includes as many early measures of relevant child and family characteristics as possible, as well as school fixed effects to condition out any school level characteristics. In addition, we include local area entry assessments of language, reading, writing and maths, administered in the first few weeks of entering school. With these measures, our model becomes:

$$(2)\ OUTCOME_{ij} = a_1 + ß_1READ_i + ß_2WRITE_i + ß_3SPELL_i + ß_4MATHS_i + \alpha_1ENTRY_i + \gamma_1CHILD_i + \gamma_2FAM_i + \gamma_3SCHOOL_i + e_{ij}$$

where $CHILD_i$, FAM_i AND $SCHOOL_i$ are sets of child and family background characteristics and individual school identifiers included in analyses to control for individual differences.

These entry assessments (α_1), while not the National Curriculum Foundation Stage assessments, are remarkably similar (see section 2.2.3) and provide earlier measures of both literacy and numeracy skills. Therefore manipulation of equation (2) leads to a kind of *change* model. This estimation provides extra power against omitted-variable bias in estimating $ß_1$, $ß_2$ $ß_3$ and $ß_4$ to the extent that unobserved age invariant features of the child, for example, IQ and/or personality; and family, for example, parental education, are differenced out of the estimation. The coefficients on the KS1 measures thus represent the effect of change in attainment over the KS1 period. This equation can be written as:

$$OUTCOME_{ij} = b_1 + \delta_1 \Delta READ_i + \delta_2 \Delta WRITE_i + \delta_3 \Delta SPELL_i + \delta_4 \Delta MATHS_i + \lambda_1 ENTRY_i + \gamma_1 CHILD_i + \gamma_2 FAM_i + \gamma_3 SCHOOL_i + \eta_{ij}$$

"Δ" indicates a simple difference between entry assessment measures and the Key Stage 1 tests, i.e. the change or development in those measures. Algebraic manipulation shows that the parameters δ_1, δ_2, δ_3 and δ_4 are identical to the $ß_1$, $ß_2$, $ß_3$ and $ß_4$ parameters of equation (2) and can be interpreted as the effects of changes in measures of school entry capabilities.

The change model is considerably more robust to omitted variable bias than is a

model using measures at a given age. However, we cannot be certain that the comprehensive set of control variables used from our rich, longitudinal data capture all of the important confounding variables on early measures of child development and so remain cautious in making strong claims of causality; it remains possible that this approach will still produce biased estimates of $ß_1$, $ß_2$, $ß_3$ and $ß_4$. Nonetheless, the change model (equation 2) should be a substantially more robust estimate of the ß-coefficients than would result from the first model (equation 1) using levels, i.e. static measures, for a given point in time.

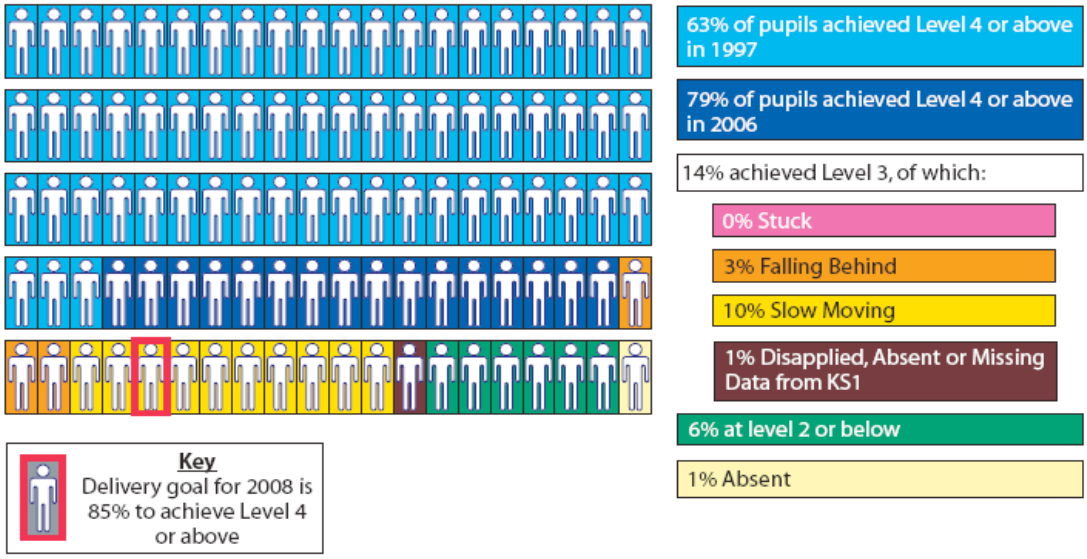
We estimate equations 1 and 2 to examine the differences between the four ß coefficients when additional controls are entered into the regression model. All estimation models control for gender and age (in days) at assessment. The sequential steps of the full regression model can be seen in Appendix Tables 1 to 3.

3. Results

3.1 A picture of attainment in primary school

To contextualise the results presented here, it is useful to compare the broad pattern of attainment of ALSPAC pupils with nationally representative statistics. The National Curriculum key stages are built on 10 levels which cover the age range 5 – 16 and are arranged so that the average student is expected to progress one level every two years. The target at KS1 is Level 2, specifically Level 2B, and Level 4 at KS2. We therefore begin by framing these data within the government's target-based expectations framework. Figure 1 shows the national average levels of attainment in KS2 English for pupils in 2006, each figure representing one pupil in every hundred.

Figure 1: Attainment in Key Stage 2 English, 2006



Source: Making Good Progress, DfES 2006

Based on a sample of 10,243 pupils who sat the KS2 English assessment earlier, in the summer terms of 2002, 2003 and 2004, our data show a comparable, but slightly higher pattern of attainment. In the current data:

- 79.4% of pupils achieved Level 4 or above
- 15.8% achieved Level 3, of which:
    - 0.1% Stuck (i.e. had not advanced from their Key Stage 1 Reading level)
    - 3.5% Falling Behind (i.e. Level 2B or 2A in Key Stage 1 Reading)
    - 12.2% Slow Moving (i.e. Level 2C or below in Key Stage1 Reading)
    - 0.1% Disapplied or absent from Key Stage 1 data
- 3.6% at Level 2 or below
- 1.2% Absent

For comparison, this N is based on pupils in the current sample who have complete KS2 English and KS1 reading data and does not omit those who may be missing assessment data in any of the other KS1 or KS2 assessments. However, for the remaining analyses we restrict our sample to those pupils who have complete data

across each of the four KS1 tests. These are necessary methodological constraints for our analyses, but limiting the sample in this way does mean that pupils in our data are, on average, performing slightly better than in the national picture of attainment in primary school.

As noted above, the National Curriculum targets for KS1 and KS2 are Levels 2B[2] and 4, respectively. Table 4 shows the different configurations of on/off target status in these data for KS1 reading, writing, spelling and maths.

Table 4: Frequencies and percentages of pupils achieving at least Level 2B in Key Stage 1 tests, by full sample and gender

	Full sample		Girls		Boys		Gender Gap
Level 2B or above in:	**N**	**%**	**N**	**%**	**N**	**%**	**%**
No subjects	*683*	7	*314*	6	*369*	8	-2
One subject:							
Reading only	*183*	2	*88*	2	*95*	2	0
Writing only	*18*	0	*9*	0	*9*	0	0
Spelling only	*330*	3	*162*	3	*168*	3	0
Maths only	*461*	5	*166*	3	*295*	6	-3
Two subjects:							
Reading and Writing	*25*	0	*16*	0	*9*	0	0
Reading and Spelling	*438*	4	*251*	5	*187*	4	1
Reading and Maths	*221*	2	*87*	2	*134*	3	-1
Writing and Spelling	*37*	0	*24*	0	*13*	0	0
Writing and Maths	*36*	0	*14*	0	*22*	0	0
Spelling and Maths	*371*	4	*134*	3	*237*	5	-2
Three subjects:							
Reading, Writing and Spelling	*467*	5	*305*	6	*162*	3	3
Reading, Writing and Maths	*101*	1	*55*	1	*46*	1	0
Reading, Spelling and Maths	*989*	10	*368*	7	*621*	13	-6
Writing, Spelling and Maths	*116*	1	*58*	1	*58*	1	0
All four Key Stage 1 subjects	*5382*	**55**	*2994*	**59**	*2388*	**50**	**9**
Total	***9858***		***5045***		***4813***		

Columns 1 and 2 report the full sample results and show that, in these data, more than half of the children, 55%, are achieving the target level of 2B or above in all four KS1 assessments. This is a more conservative estimate of target status since, unlike some of the nationally reported statistics, it does not include those achieving Level 2C. A further 10% are on target in at least two of these subjects and 17% are on target in three. However, 7% of pupils failed to reach the target level in any of these KS1

[2] The target in Key Stage 1 spelling is Level 2 as no sub-levels distinctions are made in this test.

assessments.
In line with wider interest in gender attainment gap, the final column shows the percentage point difference between girls and boys achieving Level 2B or above; a positive figure indicates girls have, on average, higher attainment than boys, a negative one shows boys to be outperforming girls. As in the summary statistics reported above (Table 3) and national averages, these descriptive results show girls are outperforming boys in all three literacy-related tests and are more likely to be on target across all four KS1 assessments. However, in contrast to the national average, boys in these data are doing better in KS1 maths.

Table 5 shows the frequency counts and related percentages for the same on/off target configurations in each of the KS2 assessments, for pupils with complete KS1 information.

Table 5: Frequencies and percentages of pupils achieving at least Level 4 in Key Stage 2 tests, by full sample and gender

	Full sample		Girls		Boys		Gender Gap
Level 4 or above in:	**N**	**%**	**N**	**%**	**N**	**%**	**%**
English only	*107*	1	*77*	2	*30*	1	1
Maths only	*42*	0	*10*	0	*32*	1	-1
Science only	*374*	4	*170*	4	*204*	5	-1
English & Maths	*51*	1	*26*	1	*25*	1	0
English & Science	*727*	8	*458*	10	*269*	6	4
Maths & Science	*465*	5	*107*	2	*358*	8	-6
All three Key Stage 2 subjects	*7150*	**78**	*3726*	**79**	*3424*	**76**	**3**
No subjects	*290*	3	*156*	3	*134*	3	0
Total	***9206***		***4730***		***4476***		

By the end of primary school, more than three quarters, 78%, are achieving the target of Level 4 or above in each of the KS2 assessments. A further 14% are on-target in at least two subjects and only 3% are off target on all three assessments. The gender gap is still evident, but is much smaller for pupils reaching the expected level of attainment in all three assessments. The rise in the percentage of pupils achieving the expected level of attainment in all assessed subjects by the end of primary school suggests improvement is being made between KS1 and KS2. However, breaking down attainment levels in this discrete target-based way masks a lot of variation and does not consider patterns of individual change and stability over the KS2 period. Moreover, we focus here just on those pupils achieving the specific Level 2B or above target and not those who are awarded Level 2C. Therefore, we now turn to our first research question and examine the relationships between Key Stage assessments over time and consider in more detail *how* pupils are progressing during primary school.

3.2 How do children progress from Key Stage 1 to Key Stage 2?

Overall, we find substantial stability over the KS2 period, but also mobility both upward and downward in the distribution. This mobility is socially stratified and shows some domain-specific variation by gender. Bivariate correlations between the all outcome and key independent variables are presented in Table 6 and are all positive and statistically significant at $p<.001$.

Table 6: Pairwise correlations between Key Stage 1 and Key Stage 2 measures

Variable	1.	2.	3.	4.	5.	6.	7.
Key Stage 2 outcomes							
1. English	1.00						
2. Maths	0.72	1.00					
3. Science	0.74	0.79	1.00				
Key Stage 1 measures							
4. Reading	0.71	0.62	0.61	1.00			
5. Writing	0.70	0.62	0.58	0.82	1.00		
6. Spelling	0.52	0.45	0.39	0.67	0.66	1.00	
7. Maths	0.60	0.75	0.63	0.71	0.71	0.51	1.00

The strengths of these relationships are greatest within the Key Stage assessment period, but the associations over time are also highly correlated, indicating a high degree of stability across primary school assessment. Correlations are also strongest within domain: KS1 reading and writing relate more to KS2 English and KS1 maths more to KS2 maths. KS1 spelling shows the weakest relationship to both the KS2 outcomes and other KS1 measures.

As noted above, the National Curriculum defines expected levels for each of the Key Stages such that most pupils will progress approximately one level every two years. Thus pupils should, on average, advance two levels between their KS1 and KS2 assessments in primary school. Transition matrices show the average probability, or the likelihood, of a pupil attaining a certain level in KS2 assessments, given the level they attained at KS1. Boxes are shaded grey to show the expected level of progression for each KS1 level attained over the KS2 period.

For maths (Table 7), for example, 71.8% of those achieving the lowest level at KS1, working towards Level 1, were still below the level of the test in their assessments at KS2. Similarly, at the top end of the distribution, 76% of pupils achieving Level 3 or above went on to get at least Level 5 in their maths assessment at the end of primary school.

Table 7: Transition matrix: Key Stage 1 maths to Key Stage 2 maths

	Key Stage 2 Maths						
Key Stage 1 Maths	Below test	Level 2	Level 3	Level 4	Level 5+	**Total %**	*Prop^n of whole*
Working towards L1	71.8	3.6	21.0	3.6	0	**100**	*1.8*
Level 1	20.8	6.7	53.6	18.0	0.9	**100**	*8.9*
Level 2C	2.4	1.7	42.3	47.9	3.9	**100**	*20.1*
Level 2B	0.4	0.3	16.6	66.4	16.3	**100**	*23.1*
Level 2A	0	0	4.6	57.5	37.8	**100**	*22.0*
Level 3+	0	0	0.6	23.5	76.0	**100**	*24.0*

Table 7 also clearly demonstrates that the majority of pupils are achieving the two level advancement over this Key Stage period emphasising that there is considerable stability in maths attainment across the two assessments: 53.6% of pupils gaining Level 1 in KS1 maths achieve Level 3 in KS2; 66.4% of those on target at Level 2B go on to Level 4 four years later. However, there is also evidence of discontinuity, with pupils both "escaping" from and "falling off" these expected trajectories. For example, the escape rate, i.e. those who advance more than the projected two levels, from Level 1 at KS1 is 18.9%, with 0.9% of these pupils getting Level 5+ in KS2. The corresponding fall off rate is 27.5%, and while 6.7% progress to Level 2, 20.8% appear to have made little or no progress and remain below the level of the test. Furthermore, 0.4% of those who were on target at KS1 have fallen to below the assessment level of the KS2 test. Overall, 61.1% of pupils advanced two levels, 15% made more than the expected level of progress, and 23.9% fell short of the expected two level progression[3].

The transitions from KS1 reading to KS2 English (Table 8) are comparable.

Table 8: Transition matrix: Key Stage 1 reading to Key Stage 2 English

	Key Stage 2 English						
Key Stage 1 Reading	Below test	Level 2	Level 3	Level 4	Level 5+	**Total %**	*Prop^n of whole*
Working towards L1	68.8	3.6	20.6	6.7	0.5	**100**	*2.1*
Level 1	19.8	5.1	45.9	28.3	0.9	**100**	*13.0*
Level 2C	2.5	1.3	35.7	55.3	5.1	**100**	*16.2*
Level 2B	0.2	0.1	13.0	74.0	12.7	**100**	*20.7*
Level 2A	0.1	0	4.9	69.3	25.8	**100**	*16.3*

[3] This figure includes pupils advancing all two level combinations between KS1 and KS2, i.e. boxes shaded grey: working towards level 1 – level 2; level 1 – level 3; level 2C, 2B, 2A – level 4; level 3 – level 5.

Level 3+	0	0	0.3	35.0	64.7	**100**	*31.7*

Over two thirds, 68.8%, of pupils working towards Level 1 at KS1 remain below the level of the test at KS2, while 7.2% of them progress to Level 4 and above. Just over three quarters, 75.1%, of those who gained only a Level 1 have progressed at least two attainment levels, with 29.2% reaching or exceeding the KS2 target of Level 4. Of those "on target" and achieving Level 2B, 74% remain at the target level in KS2 English, with 12.7% progressing to Level 5+ and only 13.3% falling "off track". In total, 62.1% of pupils made the expected two levels of progress in KS2 English, 12% made more than two levels of progress, and 25.8% made lower than the expected level of improvement. There is a greater level of stability at the extreme ends of the distribution, more so in maths than in English, and greater mobility in the middle. However, pupils of mid-level ability are more likely to remain towards the centre of the distribution.

For KS1 writing (Table 9) the mean mobility is slightly higher: pupils are more likely to advance at least the expected two levels of key stage attainment and less likely to fall off this track.

Table 9: Transition matrix: Key Stage 1 writing to Key Stage 2 English

	Key Stage 2 English						
Key Stage 1 Writing	Below test	Level 2	Level 3	Level 4	Level 5+	**Total %**	*Prop^{n} of whole*
Working towards L1	51.4	5.4	31.6	11.0	0.6	**100**	*4.7*
Level 1	15.4	4.7	46.1	32.7	1.1	**100**	*10.7*
Level 2C	1.4	0.7	26.5	63.5	7.9	**100**	*30.5*
Level 2B	0	0	4.0	66.2	29.8	**100**	*29.9*
Level 2A	0	0	0.3	37.5	62.2	**100**	*17.0*
Level 3+	0	0	0.3	14.2	85.5	**100**	*7.4*

Just over half, 51.4%, of those working towards Level 1 at KS1 remain below the level of the test at KS2, with 31.6% achieving Level 3 and 11.6% going on to attain the target of Level 4 or above. Almost half of pupils below target at Level 1, 46.1%, advanced two attainment levels and over a third progressed more than this expected rate to reach Level 4. For those pupils on target or above at KS1, the probability of remaining so is 96% for those at Level 2B and over 99% at Levels 2A and 3+.

Tables 10 & 11 show the transitions for KS2 science. We show the transition matrices from both KS1 reading and maths as we do not have an earlier science test.

Table 10: Transition matrix: Key Stage 1 reading to Key Stage 2 science

	Key Stage 2 Science						
Key Stage 1 Reading	Below test	Level 2	Level 3	Level 4	Level 5+	**Total %**	*Prop^n of whole*
Working towards L1	31.8	1.3	32.6	31.3	3.0	**100**	*2.1*
Level 1	4.2	1.3	27.8	56.8	9.9	**100**	*13.0*
Level 2C	0.4	0.1	13.7	65.9	20.0	**100**	*16.2*
Level 2B	0.0	0.1	5.4	60.5	34.1	**100**	*20.7*
Level 2A	0.0	0.0	2.5	52.9	44.6	**100**	*16.3*
Level 3+	0.0	0.0	0.2	20.0	79.8	**100**	*31.7*

Table 11: Transition Matrix: Key Stage 1 maths to Key Stage 2 science

	Key Stage 2 Science						
Key Stage 1 Maths	Below test	Level 2	Level 3	Level 4	Level 5+	**Total %**	*Prop^n of whole*
Working towards L1	39.5	1.5	34.2	23.9	1.0	**100**	*1.8*
Level 1	4.7	1.6	34.8	53.7	5.2	**100**	*8.9*
Level 2C	0.5	0.1	14.8	67.5	17.1	**100**	*20.1*
Level 2B	0.1	0.0	4.6	60.9	34.3	**100**	*23.1*
Level 2A	0.0	0.0	1.2	41.4	57.5	**100**	*22.0*
Level 3+	0.0	0.0	0.3	17.5	82.2	**100**	*24.0*

The mean mobility for these cross-domain transitions, i.e. between subjects, is higher than the within-domain maths to maths and English to English ones, indicating that both literacy and numeracy KS1 skills are important for subsequent attainment in science. For example, fewer than 40% of pupils working below the level of the test at KS1 remain below Level 2 at KS2, with over 30% making the two levels expected and 34% (reading) and 25% (maths) of pupils advancing to on-track status by the end of primary school. Upward mobility is also higher further up the distribution with very few pupils falling off-track: only 5.5% of pupils achieving Level 2B in KS1 reading and 4.7% in KS1 maths were stuck or making slow progress in KS2 science.

The preceding tables give a picture of pupil transitions between KS1 and KS2 in terms of the National Curriculum assessment framework. Table 12 summarises this information by combining the four KS1 measures and three KS2 outcomes into single factors of broad performance at the two time points. This enables consideration of a more general level of attainment in terms of stability and change in quartile positions in Key Stage scores that is less constrained by the level–bound information provided by the separate KS1 assessments.

Table 12: Transition matrix: Key Stage 1 factor score to Key Stage 2 factor score

	Key Stage 2 score				
Key Stage 1 score	Bottom quartile	Q2	Q3	Top quartile	**Total %**
Bottom quartile	50	33.5	13.5	3.0	100
Q2	20.7	38.3	30.4	10.6	100
Q3	6.0	25.1	39.9	29.1	100
Top quartile	1.1	8.8	26.6	63.5	100

As with the level-specific transition matrices, these summary data show a greater level of stability at the ends of the distribution and more fluidity in the middle. Fifty per cent of bottom quartile scorers in the KS1 tests still have scores in the bottom quarter of the distribution in the KS2 assessments. The corresponding likelihood of escape is therefore, also 50%, with 3% making it into the top quartile. At the other end of the distribution, the probability of dropping out of the top quartile between KS1 and KS2 is 36.5%. Again, middle-ability pupils are more mobile than those in the top and bottom quartiles, but are more likely to remain in the centre of the distribution than to fall into the bottom or escape to the top quartiles.

3.2.1 Moderation by gender

In line with the wider literature on pupil attainment, we also consider how these transitions over the KS2 period differ by features of the child. Tables 13 and 14 show the transitions between KS1 reading and KS2 English and KS1 and KS2 maths respectively, by gender. These results reflect the basic score and attainment level differences reported above showing girls as doing better in English-related tests in both the KS1 and KS2 assessments and boys as doing better in maths. Furthermore, these results suggest that the progression trajectories are moderated by gender: girls are more likely than boys to progress the expected two attainment levels, as well as to exceed this expectation, moving to the highest level at KS2. Conversely, girls are less likely to remain at the lower ends of the attainment distribution given their position in the bottom of the distribution at KS1 (60.8% versus 67.3%). For maths, the opposite is true: boys are more likely to remain at the top ends of the distribution and both meet and exceed target expectations than girls. In section 3.3.1, we consider whether this consistently found gender effect matters differentially for girls and boys in predicting subsequent KS2 performance.

Table 13: Transition matrix: Key Stage 1 Reading to Key Stage 2 English, by gender

	Below level of test		Level 2		Level 3		Level 4		Level 5+		Total	*Proportion of whole*	
	Boys	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**		**Boys**	**Girls**
Working towards L1	68.0	70.0	3.9	2.9	20.9	20.0	6.5	7.1	0.7	0	**100**	*2.8*	*1.3*
Level 1	22.3	15.6	5.9	3.7	45.3	46.9	25.7	32.9	0.9	1.0	**100**	*16.2*	*9.7*
Level 2C	2.7	2.3	1.8	0.7	39.1	31.3	52.2	59.5	4.2	6.3	**100**	*18.2*	*14.2*
Level 2B	0.3	0.1	0.1	0	15.4	10.5	72.0	76.2	12.3	13.2	**100**	*20.8*	*20.6*
Level 2A	0.1	0	0	0	6.1	3.7	70.3	68.4	23.4	27.9	**100**	*15.6*	*17.0*
Level 3+	0	0	0	0	0.6	0.1	41.2	30.5	58.3	69.5	**100**	*26.5*	*37.1*

Table 14: Transition matrix: Key Stage 1 Maths to Key Stage 2 Maths, by gender

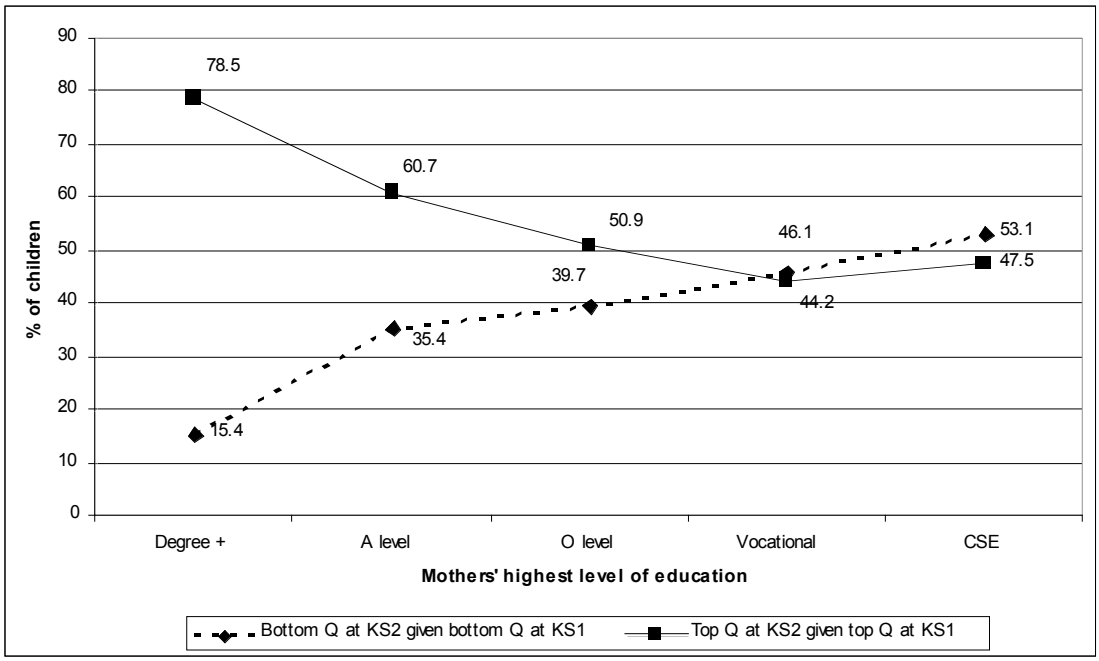
	Below level of test		Level 2		Level 3		Level 4		Level 5+		Total	*Proportion of whole*	
	Boys	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**		**Boys**	**Girls**
Working towards L1	71.1	73.0	3.3	4.1	20.7	21.6	5.0	1.4	0	0	**100**	*2.2*	*1.4*
Level 1	21.0	20.5	5.0	8.9	53.1	54.1	19.4	16.3	1.5	0.2	**100**	*9.7*	*8.2*
Level 2C	2.8	2.0	1.6	1.9	37.5	47.2	53.4	45.8	4.7	3.1	**100**	*20.1*	*20.1*
Level 2B	0.2	0.5	0.3	0.2	16.1	17.1	62.7	69.8	20.8	12.3	**100**	*21.7*	*24.7*
Level 2A	0	0	0.1	0	4.5	4.7	53.6	61.4	41.9	33.9	**100**	*21.3*	*22.7*
Level 3+	0	0	0	0	0.6	0.5	20.1	27.4	79.4	72.1	**100**	*25.1*	*23.0*

3.2.2 Moderation by social background

In line with general social mobility and underperformance from children from more disadvantaged backgrounds and those experiencing other forms of social difficulties, we also consider the influence that social background has on these patterns of change and stability. Earlier work by the Centre (for example, Feinstein, Duckworth & Sabates, 2004; Gutman & Feinstein, 2007) has demonstrated the importance of maternal education as an indicator of social background and influence on child development, and we therefore use this as our measure.

Following on the transition matrices discussed above, Figure 2 shows the probability of being in the top or bottom quartile of the KS2 English distribution, conditional on having been in the top or bottom quartile in the KS1 English tests, i.e. stability or continuity in quartile position, broken down by mothers' highest level of education.

Figure 2: Top and bottom quartile continuity in English, by mothers' education



While the average probability for all pupils of being in the top quartile at KS2 given being in the top quartile at KS1[4] is 57.4%, for children whose mother has a degree or higher this probability rises to 78.5% and for those whose mothers have only CSEs or lower it falls to 53%. Similarly, the likelihood of remaining in the bottom quartile at the end of primary school given you were in the bottom 25% in the Key Stage 1 literacy tests is 43.5% on average, but drops down to just 15.4% for children of the mothers with the highest levels of education in our sample and rises to 47.5% for those with the lowest.

We can also look at sub-group differences by considering both mothers' and fathers' (or partners') education to gain an overall measure of social background of the

[4] We use the quartile position in the distribution of the score for the first factor of the three Key Stage 1 English assessments, i.e. the combination of reading, writing and spelling, to explore progression in literacy-related skills during primary school.

household. In the subsequent analysis, parents' education is measured by the combination of both parents' highest level of qualification and is defined as high/medium/low broadly corresponding with DCSF qualification distinctions.

Table 15: Parents' combined level of education

	Father/partner's education:				
	CSE & lower	> Level 2 Vocational	O-Level	A-Level	Degree +
Mothers' education:					
CSE & lower	LOW				
> Level 2 Vocational					
O-Level (Level 2 equiv)			MEDIUM		
A-Level (Level 3 equiv)					
Degree +				HIGH	

Low education is broadly defined as both parents not being at Level 2 (5 O-Levels/ GCSEs or vocational equivalent) and *high education* as both parents having at least Level 3 (2 A-Level passes or equivalent) qualifications. *Medium education* is defined as every other combination between low and high. Where father/partner's education is not available, we use mother's education defined using the same cut offs.

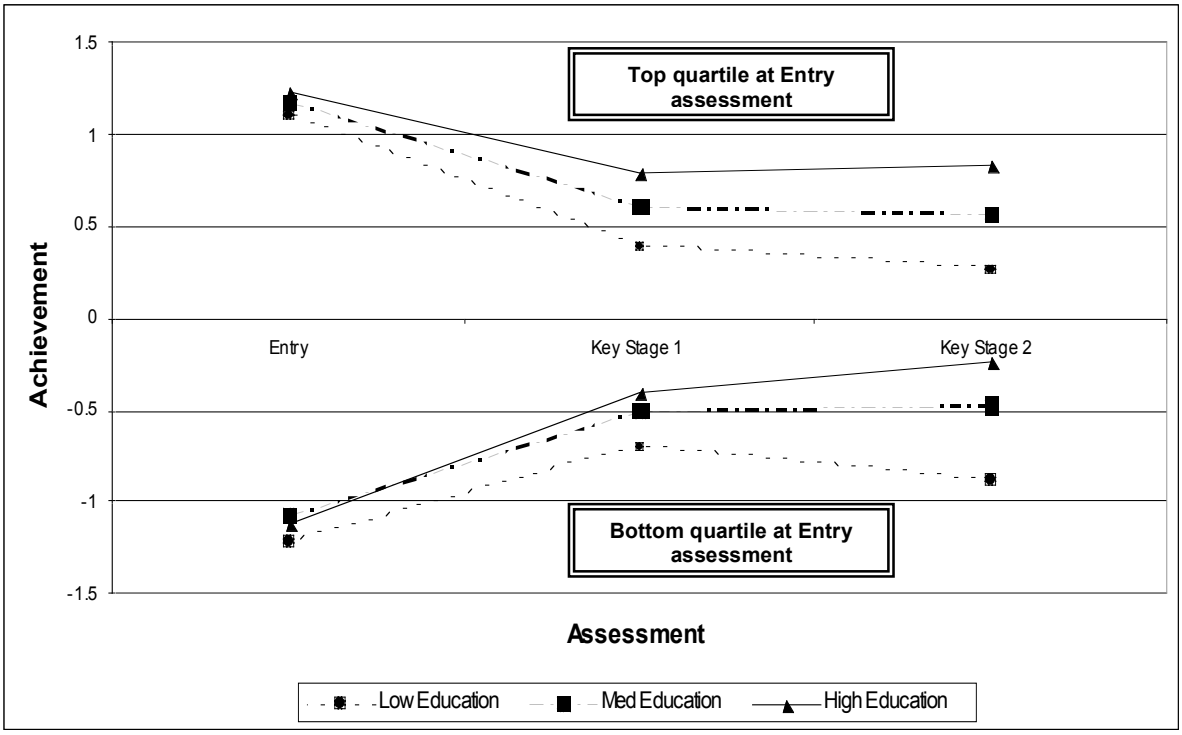
Our analysis shows that in addition to the strong association of social background at the initial position at school entry (Table 16), there are also substantial relationships between indicators of family background and movements over the primary school years.

Table 16: Summary statistics for Entry Assessment, by parents' highest level of education

	High			Medium			Low		
	Obs	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**St. Dev.**	**Obs**	**Mean**	**Std. Dev.**
Entry Assessment									
Language	3693	3.74	(.80)	4592	3.61	(.87)	3425	3.37	(.96)
Reading	3693	3.26	(.66)	4592	3.11	(.68)	3425	2.89	(.76)
Writing	3693	3.04	(.64)	4592	2.93	(.69)	3425	2.77	(.77)
Maths	3693	3.52	(.79)	4592	3.33	(.87)	3425	3.05	(.93)

Figure 3 demonstrates the compounding effects of social background in assessment as pupils move through primary school. At each assessment point, we take the first factor of the combined tests. Figure 3 then shows the mean of the factor scores at KS1 and KS2 for pupils who were in the top and bottom 25% of the entry assessment score, by parents' level of education (see Table 3 for KS1 & KS2 sub-group summary statistics).

Figure 3: Stratification in school assessment: Mean KS1 and 2 factor scores by top and bottom quartile position in Entry Assessments, by parents' education



In both the upper (top 25% at entry assessment) and the lower (bottom 25% at entry assessment) parts of this graph it can be seen that the relative performance of children from more educated households is consistently higher than those from households with medium and low levels of education, regardless of position in either the top or the bottom quartile. Moreover, over the course of primary school this gap in performance widens, highlighting that the attainment gap is more than just a one-off phenomenon in terms of relative performance at any one point in primary schooling, but is consistent and has continuing effects that persist and grow over time. A similar pattern also emerges if we consider social background as defined by parental socio-economic group.

3.3 What Key Stage 1 tests are most important in predicting success at Key Stage 2?

The results discussed above explore descriptively how pupils progress across the KS2 period and highlight gender differences and the compounding effects of social background on assessments throughout the primary school years. Our second research question examines what skills predict success at KS2 and uses regression modelling to estimate the importance of changes measured in literacy and numeracy ability over the KS1 period for predicting success at KS2 in English, maths and science. Overall, we find that prior attainment of pupils measured by KS1 tests is the most important predictor of performance in KS2 assessments and that these effects are higher within-subject than between-subject.

Table 17 shows the regression of the KS2 assessments on the four KS1 tests. For each

KS2 outcome, model 1 shows the effects for each KS1 test controlling for age and gender (see discussion of equation 1 in the Analysis Plan, section 2.3 above). Model 2 shows the estimates for the same KS1 tests, conditional on entry assessments and other comprehensive measures of child and family background characteristics, including school fixed effects. These estimates can therefore be interpreted as the effect of change in attainment over the KS1 period (see equation 2). These coefficients are shown in the natural metric of the Key Stage information summarised in section 2.2.2 above, such that each unit increase in the KS1 test, i.e. for each extra level attained in the four tests, leads to a corresponding increase in points of the KS2 score.

Table 17: Coefficients and standard errors from regression models of Key Stage 2 scores on Key Stage 1 tests

	Key Stage 2 assessments					
	English		Maths		Science	
	(1)	(2)	(1)	(2)	(1)	(2)
Key Stage 1 tests:						
Reading	3.75***	3.19***	1.78***	1.32***	2.43***	1.95***
	(.14)	(.13)	(.18)	(.18)	(.11)	(.11)
Writing	3.67***	3.02***	2.77***	1.85***	1.75***	1.01***
	(.16)	(.16)	(.21)	(.22)	(.13)	(.13)
Spelling	1.25***	1.79***	1.10***	1.78***	-1.13***	-0.68***
	(.23)	(.23)	(.31)	(.31)	(.20)	(.20)
Maths	1.94***	1.58***	8.92***	8.29***	3.46***	3.04***
	(.12)	(.12)	(.16)	(.17)	(.11)	(.11)
Control variables		X		X		X
Obs	9142	9142	9123	9123	9152	9152
R^2	0.49	0.58	0.52	0.60	0.40	0.51

* p <.05, ** p <.01, *** p <.001

These results show a very strong pattern of stability within the Key Stage assessments with all four KS1 tests being substantial predictors of all three outcomes assessed at KS2. For each of the outcomes considered, at least 40% of the variance is explained by the KS1 tests alone (model 1) and for KS2 maths it is more than half, 52%. Moreover, the size of these estimates hardly changes when additional control sets for child, family and school contexts are introduced into the model, with all KS1 predictors remaining highly statistically significant in model 2 ($p<.001$) and the proportion of variance explained does not substantially increase. These findings are discussed in detail below. The full stepwise OLS regression tables are reported in Appendix Table 1 to 3.

The strength of these relationships is strongest within subject areas reflecting strong within-domain continuity. Thus the conditional effects (model 2) of KS1 literacy skills are greatest for KS2 English and KS1 maths is most important in predicting

subsequent attainment in maths. Change in maths ability over the KS1 period carries particularly strong associations with later maths performance: conditional on ability assessed at entry to school, development in the pre-school years, family background and school effects, each increase in the level of attainment in KS1 maths results in an additional 8.29 points in the KS2 maths score.

There is also evidence that development in literacy and numeracy skills over the KS1 period is also important for later cross-domain outcomes, i.e. early literacy skills are also important for later numeracy skills and vice versa. For example, an increase in one level at KS1 in maths results, on average, in an increase of 1.58 points in KS2 English. However, these subject effects do not interact to accelerate attainment and progress in maths does not depend on progress in English-related subjects. These results also suggest that early numeracy skills are more important in predicting later science attainment than KS1 English-related tests, but that of these literacy skills higher levels of reading are more predictive of the increases in science attainment. As in the correlation matrix above (Table 6), spelling is the least predictive of the four KS1 tests, of success at KS2. This finding is further discussed below.

3.3.1 Moderation by gender

The summary statistics and transition matrix results reported above show that, on average, girls do better and are more likely to remain at the top end of the KS2 distribution conditional on their attainment at KS1 in English-related subjects, whereas boys perform better and show greater progress in maths. The results of the regression analysis, however, indicate that there are no significant gender differences in the predictive importance of KS1 tests for performance at the end of the KS2 period (see Appendix Table 4). That is, while girls are doing better in English and boys better in maths, performance in the KS 1 tests does not matter in different ways for girls and boys and there are no interaction effects between child gender and KS1 performance.

3.3.2 Moderation by social background

Table 18 presents the results for the sub-group analysis for all three KS2 outcomes according to parents' highest level of education (see section 3.2.2). All the results presented include the full set of controls for prior ability as well as child and family background characteristics and school fixed effects.

As in the full sample analyses, KS1 tests in reading, writing and maths continue to strongly predicts KS2 attainment for each of the three groups and the strength of the within-domain continuity also remains very clear. The pattern of results here suggests that the returns to KS1 tests are generally greater for children whose parents have lower levels of education, i.e. doing well in KS1 tests is more predictive of later attainment than for children of families with medium and high levels of education. Most interesting is that for both English and maths outcomes, the effect of KS1 maths is substantially and statistically significantly higher for pupils from homes where parents' education is low. On average, for these pupils each one unit increase in the level of KS1 maths attainment results in 2.04 extra points in KS2 English and 9.39 extra points in KS2 maths (both $p<.001$). Given that in these analyses there are six possible attainment levels in these KS1 tests (W, L1, L2C, L2B, L2A, and L3+), this difference means, for example, that a pupil achieving Level 3 or higher in KS1 maths

will, on average, score 12 points higher in KS2 English and 56 points higher in the maths assessment than his friend who is only working towards Level 1. In comparison, the gain for children from homes where parents have high levels of education is 7 points for English and 49 points for maths. This gain is considerable, particularly given the level boundaries outlined in Table 1 above.

For KS2 English and maths, KS1 spelling is more important for those from homes where parents have medium and high levels of education, whereas for science for children of low and middle educated households it predicts negatively. As in the full sample analyses, this result does not reflect a negative bivariate relationship between spelling and later KS2 outcomes, but again suggests that good spelling may be a particular effect of within-domain subject specialisation and that, for the average pupil, is not as important as the other KS1 tests in predicting later educational success. We discuss these findings in further detail in section 4.2.

Table 18: Coefficients and standard errors from regression models of Key Stage 2 scores on Key Stage 1 tests, by parents' highest level of education

	Key Stage 2 English			Key Stage 2 Maths			Key Stage 2 Science		
Parents' education:	**High**	**Medium**	**Low**	**High**	**Medium**	**Low**	**High**	**Medium**	**Low**
Key Stage 1 tests:									
Reading	3.06 ***	2.95 ***	3.40 ***	0.85 *	1.26 ***	1.21 **	1.63 ***	1.62 ***	1.93 ***
	(.31)	(.24)	(.31)	(.40)	(.33)	(.41)	(.24)	(.20)	(.27)
Writing	2.54 ***	3.26 ***	3.25 ***	1.26 **	1.95 ***	2.08 ***	0.79 **	0.97 ***	1.48 ***
	(.35)	(.30)	(.39)	(.43)	(.39)	(.56)	(.26)	(.24)	(.35)
Spelling	2.38 ***	1.64 ***	0.84	2.51 ***	1.18 *	1.04	-0.24	-0.69 *	-1.28 **
	(.50)	(.41)	(.50)	(.65)	(.56)	(.71)	(.40)	(.34)	(.46)
Maths [a b]	1.22 ***	1.50 ***	2.04 ***	8.14 ***	7.95 ***	9.39 ***	2.71 ***	3.04 ***	3.28 ***
	(.29)	(.22)	(.27)	(.38)	(.29)	(.37)	(.23)	(.18)	(.26)
Controls	X	X	X	X	X	X	X	X	X
Obs	2080	3161	2264	2076	3153	2267	2079	3161	2272
R^2	0.62	0.58	0.60	0.65	0.60	0.63	0.55	0.52	0.52

* p <.05, ** p <.01, *** p <.001

[a] Low Education coefficient is significantly different from High Education coefficient for Key Stage 1 English at p<.05

[b] Low Education coefficient is significantly different from High Education coefficient for Key Stage 1 Maths at p<.05

Moderation by social background: Sensitivity analysis

The finding that KS1 results are more predictive for children from households with lower levels of parental education may at first glance appear to contrast with previous work on the importance of development in the pre-school years, which finds that children from low SES families fail to realise the benefits of positive development between the ages of 3 and 5 (Feinstein & Duckworth, 2006). To test whether this result is just picking up the fact that the KS1 maths assessment is particularly good at discriminating amongst those in this low household education group between those with more and less achieved intelligence, we condition for approximately concurrent IQ, measured at age 8 (mean = 104.24, S.D. = 16.4), to see if this effect still remains.

The results of this sensitivity check are summarised in Table 19 and suggest that IQ is not the mechanism for the observed stability between KS1 and KS2. Furthermore, the general pattern of results for the sub-group analyses holds and, for KS2 English, the difference in the effect of KS1 maths between those from lower educated households remains statistically significantly higher. These findings indicate that it is more than just underlying intelligence driving stability in attainment during primary school. Rather, doing well in these early assessments seems to act as some kind of signal, heightening self-efficacy or aspirations for the pupil and/or an increasing support offered by teachers that translates into later successful KS2 performance. We discuss these findings and their possible implications in greater detail below.

Table 19: Coefficients and standard errors from regression models of Key Stage 2 scores on Key Stage 1 tests, controlling for IQ, by parents' highest level of education

	English			Maths			Science		
	High	Med	Low	High	Med	Low	High	Med	Low
Key Stage 1 tests:									
Reading	2.86 ***	2.43 ***	3.49 ***	0.26	0.45	1.00	1.32 ***	1.15 ***	1.18 **
	(.38)	(.32)	(.51)	(.48)	(.43)	(.68)	(.28)	(.25)	(.44)
Writing	2.58 ***	3.03 ***	2.12 ***	0.75	1.58 **	0.96	0.51	0.06	1.16 *
	(.43)	(.38)	(.59)	(.52)	(.50)	(.91)	(.31)	(.28)	(.53)
Spelling	2.56 ***	2.05 ***	1.74 *	2.39 ***	0.87	1.49	-0.37	-0.14	-0.81
	(.61)	(.54)	(.81)	(.78)	(.75)	(1.13)	(.46)	(.43)	(.69)
Maths [a]	0.01	0.37	1.18 **	6.18 ***	6.48 ***	7.30 ***	1.72 ***	1.89 ***	1.73 ***
	(.37)	(.31)	(.42)	(.48)	(.41)	(.62)	(.30)	(.24)	(.39)
WISC IQ	0.18 ***	0.21 ***	0.23 ***	0.31 ***	0.37 ***	0.42 ***	0.18 ***	0.25 ***	0.31 ***
	(.02)	(.02)	(.03)	(.03)	(.03)	(.05)	(.02)	(.02)	(.03)
Controls	X	X	X	X	X	X	X	X	X
Obs	1501	1824	1077	1495	1820	1076	1498	1820	1083
R^2	0.66	0.68	0.72	0.70	0.68	0.73	0.62	0.63	0.69

* p <.05, ** p <.01, *** p <.001

[a] Low Education coefficient is significantly different from High Education coefficient for Key Stage 1 English at *p<.05*

4. Discussion

This study explored Key Stage assessments between the ages of 7 and 11 years to better understand progression over the primary school years and the advantages that pupils gain from higher attainment during the first years of formal schooling, above and beyond ability at entry to school. Our primary focus was in estimating the importance of progress over the Key Stage 1 period as it relates to attainment in Key Stage 2 assessment. To do this, we examined the extent to which the change in literacy and numeracy skills measured at entry to school and Key Stage 1 mattered for subsequent Key Stage 2 success. In this final section, we discuss the substantive themes that emerge from these results in more detail and put forward possible interpretations for what they might mean in terms of policy and practice. We also raise some other points that emerge from these findings that pose further questions and require additional analysis.

4.1 Stability and change in primary school attainment

First and foremost, the results of this study emphasise the high degree of stability in attainment as measured by the Key Stage assessments. Examination of the correlations, transition matrices, and regression modelling used here, all indicate a great deal of stability in attainment in the primary school National Curriculum tests. The majority of pupils are advancing the expected two levels over the Key Stage 2 period, more so at the top end of the Key Stage 1 distribution, and, in these data, 78% of pupils are on target in all three Key Stage 2 assessments. The results of the multiple regression analyses also highlight that the substantial continuity observed here is fairly independent of wider developmental contexts. The four Key Stage 1 scores alone account for 49%, 52% and 40% of the variance in Key Stage 2 English, maths and science assessments respectively and when comprehensive controls for family background, parenting, earlier child-level skills and capabilities and school fixed-effects are entered into the regression estimation, the proportion of variance explained does not substantially increase. That is not to say that features of the child and their family background do not matter for attainment in primary school, but that the prior attainment of pupils measured by Key Stage 1 tests contributes the most in explaining the variance in their performance at Key Stage 2. We discuss the importance of other influences further in section 4.3.

At an average level then, progression between Key Stage 1 and 2 appears remarkably stable: in terms of position in the distribution, the majority of pupils remain in the same quartile at Key Stage 2, as they were in Key Stage 1. However detailed consideration of the transition matrices suggests that there is also change and discontinuity in pupils' progression during primary school. That is, although the majority of pupils are advancing the expected two levels over the Key Stage 2 period, there are others who buck this trend, both exceeding and falling off these average "target" trajectories. A small percentage of pupils, for example, who at Key Stage 1 were only "Working towards Level 1", manage to advance considerably over Years 3 – 6 and are on target, i.e. Level 4 or above, in the Key Stage 2 assessments. Equally, some pupils, even those on target at Key Stage 1, are dropping down in the distribution.

Thus, while the average picture here is one of stability in attainment, there is also mobility in academic achievement and attainment trajectories are not set in stone. This is unlikely to be adversely affected by measurement error, for example, being unwell or having a bad day on the day of the test, since mobility is evident throughout the distribution and not systematic at any one point. The high correlations between the level awarded by the test and the teacher-based assessments also indicate that it is unlikely that error in the assessment data drives these results. Furthermore, the comprehensive controls afforded by the rich, longitudinal information in the ALSPAC enable us to condition out earlier ability and other background characteristics that might also influence Key Stage 2 performance. The estimates reported here are therefore considerably more robust than those controlling for PLASC information alone, i.e. special educational needs status, free school meal claimants and English as a first language.

These findings support earlier work by Feinstein (2004) which reports similar mobility in the relative position of pupils' attainment in primary and secondary school for the 1958 National Child Development Study (NCDS) and 1970 British Cohort Study (BCS). Again, using quartile-based transition matrices, he shows that, in the NCDS, 35 per cent of those in the bottom quartile of general academic ability at age 7, have "escaped", i.e. are no longer in, the bottom quartile by age 11. Conversely the probability of "dropping out" of the top quartile between 7 and 11 is 44 per cent. For the 1970 cohort, the escape rate from the bottom quartile between 5 and 10 years was 46 per cent and the drop out rate from the top quartile was 50 per cent. The results of the current study show a slightly greater likelihood of escaping the bottom quartile, with 50 per cent of pupils moving out of the bottom quartile by Key Stage and a lower probability, 35.5 percent, of exiting the top quartile. This may reflect genuine differences in attainment mobility over time, but is more likely to reflect differences in the type of data used: the Key Stage 1 data are bound by discrete levels of attainment rather than being a truly continuous score and so differ from the cognitive tests of ability used in the cohort studies which are expressed in absolute rather than relative terms.

Finding both change and stability in primary school attainment is not surprising and is expected from the literature reviewed in our Introduction concerning the developmental continuities and discontinuities that are so prominent in middle childhood. It does, however, highlight the need to understand better the balances of stability and change, continuity and discontinuity in order that all children are to be able to make the most of their learning opportunities as well as make the expected levels of progression across their school careers. Recognition and appreciation of the shifts and fluctuations that take place during primary school are central to a focus on personalised learning and fundamental to closing the attainment gaps.

4.2 Effects within and between subjects

In line with other research (Melhuish, *et al.,* 2006), our second major finding relates to the strength of within-subject influences of pupils' literacy and numeracy skills. We find that performance in a given subject at Key Stage 1 is highly predictive of performance at Key Stage 2 in that subject, reflecting the high degree of stability in pupils' trajectories noted above. Reading, writing and spelling at Key Stage 1 are

more important for Key Stage 2 English than maths and, likewise, doing well in maths is the best gauge for subsequent maths performance, more than four times as important as Key Stage 1 writing and over five times more important than the earlier assessment of reading. Literacy and numeracy skills thus seem to become increasingly entrenched in their subject specific domains as children move through primary school.

Nevertheless, we do find evidence of smaller, but statistically significant cross-over effects between skill areas, with performance in Key Stage 1 maths showing a positive relationship with Key Stage 2 English and science results. Similarly, performance in Key Stage 1 reading and writing both have positive associations with Key Stage 2 attainment in maths and science. Key Stage 2 science scores are more strongly related to Key Stage 1 reading than writing skills. However, the benefit of increases in attainment in Key Stage 1 maths for and Key Stage 2 maths does not depend on attainment in Key Stage 1 English tests. The gains in Key Stage 1 attainment for Key Stage 2 performance are additive and do not create complementarities across subjects over time. For example, the benefit of increases in attainment in Key Stage 1 maths for subsequent Key Stage 2 English performance is the same for low and high levels of attainment in Key Stage 1 English-related tests.

The counter-intuitive negative relationship between Key Stage 1 spelling and Key Stage 2 science may also follow from the subject specialisation discussed above. For example, it is possible that once the effects of reading, writing and maths have been taken account of, doing well in spelling has no further benefit for a broader subject such as science which is influenced more equally by earlier literacy and numeracy skills. Other possible explanations as to why spelling is the least predictive of all the Key Stage 1 skills may also reflect who takes the test (pupils typically have to achieve Level 2 in Key Stage 1 writing to take the spelling test), the fewer number of attainment levels are awarded for spelling than for the other Key Stage 1 tests (three rather than the six), and the resulting predictive importance of spelling when pupils are 6 to 7 years old.

It is interesting to compare these findings on domain specificity with those from previous work by the Centre on early development and school readiness (Feinstein & Duckworth, 2006). In that study positive development in cognitive ability between the ages of 3 and 5, measured using a copying test, was shown to be beneficial for both reading and maths at age 10, and yielded higher gains than, for example, did early development in vocabulary. As research shows that early developmental skills lay the foundations for deeper understanding of more complex skills and problem solving abilities, we argued our results reflect that, in the pre-school period, the copying test better tapped the foundations of both pre-literacy and pre-numeracy skills, such as decoding, upon which later abilities are built. As children move from early childhood into middle childhood and gain increasing experiences of formal education, they become more adept at acquiring new information and using this to develop more mature modes of thought. Consequently, specific literacy and numeracy skills become more advanced over the 7 – 11 Key Stage 2 period and thus show greater within-domain continuity over time. Feinstein (1998) also finds evidence of cross-domain skills in the early years that lessen with experience of formal schooling.

4.3 Who gets the benefits? Sub-group differences

The final theme emerging from these results concerns differences in who realises the benefits of early attainment. We find that for all Key Stage 2 assessments, the prior attainment of pupils is the most important predictor of later performance and that adding in controls for child, family and other background characteristics does not substantially change the size of the Key Stage 1 effects or increase the proportion of variance accounted for. However, our sub-group analyses do find significant moderation by socio-economic characteristics and differences in attainment by gender.

Our results highlight the important influence of social background on attainment as children progress through schools and, in line with other research, further demonstrate that the attainment gap is not simply a one-off phenomenon evident at entry to school but continues throughout pupils' educational trajectories. For example, pupils whose parents have higher levels of education are more likely to remain at the top of the distribution at the end of Key Stage 2, conditional on being in the top 25% at Key Stage 1, than those from households where parents' education is lower. The converse is true for bottom quartile continuity: children are less likely to remain in the bottom quartile in their Key Stage 2 assessments given relative position in the bottom quartile at Key Stage 1 if their parents have higher levels of education themselves.

However, doing well in maths is more important (i.e. a better predictor of later success) for children from more socially disadvantaged backgrounds. The results of the multiple regression sub-group analysis reveal that the returns for Key Stage 1 maths in relation to both Key Stage 2 English and maths are statistically significantly higher for children from households with low levels of parental education than for children from well-educated households. In other words, although children from more advantaged backgrounds are, on average, doing better in both the Key Stage 1 and Key Stage 2 assessments, doing well in Key Stage 1 maths carries particular gains for children from less advantaged families. On average, for a child from a low-education household, an increase of one level in the Key Stage 1 test results in an extra 2 points in Key Stage 2 English and an extra 9.4 points in Key Stage 2 maths. For a child from a well-educated household, the increases at Key Stage 2 for each level increase at Key Stage 1 are 1.2 points in English and 8.1 points in maths respectively. The converse of this, however, is that doing badly in Key Stage 1 maths has a particularly negative impact on Key Stage 2 outcomes.

This result may appear at first glance to conflict with previous work by the Centre (Feinstein & Duckworth, *op cit*) which explored, in different data, the benefit that children gain from arriving at school with particular personal characteristics that have developed in the previous pre-school years, above and beyond innate capabilities, for subsequent school success as well as for later adult outcomes. In our earlier study, the same sub-group analyses revealed that low SES children with high copying skills, an early test of cognitive ability, were not realising the same positive benefits for either maths or reading measured at age 10 and for the adult outcomes considered.

Why, unlike these earlier findings, might this signal of early ability carry through to attainment at the end of primary school? While it is unlikely to be simply an artefact of historical shifts in mobility, it may, in part, reflect differences in both the age of

children and the developmental period considered as well as the nature of the assessment itself. For example, in the earlier study, the cognitive tests were administered as part of the study design, not as individual records of performance, and so, unlike the Key Stage assessments examined here, the outcomes were not known to children, their parents and, later, their teachers. It may be then, that some element of self-, teacher- or parent-knowledge of Key Stage 1 attainment might create the observed continuity.

We tested this hypothesis by entering IQ into our regression estimation in an attempt to condition out a broader measure of achieved intelligence measured around the age of the Key Stage 1 assessments. These sensitivity results suggest that IQ is not the mechanism for an effect of Key Stage stability as, when it is entered as a control, Key Stage 1 attainment continues to be a better predictor of later success for children from household with lower levels of parental education. Moreover, for Key Stage 2 English the difference in the effect of Key Stage 1 maths between those from lower educated households remains statistically significantly higher. Thus it is not just general intelligence or broader cognitive functioning that drives these results. This lends support to our hypothesis that knowledge and awareness of the assessment outcomes may be an important feature of evaluating understanding and monitoring performance in primary school.

The importance of children's perceptions of primary school assessments and how they contribute to their understanding and interpretation of themselves as learners is demonstrated in a qualitative study by Reay and Wiliam (1999). Interviews with children in year 6 reveal, for example, children's anxiety about failing their SATs, how success or failure in these assessments says something about their own intrinsic worth and is conflated with "goodness", cleverness and future prospects, as well as serving as informal assessments of positive and negative attributes of their peers. The authors argue that both pupil and teacher identities and practices are modified through the Key Stage assessment processes and that children's voices themselves are a particularly salient part of any picture investigating the importance and social consequences of measures of educational effectiveness.

The differences between our two studies may also reflect the quality of the parenting resources available to less advantaged parents. That is, children from lower socio-economic backgrounds who were showing positive development between ages 3 and 5 and were scoring highly in the copying test may fail to carry this through to age 10 outcomes because of lower levels of parental input and lack of scaffolding of these skills during the pre-school years. However, if once in school, this ability is recognised, for example signalled by positive achievement in the Key Stage 1 assessments in year 2, the resulting support and learning opportunities provided by teachers can ensure that early positive development translates into later academic success.

There is considerable social stigma attached to being "bad" at maths, both for children and adults, which is often seen as interchangeable with intelligence. Achievement in maths then, may be seen as a more objective measure of ability than achievement in other subjects and the related value that children, teachers and parents attribute to doing well in these tests may, therefore, further increase the benefit of higher attainment in Key Stage 1 assessments. More in-depth analysis of teacher-pupil

processes in the classroom and the interactions between family and school contexts is required to unpack why attainment (or underachievement) in maths may be a particularly important signal for later academic success.

Our results also suggest a greater tendency for girls to be "upwardly mobile" in English and boys to be so in maths. However, this does not result in a significant difference in the predictability of Key Stage 2 results based on Key Stage 1: Key Stage 1 results are equally predictive of Key Stage 2 success (or failure) for both boys and girls. Rather, it reflects the differential attainments whereby girls have tended to do better in English and literacy-related skills and boys in maths.

4.4 Formative assessment and the broader learning experience

The research put forward here builds on earlier work by the Centre that explored development in the pre-school years and reports some complementary findings as well as raising some interesting questions about middle childhood development and the nature of primary school assessments. We know that Key Stage 1 tests are the best predictor of subsequent attainment at Key Stage 2, but we also know that there is both stability and change in middle childhood. To best support progression in learning, practitioners and policy makers need to have a clear understanding of this balance. This research is about how we interpret this finding recognising that this statistical correlation masks a lot of underlying discontinuity in how children perform. The stability is important because early achievement matters and is the foundation of subsequent success. Change and discontinuity are also important because if people believe that ability is fixed and innate, then the response to poor educational performance, both by teachers and pupils, will be to disengage as nothing can be done. Understanding and appreciating the substantial malleability in performance that is observed in this report will enable a broader appreciation of the ever-present potential for improvement and the performance and capability of pupils and learners.

Critics of the National Curriculum assessments, particularly those administered in primary school, have questioned their value and purpose, arguing that they provide simplistic judgements about pupils' ability and act only as "performance indicators of teacher effectivity" (Ball, 1994, p. 41). They argue that over-assessment of pupils can adversely affect child well-being and impact on educational engagement and intrinsic motivation, as well as encourage "teaching to the test". However, the results of the current study suggest that they may provide useful signals both for pupils doing well as well as those who might be at risk of under achieving. In particular, failure to place sufficient emphasis on basic numeracy skills may run counter to the interests of children from socially disadvantaged backgrounds. For example, we hypothesise that doing well in Key Stage 1 maths is particularly beneficial for children from more socially disadvantaged background because the public awareness of the test result and corresponding self-knowledge signals ability that might otherwise go undetected. In this regard, we believe that a greater understanding about why early maths carries such weight in predicting success at later scholastic attainment is warranted, particularly for those who might otherwise be at risk of underachievement. This is likely to require a greater appreciation of children's own interpretations of what SATs mean, how they are valued by children and teachers, as well as a more in-depth understanding of how different configurations of school and family contexts relate to

stability and change in attainment.

While we believe that the results presented here lend support to calls for personalised learning, we are cautious of interpretations that might recommend the increased use of "educational triage" to concentrate on boosting the performance of children just under the National Curriculum target levels in an effort to increase the numbers of children reaching this threshold. Rather, we recommend a more holistic approach for children across the full spectrum of achievement. Pedagogy and assessment must continue to work alongside each other to provide more detailed pictures of the sorts of learners that individual pupils are and can be, as well as what they have learnt. Assessment practices, both summative and formative, need to be further developed to monitor progress and identify the next sets of learning objectives in order that they support learning as well as provide meaningful measures of educational effectiveness. Strategies that enable pupils to see how they are progressing in frequent and meaningful ways increase engagement and motivation and help develop pupils' appetites for and attitudes to lifelong learning.

Further research is required to establish the longitudinal stability and validity of the key stage curriculum assessments and whether the high degree of stability observed in these assessments signals genuine continuity in development and resulting attainment or is merely a product of common assessment. Future research also needs to focus on how different children progress as they make the transition into secondary school and whether the results observed here hold for Key Stage 3 attainment. In particular, we are interested to see how much the compounding effects of social background continue, examine in further detail what predicts the sub-group differences observed here, and explore change and stability in different subjects.

Finally, academic attainment is only one facet of educational success and only part of the Government's Every Child Matters agenda. Improvements in problem behaviour or social skills may better predict other important school outcomes, such as a child's engagement in school and motivation for learning, relationships with peers and teachers as well as their overall self-concept and school adjustment. These other features of school success need to be reflected in configurations of pupil abilities so as to understand individual learning trajectories more fully.

References

Aunola, K., Leskinen, E., Lerkkanen, M., & Nurmi, J. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*, 699-713.

Ball, S. (1994). *Education reform: a critical and post-structural approach.* Buckingham: Open University Press.

Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructing Adaptive Expertise Studies*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology, 65*, 1-24.

Collins, W. A. (1984). *Development during middle childhood*. Washington, DC, National Academy Press.

Department for Children, Schools and Families (2007). *Building on Achievement, Meeting New Challenges*. Press Notice, 10 July 2007. http://www.dcsf.gov.uk/pns/DisplayPN.cgi?pn_id=2007_0126

Department for Education and Skills. (2004). *Five Year Strategy for Children and Learners.* (CM 6272). London: HMSO.

Department for Education and Skills. (2005). *National Curriculum Assessment, GCSE and equivalent attainment and post-16 attainment by pupil characteristics in England.* 2004: www.dfes.gov.uk/rsgateway/DB/SFR/s000564/SFR08-2005v2.pdf.

Department for Education and Skills. (2006). *Making Good Progress: How can we help every pupil to make good progress at school?* Consultation document.

Department for Education and Skills. (2006b). *Social Mobility: Narrowing social class educational gaps*: www.dfes.gov.uk/rsgateway/DB/STA/t000657/index/shtml.

Department for Education and Skills. (2007). *2020 Vision*. Report of the Teaching and Learning in 2020 Review Group. http://www.teachernet.gov.uk/docbank/index.cfm?id=10783.

Department for Education and Skills. (2007b). *Gender and education: The evidence on pupils in England*: HMSO.

Entwisle, D. R., & Alexander, K. L. (1990). Beginning School Math Competence: Minority and Majority Comparisons. *Child Development, 61*, 454-471.

Feinstein, L., & Duckworth, K. (2006). *Ready for School? The components of school readiness and their importance for school performance and adult outcomes* (No. 20). London: Centre for Research on the Wider Benefits of Learning.

Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 73-98.

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

Ferrari, M., & Sternberg, R. J. (1998). The development of mental abilities and styles. In W. Damon (Ed.), *Handbook of Child Psychology: Child Psychology and Practice* (5th ed., Vol. 2, pp. 899-946). New York: John Wiley & Sons.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35,* 39-50.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74,* 213-239.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345-362.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number.* Cambridge, MA: Harvard University Press.

Gutman, L. M., & Feinstein, L. (2007). *Parenting behaviours and children's development from infancy to early childhood: Changes, continuities and contributions.* (No. 22). London: Centre for Research on the Wider Benefits of Learning.

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relation between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology, 79*, 192-227.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction, 14*, 251-283.

http://www.qca.org.uk/eara/documents/KS2_ARA.pdf

Kail, R., & Hall, L. K. (1999). Sources of developmental change in children's word-problem performance. *Journal of Educational Psychology, 91*, 660-668.

Kowaleski-Jones, L., & Duncan, G. J. (1999). The Structure of Achievement and Behavior across Middle Childhood. *Child Development, 70(4),* 930-943.

Kulak, A. G. (1993). Parallels between math and reading disability: Common issues and approaches. *Journal of Learning Disabilities, 26*, 666-668.

Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J. (2006). Development of reading and spelling Finnish from preschool to Grade 1 and Grade 2. *Scientific Studies of Reading, 10(1)*, 3-30.

Lerkkanen, M., Rasku-Puttonen, H., Aunola, K., & Nurmi, J. (2005). Mathematical performance predicts progress in reading comprehension among 7-year olds. European *Journal of Psychology of Education, 20(2),* 121-137.

Macaruso, P., & Sokol, S. M. (1998). Cognitive neuropsychology and developmental dyscalculia. In C. Donlan (Ed*.), The development of mathematical skills* (pp. 201-225). East Sussex: Psychology Press.

Machin, S., & McNally, S. (2003). *The Literacy Hour*. Centre for the Economics of Education, London School of Economics. Discussion Paper No. 43*

McCall, R. B., Appelbaum, M., & Hogarty, P. S. (1973). Developmental changes in mental performance. *Monographs for the Society for Research in Child Development, 38*(3, Serial No. 150).

Melhuish, E., Romaniuk, H., Sammons, P., Sylva, K., Siraj-Blatchford, I., & Taggart, B. (2006). *The effectiveness of primary schools in England in Key Stage 2 for 2002, 2003 and 2004.* London: DfES / Institute of Education, University of London.

Muthén, B. O., & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, 10*, 73-102.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading, 27(4)*, 342-356.

Nesher, P. (1986). Learning mathematics: A cognitive perspective. *American Psychologist, 41,* 1114-1122.

NICHD, E., Child, Care, Research Network,. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428-442.

Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology, 94(1),* 3-13.

Pungello, E. P., Kupersmidt, J. B., Burchinal, M. R., & Patterson, C. J. (1996). Environmental Risk Factors and Children's Achievement from Middle Childhood to Early Adolescence. *Developmental Psychology, 32(4),* 755-767.

Reay, D., & Wiliam, D. (1999). "I'll be a nothing": Structure, agency and the construction of identity through assessment. *British Educational Research Journal, 25(3)*, 343-354.

Rourke, B. P., & Conway, J. A. (1997). Disabilities of arithmetic and mathematical reasoning: Perspectives from neurology and neuropsychology. *Journal of Learning Disabilities, 30,* 34-46.

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy Research*. New York: The Guilford Press.

Share, D. L., Moffitt, T. E., & Silva, P. A. (1988). Factors associated with arithmetic-and-reading disability and specific arithmetic disability. *Journal of Learning Disabilities, 21,* 313-320.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, D.C.: National Research Council, National Academy Press.

South Gloucestershire Professional & Curriculum Support Service. (1996). *Entry Assessment in South Gloucestershire*. Weston-super-mare: Woodspring Educational Resource Centre.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934-947.

Strand, S. (1999). Ethnic group, sex and economic disadvantage: Associations with pupils' educational progress from Baseline to the end of Key Stage 1. *British Educational Research Journal, 25(2)*, 179-202.

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification ability on children's work problem solution. *Journal of Experimental Child Psychology,* 55, 374-395.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848-872.

Williamson, G. L., Appelbaum, M., & Epanchin, A. (1991). Longitudinal analyses of academic achievement. *Journal of Educational Measurement, 28*, 61-76.

Wilson, R. S. (1983). The Louisville Twin Study: Developmental Synchronies in Behaviour. *Child Development, 54*, 298-316.

Appendices

Appendix Table 1: Coefficients and standard errors from regression models of Key Stage 2 English scores on Key Stage 1 tests, full stepwise model

	Key Stage 2 English					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Key Stage 1 tests						
Reading	3.75 ***	3.45 ***	3.38 ***	3.34 ***	3.23 ***	3.19 ***
	(.14)	(.13)	(.13)	(.13)	(.13)	(.13)
Writing	3.67 ***	3.15 ***	3.11 ***	3.06 ***	2.93 ***	3.02 ***
	(.16)	(.16)	(.16)	(.16)	(.16)	(.16)
Spelling	1.25 ***	1.59 ***	1.65 ***	1.64 ***	1.69 ***	1.79 ***
	(.23)	(.23)	(.23)	(.23)	(.23)	(.23)
Maths	1.94 ***	1.66 ***	1.68 ***	1.68 ***	1.47 ***	1.58 ***
	(.12)	(.12)	(.12)	(.12)	(.12)	(.12)
Controls: Distal factors		X	X	X	X	X
Controls: Maternal chars & prox process			X	X	X	X
Controls: Pre-school development				X	X	X
Controls: Entry Assessments					X	X
Controls: School fixed effects						X
Obs	9142	9142	9142	9142	9142	9142
R^2	0.49	0.53	0.54	0.54	0.55	0.58

Appendix Table 2: Coefficients and standard errors from regression models of Key Stage 2 maths scores on Key Stage 1 tests, full stepwise model

	Key Stage 2 Maths											
	(1)		(2)		(3)		(4)		(5)		(6)	
Key Stage 1 tests												
Reading	1.78	***	1.46	***	1.47	***	1.45	***	1.26	***	1.32	***
	(.18)		(.18)		(.18)		(.18)		(.18)		(.18)	
Writing	2.77	***	2.17	***	2.15	***	2.18	***	2.00	***	1.85	***
	(.21)		(.21)		(.21)		(.21)		(.21)		(.22)	
Spelling	1.10	***	1.47	***	1.53	***	1.56	***	1.62	***	1.78	***
	(.31)		(.31)		(.31)		(.31)		(.31)		(.31)	
Maths	8.92	***	8.55	***	8.49	***	8.40	***	8.02	***	8.29	***
	(.16)		(.16)		(.16)		(.17)		(.17)		(.17)	
Controls: Distal factors			X		X		X		X		X	
Controls: Maternal chars & prox process					X		X		X		X	
Controls: Pre-school development							X		X		X	
Controls: Entry Assessments									X		X	
Controls: School fixed effects											X	
Obs	9123		9123		9123		9123		9123		9123	
R^2	0.52		0.55		0.55		0.56		0.57		0.60	

Appendix Table 3: Coefficients and standard errors from regression models of Key Stage 2 science scores on Key Stage 1 tests, full stepwise model

	Key Stage 2 Science											
	(1)		(2)		(3)		(4)		(5)		(6)	
Key Stage 1 tests												
Reading	2.43	***	2.13	***	2.09	***	2.07	***	1.96	***	1.95	***
	(.11)		(.11)		(.11)		(.11)		(.11)		(.11)	
Writing	1.75	***	1.27	***	1.23	***	1.19	***	1.08	***	1.01	***
	(.13)		(.13)		(.13)		(.13)		(.13)		(.13)	
Spelling	-1.13	***	-.84	***	-.85	***	-.84	***	-.79	***	-.68	***
	(.20)		(.20)		(.20)		(.20)		(.20)		(.20)	
Maths	3.46	***	3.17	***	3.16	***	3.12	***	2.91	***	3.04	***
	(.11)		(.11)		(.11)		(.11)		(.11)		(.11)	
Controls: Distal factors			X		X		X		X		X	
Controls: Maternal chars & prox process					X		X		X		X	
Controls: Pre-school development							X		X		X	
Controls: Entry Assessments									X		X	
Controls: School fixed effects											X	
Obs	9152		9152		9152		9152		9152		9152	
R^2	0.40		0.44		0.45		0.46		0.47		0.51	

Appendix Table 4: Coefficients and standard errors from regression models of Key Stage 2 scores on Key Stage 1 tests, by gender

	Key Stage 2 English		Key Stage 2 Maths		Key Stage 2 Science	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Key Stage 1 tests						
Reading	3.10 ***	3.30 ***	1.23 ***	1.42 ***	1.81 ***	2.11 ***
	(.19)	(.20)	(.26)	(.27)	(.18)	(.16)
Writing	3.13 ***	2.83 ***	2.07 ***	1.71 ***	1.27 ***	0.75 ***
	(.23)	(.26)	(.31)	(.33)	(.20)	(.20)
Spelling	1.75 ***	1.92 ***	1.68 ***	1.96 ***	-0.23	-0.98 ***
	(.33)	(.34)	(.45)	(.46)	(.29)	(.28)
Maths	1.53 ***	1.65 ***	8.21 ***	8.26 ***	3.07 ***	2.98 ***
	(.18)	(.19)	(.24)	(.25)	(.16)	(.16)
Obs	4709	4433	4698	4425	4705	4447
Controls	X	X	X	X	X	X
R^2	0.60	0.58	0.64	0.60	0.55	0.53

WIDER BENEFITS OF LEARNING RESEARCH REPORT NO.23

What role for the three Rs?

Progress and attainment during primary school

While there is a wealth of data showing that children's achievements in test scores are strongly related to their prior cognitive functioning and attainment of basic skills in numeracy and literacy, children also demonstrate shifts and fluctuations in the trajectories of their skills and abilities, particularly during the years covering primary school.

This study builds on earlier work by the Centre for Research on the Wider Benefits of Learning, and explores the balance of change and stability in children's school test scores during primary school. We consider how the foundation stones of literacy and numeracy operate to support progress in Key Stage attainment, looking at within- and between-subject effects for literacy, numeracy and science.

In line with other research, we find that Key Stage 1 tests are the best predictors of attainment at Key Stage 2. Overall, there is substantial stability over the Key Stage 2 period, but also mobility, both upward and downward, in pupils' performance. Across subjects, the majority of pupils advanced the expected two levels of progress between Key Stage 1 and Key Stage 2 as laid out in the National Curriculum, with some exceeding these expectations. However, approximately a quarter of pupils failed to progress the expected two levels.

Progress and attainment were also moderated by socio-economic factors. Children from more socially-advantaged homes do better, on average, in both Key Stage 1 and 2 assessments. However, for children of parents with lower levels of education, doing well in Key Stage 1 tests, particularly in maths, is more important (i.e. more predictive of later attainment) than for other groups.

This project highlights the importance of appreciating both change and stability in attainment during primary school across the full spectrum of achievement. Stability is important because early achievement matters and is the foundation of subsequent success. But equally, recognising change and discontinuity is also necessary because if people over-estimate the extent to which ability is fixed and innate, then the response to poor educational performance, both by teachers and pupils, will be to disengage excessively.

The results presented here lend support to calls for personalised learning, in which assessment practices – both formative and summative – are developed to provide a more complete picture of children's development and are used to support learning as well as measure it. Strategies that enable pupils to see, in frequent and meaningful ways, how they are progressing, will increase engagement and motivation and help develop their appetites for lifelong learning.

Kathryn Duckworth is a Research Officer at the Centre for Research on the Wider Benefits of Learning.

ISBN 978-0-9552810-3-7