The long-term impact of effective teaching

Abstract

This paper investigates the impact of effective schooling in the first year of elementary school on later academic outcomes and equal educational opportunity.

A large longitudinal dataset from England was used to estimate the importance of the first year of elementary school for academic outcomes up to age 16. Multi-level models, controlling for baseline assessment, deprivation, sex and ethnic status showed that classes in the first year differed substantially in their progress but did not vary in their impact on equity. Those classes defined as effective and students from those classes were tracked on three further occasions up to the age of 16 and compared with others.

Being in an effective class in the first year of school, when the children were aged 4 – 5 years, was significantly related to later attainment at age 16 (Effect Size = 0.2). However, it was unrelated to equity at age 16.

Key words: first school year, longitudinal, equity, impact, effectiveness, elementary school

Introduction

Links between early childhood development, home background and later outcomes in life are extensively documented. Development in the first few years of life is rapid and it has long been suggested that the earlier interventions are implemented, the better (see, for example Farrington, 1994). Longitudinal studies, which have followed children who received pre-school interventions, have quantified the impact of high-quality provision on outcomes right into adulthood. Less is known about the impact of an effective first year of school, which is a period of rapid cognitive development; Tymms et al. (2016) found that, in the UK, many children start school with some knowledge of letter and number identification, able to count and manipulate small quantities, and leave that first year able to read simple sentences, perform calculations and solve a variety of mathematical problems. The present study builds upon earlier research, and suggests that an effective first year of school, where children have made more progress than their peers in other schools, has a long-lasting impact. Such a finding has important policy considerations to ensure that this phase of education is well supported with high-quality staff and resources.

Previous Research

Early development in the first five years of life lays the foundations for lifelong learning (Shonkoff and Phillips, 2000; Feinstein and Duckworth, 2006), for later academic success in school (Duncan et al., 2007), at college (McClelland, Acock, Piccinin, Rhea and Stallings, 2013) and health outcomes (Pagani and Fitzpatrick, 2013). The effects of disparities in home background, care and educational opportunity between children from disadvantaged and more affluent backgrounds are apparent amongst even the very young (Hart and Risley, 2003; Lee and Burkham, 2002). More broadly, the importance of school characteristics in terms of facilities, teachers, materials and curriculum as a means of providing equal opportunities and potential impact on skills and knowledge have been the subject of large scale studies

(Coleman et al., 1966; Borman and Dowling, 2010) and about 50 years of school improvement and school effectiveness research (Scheerens 2017).

Educational provision and care present opportunities to counter the risks of poor outcomes in later life. A high-quality pre-school has been causally linked to later outcomes and this is picked up later. But it is noted here that the pre-school environment has recently been positively associated to later academic outcomes. The Effective Pre-school Primary and Secondary Education (EPPSE) project in England investigated the influence of pre-school on later outcomes (Sammons et al., 2014). The study sample consisted of 3,110 children, of which 2,800 attended a pre-school setting in England. At school entry, pre-school attendance was found to be positively related with academic, social and behavioural outcomes, and the number of months spent in pre-school was important. Attending a pre-school was linked to higher examination results at the end of compulsory schooling[1] in English and mathematics (Effect Size = 0.23 and 0.21 respectively). Attendance at an effective pre-school was associated with higher GCSE English and mathematics grades (Effect Size = 0.31 and 0.35 respectively). In a longitudinal study which followed 1364 children up to Grade 6, age 12 years. Belsky et al. (2007) found a positive association between high quality early childhood care of any kind, and later vocabulary. They also found a correlation between time spent in pre-school centre settings and problem behaviours at age 12, with those children who spent more time in centre settings showing more problem behaviours, which is a concerning outcome. The authors noted the limitation of the sample not being nationally representative.

Loeb and Bassock (2007) presented evidence in support of the deeply entrenched and dramatic socio-economic gap between children from affluent and deprived backgrounds as they enter kindergarten. They suggested that this disparity emerges in toddlers as young as

[1] General Certificate of Secondary Education (GCSE) is the examination taken at age 16.

eighteen months and widens throughout early childhood. Many early years interventions have focused on children in deprived circumstances because of this link between socio-economic status and academic achievement (see, for example Bourdieu & Passeron, 1977; Raffo *et al*., 2007). Additionally, development during early childhood is particularly malleable and studies have repeatedly shown interventions implemented during this period to yield long-lasting, cost-effective impact (Heckman, 2006). The Perry Pre-school Program is often cited as an effective early intervention for children living with disadvantage, with Schweinhart et al., (2005) reporting long-term beneficial effects well into adulthood. Heckman et al., (2010) re-examined the findings and concluded that there was still a significant positive cost-benefit to the program. Duncan and Magnuson (2013) found long-term positive impact for many pre-school programs. However, Ramey and Ramey (1998) cautioned that early interventions should be targeted to needs and that not all are guaranteed to be successful. They discussed the mechanisms by which early interventions may contribute to long-term success including increasing a child's ability and skills to gain more from later experiences, motivation such that a child seeks out advantageous learning experiences and access to more supportive learning environments. Ramey and Ramey (1998) proposed that interventions in the early years, such as high-quality preschool provision, are insufficient by themselves to maintain the positive gains seen in children from disadvantaged backgrounds. Whilst these gains do not disappear entirely, they do fade over time and therefore developmentally-appropriate interventions need to continue. Demetriou et al. (2017) advised that interventions should be timed to focus on enhancing specific abilities that are developing in 'time windows'. Their study found children's early reading and mathematics abilities significantly developed during their first year at school in England, when they were aged 4 – 5 years, and thus effective education during this 'time window' could have long-term benefits.

Whilst links between pre-school and later outcomes have been widely reported, as have the impact of a number of early interventions, there are few large-scale longitudinal studies which have investigated the impact of the first year of elementary school on later outcomes, including the potential to reduce the socio-economic gap in attainment. There are examples of large-scale short-term studies which follow children through the first few grades of school (Lonigan et al., 2008). Tymms et al. (2009) looked at children's progress from the beginning of school (age 4 years) up to the end of primary school (age 11 years) in England. They found that the children who were in effective classes during their first school year, that is classes with positive value-added, maintained that 'boost' in attainment up to the end of elementary school. Their advantage over their peers decreased with time but was still statistically significant at age 11. The advantage of an effective first year of school was found to be more significant than any advantages gained by being in effective classes in later years. They also found that membership of more than one effective class had an additive effect but being a member of several effective classes was quite rare; within a school the quality of teaching varies from teacher to teacher. Also of interest are relationships between children's home background, an effective first year of school and later outcomes. Merrell et al. (2014) analysed data from children starting school in England between the years 2000 and 2006; between 34,000 and 67,000 children per cohort. The cohorts were all nationally representative. Background variables were collected, including entitlement to free school meals, which was taken as an indicator of deprivation. Children who were entitled to free school meals, started school with lower early reading and mathematics development than children from more affluent backgrounds. Entitlement to free school meals was still significantly associated with attainment at age 11; there was little evidence to suggest that this gap between the two groups narrowed over time. The study did not investigate the gap

between children from affluent compared with disadvantaged backgrounds for those who are in effective schools, and if differences are found, this has important implications for policy.

Research Questions

Studies which investigate the long-term importance of children's first year of school, known as the Reception year in England, such as the link to grades at the end of compulsory education, may contain important messages for policy. If the first year of school rivals early years provision in terms of long-lasting effects on children's educational outcomes, this should inform policies relating to school starting age and resourcing. This study has analysed data from a large (around 45,000 students) longitudinal dataset from England, which followed children from the start of school to the end of compulsory education at age 16, to investigate the importance of the first year of elementary school, answering the following questions:

1) How well can a baseline assessment at the start of school (age 4) predict later mathematics and English results up to the national leaving exam results at age 16?
2) To what extent do students stay together in the same schools from age 4 to 16? In England, students typically remain together in classes during elementary school but at age 11 they move to secondary school. At this stage they may disperse and attend a number of different secondary schools, altering peer relations and consequently influencing progress.
3) To what extent does membership of an "effective" Reception class (defined in terms of relative progress) during the first year of school impact on later success up to the age of 16?
4) Do some Reception classes reduce the gap in attainment by the end of the first year of school between children from deprived social backgrounds compared with their peers from affluent backgrounds?

5) Does effective education during the Reception year reduce inequality at age 16?

Methodology

Data sources

The data came from schools in England through two sources; The Performance Indicators in Primary Schools (PIPS) monitoring system in which children were assessed at the start of their first year in elementary school and again at the end of their first school year. The PIPS monitoring system is run by the Centre for Evaluation and Monitoring (CEM) (www.cem.org) at Durham University, UK. It provides detailed information to schools about the attainment and progress of their students for self-evaluation purposes. The schools volunteered to participate in the PIPS monitoring system (see Tymms, 1999 for more information), and paid an annual registration fee to do so. The scores of the children in the sample were matched, through the National Pupil Database, to the later English statutory assessments taken at ages 7 (end of Key Stage 1; KS1) and 11 (end of Key Stage 2; KS2), and the GCSE examinations which are taken at age 16. The cohort started school in the 2000/01 academic year and sat their GCSE leaving examination in the 2011/12 academic year.

Ethical approval for the study has been granted from the School of Education Ethics Committee, Durham University, UK.

Sample

The sample included children who attended English elementary and secondary schools.

Table 1 gives details of the background variables at the start of school. The Income Deprivation Affecting Children Index (IDACI) is an English index of deprivation, published by the Department for Communities and Local Government, which measures the proportion of children aged between 0 and 15 years living in deprived families (DCLG, 2015). The term

of entry refers to the time of year when children started school. Most children start school in September at the beginning of the academic year but some start in January or April of the academic year. The English Special Educational Needs Code of Practice includes three different types of educational support for children with special educational needs. School Action refers to a child who is receiving in-school support to address their needs. School Action Plus refers to a child receiving specialist external support to meet their needs. A child with a Statement of Special Educational Needs has persistent and severe special educational needs that require on-going support both in-school and from external specialists. A small percentage (5.7%) of children started school with English as an Additional Language (EAL).

[Table 1 near here]

Measures

Children were assessed within the first few weeks of them starting school and then again at the end of the first school year with the PIPS Baseline and Follow-up Assessment which is a computer-adaptive assessment that is administered by an adult (usually the class teacher or teaching assistant) working with one child at a time. The assessment includes several sections: Name-writing, vocabulary acquisition, concepts about print, letter and word recognition, reading and comprehension, which are combined to give an overall Reading score; Ideas about mathematics, counting, number identification, shape identification, informally presented number problems and formal sums, which are combined together to give an overall Mathematics score. Each of these sections includes questions which the software selects to present to the child through sound files and pictures. The child responds by either saying or pointing to the answer and the adult records whether the response is right or wrong. The questions within each section are ordered in difficulty and when a child makes a certain number of mistakes, the software moves on to the beginning of the next section. The whole assessment takes around 15 minutes. At the end of the Reception year, the assessment

re-starts from around the place in each section where the child started to make mistakes at the beginning of the year. The assessment has been found to have high internal consistency and good predictive validity. For details, see Tymms, (1999) and Tymms et al., (2012).

The end of Key Stage 1 statutory assessments were administered in schools in May 2003 and included assessments of reading and mathematics. Their internal reliability was estimated to be 0.9 (Tymms and Dean, 2004). The end of Key Stage 2 statutory assessments were administered in schools in May 2007 and included tests of English and mathematics. The internal reliability (Cronbach's Alpha) was 0.89 for English and 0.92 for mathematics (Merrell, 2009). The General Certificate of Secondary Education (GSCE) was taken by students at age 16 in 2012, just before the end of compulsory education. Examination boards are responsible for setting the examinations and awarding the certificates.

Representativeness of the sample

The data were broadly representative of England although the sample was slightly higher attaining than the population. This is evident from the mean grades of students in GCSE in English and mathematics which were 0.13 and 0.27 standard deviations above the national averages respectively.

Missing data

Only for ethnicity and special educational needs, of the variables in Table 1, were the proportions of missing data more than 5%. For both variables, the missing data were from cases with significantly higher End of Reception (EOR) scores in reading and mathematics than the mean scores. For these two variables, a dummy was included to indicate "not recorded" for the modelling which is described in the 'Analyses' section below.

It should be noted that the numbers of cases available for the attainment measures in Table 2a decreased as the students aged. This was most pronounced for GCSE English and

mathematics where grades were not available for 11% and 10% of students respectively; Tables 2b and 2c set out details.

[Table 2a near here]

[Table 2b near here]

Those students for whom data were missing at GCSE tended to have lower scores on entry to school in reading and mathematics. The Effect Size differences between the group of children who had GCSE scores compared with those with missing GCSE scores were around 0.2 to 0.3 for reading and mathematics on entry to school respectively. Students with missing GCSE scores also tended to come from more deprived neighbourhoods (ES~0.3) although there was no age difference between the two groups. Of importance are the last two lines of Table 2b which relate to Reception class effectiveness measures that are described later in the paper. The differences in those measures, between those with and without missing GCSE results were very small (ES<.04). Further, just one of the four comparisons reached statistical significance ($p<.05$).

[Table 2c near here]

Table 2c, for the categorical variables, indicates that a slightly higher proportion of males had missing GCSE data. The converse held for females.

The term of entry to school was unrelated to missing data. By contrast ethnicity was; GCSE results were more likely to be missing for White students than from other ethnic minorities. Students whose first language was not English were more likely to be missing at GCSE (around 18%) than not

Seven percent of those students with missing GCSE data had a statement of Special Educational Needs (SEN) compared with one percent of the students with GCSE data.

We do not know exactly why there were no GCSE scores for some of the sample but we suggest some possible reasons. There must be some students who have left the system because their families emigrated, changed their names, or whose records were corrupted. Then there are others who will not have been entered for GCSE or missed the examinations for a variety of reasons We cannot know

which students fit into which category but the data seem to indicate that non-entry is linked to one or more of SEN, low cognitive development at the start of school, low SES classification and EAL. Crucially we did not find an educationally significant link between the characteristics of the students with missing GCSE scores and the effectiveness of Reception classes, which was one of the main areas of focus for this study.

For this paper a dataset based on the same set of pupils on who all data were available was used.

Analyses

The Start of Reception scores, which were collected shortly after the children started school in September, January and April, were age corrected to give the expected score on 1st Sept 2000 and then normalised. All other test score variables and IDACI, were normalised. The distribution of ages at the start of school and the correlations between variables were examined.

Then the flow of students was explored, seeking to establish the extent to which they remained as a group as they moved through the key stages.

For the Reception year, and at the end of each Key Stage during elementary schooling, those students who were in effective classes or schools were identified. To this end, a series of multi-level models (MLMs) were constructed which took as their outcomes the results from assessments of reading/English and mathematics at the end of Reception (EOR), end of Key Stage 1 (KS1) and end of Key Stage 2 (KS2). Each time, the models were based on students nested in the school (or class for Reception), that the students were in at the time of the assessed outcome. The models included a series of background variables as well as all the

prior academic measures. Of particular interest was the link to deprivation, which was allowed to vary from class to class over the Reception year.

An effective class or school was defined as one for which the class or school level residuals were two standard deviations above the mean. The results of the students who were in effective classes or schools were then plotted over time.

Correlations between the effectiveness scores at the three time points were used to investigate the likelihood of students experiencing particularly effective education repeatedly. And, at this point it was possible see if the effective classes/schools tended to have more or less able students in the beginning.

Further MLMs were constructed with the two GCSE results as the outcomes. Membership of an effective Reception class was included, as an additional variable, to see if it would add to the prediction of attainment and if it reduced the slope associated with deprivation (equity).

Results

The distribution of students' ages in September 2000 is shown in Figure 1. It is almost rectangular, indicating that most children start school when they are aged 4. A small proportion had a delayed entry to the second term, starting in January, and a smaller proportion delayed entry for a further term.

[Figure 1 near here]

Correlations

The correlations between age at the start of school, deprivation level and attainment are reported in Table 3.

[Table 3 near here]

The strongest correlations were for reading and mathematics respectively between the start and end of Reception (SOR and EOR); they were 0.72 and 0.71. The link between SOR scores and later attainment become weaker over time but there remained substantial links between reading and mathematics at the start of school and GCSE results at age 16; the correlations were around 0.5.

The relationship with deprivation level was negative; the higher the IDACI score, the higher the deprivation level and the lower the attainment. The correlations between deprivation and attainment remained relatively stable, between -0.2 and -0.3, from the start of elementary school to the end of secondary education.

The correlation between age and attainment was similar in size as the link to deprivation at the start of school but it became weaker as the children matured; the correlation dropped from -0.28 for maths at the start of school to -0.04 for GCSE English.

To what extent do students stay in the same school from age 4 to 16?

Table 4 shows the number of students who remained within the same schools as they moved through elementary and secondary school. In looking at the table, it should be noted that most children in England changed schools at age 11, the end of KS2, when they move from primary (elementary school to much larger secondary schools. School membership remained fairly stable whilst the children were in elementary school up to age 11 (End of KS2) with 66% of children who started the Reception year remaining together in the same school until they transferred to secondary school. Although students became more dispersed in secondary school, still almost 40% of students at age 16 started the same elementary school together.

[Table 4 near here]

Multi-level models to find effective classes/schools

The output from the MLMs with English/reading and mathematics attainment outcomes up to KS2 are shown in Tables 5a and 5b respectively. These models formed the basis for identifying effective Reception classes/schools. In all cases, the independent variables accounted for a substantial part of the variance of the outcome; at least 40% at the pupil level and 20% at the class/school level, with one exception. For mathematics at the end of KS1, 5.5% of the variance of the outcome was accounted for at the school level.

[Tables 5a and 5b near here]

The most important predictors were the prior cognitive measures and deprivation. Age was a negative predictor indicating that the advantage of being older on entry to school gradually faded. Special Educational Needs continued to be associated with less progress at all three time points. Girls steadily gained in reading/English but not so in maths. The ethnic minority groups generally made greater or similar progress to their "white British" peers. The pupils who started school in January or April were behind others with similar characteristics at the End of Reception but made up some of the ground by the end of Key Stage 2. Children with English as an additional language generally made more progress than others during the elementary years.

The MLMs were used to identify effective classes/schools by extracting the school level residual for the six models described above and calculating their means and standard deviations. Classes/schools which were more than two standard deviations above the mean were defined as effective; they were the units with their higher outcomes having controlled for the major predictors.

To what extent does membership of an "effective" class or school impact on later success up to the age of 16?

Figure 2 shows the mean standardised scores for students who were in effective Reception classes/schools. On average, these students started below the mean by 0.1 and 0.06 SD for reading and maths respectively. They then experienced a large boost in their attainment in Reception, which declined by the end of KS1 (age 7) but then remained more or less constant up to GCSE (age 16). The gain from age 4 to 16 amounts to 0.23 and 0.18 SDs for English and maths.

[Figure 2 near here]

Similar charts (Figures 3 and 4) show the results for boosts in attainment associated with membership in an effective school (as defined earlier) at the end of KS1 and KS2. The starting points of these students at the end of KS1 was a little below average but for the end of KS2 they started a point slightly above average. The boosts in attainment associated with effective schooling were not as large as for Reception in the short-term but had similar long-term impact. For the end of KS1 the impacts were 0.16 and 0.21 and for the end of KS2 0.23 and 0.26 SDs for English and maths respectively.

[Figures 3 and 4 near here]

Table 6 shows the correlations between the effectiveness scores from the MLMs; the residuals at the class/school level. They show that at a single time point there are modest correlations of between 0.55 and 0.67 for reading and mathematics. But from one time-point to the next the correlations are very low and always below 0.2 in magnitude. It seems that being in an effective class or school at one time-point is not associated with being in an effective unit a second or third time.

[Table 6 near here]

The trends illustrated in Figure 2 are further explored in MLMs reported in Tables 7a and 7b. These show the prediction of GCSE results from both the Start of Reception scores and the scores derived from the effectiveness of schooling over time.

[Tables 7a and 7b near here]

The results from the MLMs show that age is a negative predictor, as noted earlier. Deprivation is associated with lower outcomes. The next two variables indicate that pupils who started Reception a term later than the majority (in January) had significantly lower GCSE grades by about 0.1 SD for both English and mathematics. Those who started two terms later (in April) had lower GCSE grades by about 0.2 SD. As before, girls made more progress in English but less in maths whilst those for whom English was and additional language (EAL) made more progress than their peers by about 0.2 SD. All the ethnic minority groups made more progress than their white peers. Being identified as have some SEN, at whatever level, was an indicator of lower GCSE grades by about 0.5 SD.

The Start of Reception measures of early reading and maths were significant predictors of the GCSE subjects and membership of an effective Reception class was also predictive of higher GCSE results at age 16, by 0.16 and 0.07 SD units in English and maths respectively.

Reducing the attainment gap

Of particular interest during the Reception year is whether the slight link to deprivation varied from class to class but there was no evidence that it did so. This was established using the MLMs with EOR reading and maths in the tables 5a and 5b. The variable IDACI (deprivation) was allowed to vary at the class level but the error on the variance of the slope was large compared to the variance in both cases. For reading the variance (error) was 0.021 (0.028) and for maths it was 0.040 (0.033); it is concluded that the variance was not significant at the 5% level in both cases.

At GCSE, the larger link to deprivation, in tables 7a and b, did vary from school to school for English and mathematics. For English, the variance (error) was 0.190 (0.035) and for maths it was 0.211 (0.038); it is concluded that the variance was significant at the 5% level in both cases. However, the Reception effectiveness measures for classes did not relate to this slope variation. Its introduction into the model only made a difference to the slope of the deprivation in the third decimal place; by -0.003 and 0.001 for English and mathematics respectively. It is concluded that the effectiveness of schooling during the first year at school is not associated with equity at GCSE.

Summary

This study followed a single cohort of children from the start of school in England, age 4 years, to the end of compulsory education at age 16. At the start of school, the sample comprised almost 48,000 children and by age 16, the sample was smaller; almost 43,000. The sample was broadly representative of the national data although the attainment was slightly higher at age 16.

To summarise the results in relation to the research questions that we set out earlier:

How well can a baseline assessment at the start of school predict later mathematics and English outcomes at age 16?

> The correlation between the baseline assessment at the start of school at age 4 and later attainment declined as children grew older. The correlations in attainment between the start of school and age 16 varied from 0.45 – 0.48 and were all statistically significant ($p \leq 0.01$).

To what extent do children stay together in the same schools from age 4 to 16?

This gives an indication of the mobility of students in the school system and with the longitudinal nature of the analysis in this paper, is useful to consider when interpreting the results. The multi-level models take account of the nesting of students within schools and for some of the analyses there were two or more years between the prior measures and outcomes. Students do gradually disperse as they move through school with 66% of those who started school together remaining in the same schools throughout the elementary phase. Forty four percent of students who started school together were still together in the same schools at age 16.

To what extent does membership of an 'effective' Reception class/school impact on later attainment up to age 16?

Membership of an effective Reception class/school was associated with a boost in attainment that was still apparent at age 16. This amounted to 0.23 and 0.18 SDs for English and maths respectively. Further boosts in attainment were seen for those students who attended effective schools during Key Stage 1 and Key Stage 2. Each of these is additive, however, we did not find evidence that students who experienced an effective year were more or less likely to experience an effective experience in the following educational phase.

Do some Reception classes reduce the gap in attainment by the end of the first year of school between children from deprived social backgrounds compared with their peers from affluent backgrounds?

Deprivation was negatively related to attainment during the Reception year after prior attainment and other variables had been taken into account but this relationship was weak. This link to deprivation did not vary from class to class. This suggests that the link is not a school or class effect but rather a system wide effect. It seems that

deprivation has a small negative impact during the first year at school but that teachers and schools should not be held accountable for that.

Does effective education during the Reception year reduce inequality at age 16?

The impact of home background on educational outcomes after controlling for other variables was substantial and did vary across schools. However, being part of an effective first year class had no measurable impact on this variation.

Discussion and conclusions

This study has extended knowledge and understanding of the link between young children's academic progress during the first year of school and their later outcomes at the end of compulsory schooling at age 16.

Many previous studies have suggested that pre-school interventions, such as the Perry Pre-school Program, are causally linked to improvements in later life outcomes. The results of the present study add to our understanding of the importance of education in the early years, suggesting that the first year of school is also significantly associated with long-term academic outcomes. We found evidence that the boost in attainment from an effective first year of school remained with students right through to the end of secondary school. This finding leads us to suggest that good-quality educational provision in this phase of a child's school career has lasting benefits. Boosts in attainment from effective classes in Key Stages 1 and 2 also had long-term benefits but not as large as those seen in the first year of school. These findings have important implications for policy; the Reception year presents an opportunity to positively impact on children's long-term academic outcomes and we suggest that there should be a focus on the placement of high-quality of teachers and resources to ensure that all children experience an effective first year of school

A further important finding of the present study is the lack of evidence that schools in England reduce the attainment gap between children from affluent backgrounds and their less affluent peers. This gap remains a persistent problem not just in England but in many other countries too. The complex interactions of all of these factors are, to a large extent, beyond the direct influence of schools. In his discussion of theories of educational effectiveness and ineffectiveness, Scheerens (2016) considered the complexity of educational structures and suggested that policy implementation does not follow linear and predictable routes to success; standardised approaches are not necessarily effective in addressing localised issues. Perhaps more widespread acknowledgement of this complexity is a first step towards addressing this persistent problem. In 2011, the UK Government introduced the Pupil Premium to provide additional school funding for those children classed as having deprived backgrounds along with evidence-based advice on how to use the funding effectively (Jarrett et al., 2016).

This study is based upon a large-sample of students over a long period and the findings have important implications for policy. However, by their very nature, longitudinal studies reflect changes over time (for one cohort), but, if they are single-cohort studies, not from cohort to cohort. It is important to continue to monitor changes, analysing the progress of subsequent cohorts if the complexity of education is to be more fully understood and the outcomes of students improved.

References

Belsky, J., Vandell, D. Burchinal, M. Clark-Stewart, K.A., McCartney, K., Owen, M. & The NICHD Early Child Care Research Network. (2007). 'Are there long-term effects of early child-care?'. *Child Development,* 78, 681-701.

Borman, G. & Dowling, M., 2010. Schools and inequality: A multilevel analysis of Coleman's equality of educational opportunity data. *Teachers College Record*, *112*(5), pp.1201-1246.

Bourdieu, P. & Passeron, J. C. (1977) *Reproduction: in education, society and culture* (London, Sage).

Bradley, S. & J. Taylor (2004). *Ethnicity, Educational Attainment and the Transition from School*. The Manchester School 72(3).

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L., (1966). *Equality of educational opportunity*. National Center for Educational Statistics.Washington, DC: U.S. Government Printing Office.

DCLG (2015). *The English Indices of Deprivation 2015 Statistical Release*. London: Department for Communities and Local Government. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/465791/English_Indices_of_Deprivation_2015_-_Statistical_Release.pdf

Demetriou, A., Merrell, C. & Tymms, P. (2017). Mapping and Predicting Literacy and Reasoning Skills from Early to Later Primary School, *Learning and Individual Differences*, 24(2017), 217 – 225.

Duncan, G. J. & Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives,* 27(2), 109–132.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Sexton, H. (2007). School readiness and later achievement. *Developmental psychology*, *43*(6), 1428.,1428-1446.

Farrington, D. (1994) *Delinquency Prevention in the First Few Years of Life*, Plenary Address from 4th European Conference on Law and Psychology, Barcelona, Spain.

Feinstein, L. & Duckworth, K. (2006). *Development in the early years: its importance for school performance and adult outcomes*. Pub. London: Centre for Research on the Wider Benefits of Learning Institute of Education.

Fryer, R. G. and S. D. Levitt (2004). Understanding the Black-White Test Score Gap in the First Two Years of School. *The Review of Economics and Statistics*, 86(2).

Hart, B., & Risley, T. (2003). The early catastrophe: The 30 million word gap by 3. *American Educator, 27*(1), 4-9.

Heckman, J. J., Moon, H.S., Pinto, R., Savelyev, P. and Yavitz, A. (2010). *A new cost-benefit and rate of return analysis for the Perry Preschool Program: A summary*. NBER Working Paper No. 16180, July 2010. JEL No. H43,J13,J24.

Heckman, J.J. (2006). Skill formation and the economics of intervening in disadvantaged children. *Science*, 312 (5782), 1900 – 1902.

Jarrett, T., Long, R. &Foster, D. (2016). School Funding: Pupil Premium. House of Commons Library Briefing Paper, Number 6700, 21st November 2016. http://researchbriefings.files.parliament.uk/documents/SN06700/SN06700.pdf Accessed 11th June 2017.

Lee, V., & Burkham, D. (2002). Inequality at the starting gate: Social background differences in achievement as children begin school. Washington, DC: Economic Policy Institute.

Loeb, S & Bassock, D. (2007). Early childhood and the achievement gap. In H.F. Ladd & E.B. Fiske (Eds.), *Handbook of Research in Education Finance and Policy*, 514 – 534. Routledge Press.

Lonigan, C.J., Schatschneider, C. & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing and spelling. In *Developing Early Literacy: Report of the National Early Literacy Panel* (55 – 106). Washington DC: National Institute for Literacy.

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly*, *28*(2), 314-324.

Merrell, C. (2009) The reliability of the statutory end of Key Stage 2 assessments and their use in England, *Education Review*, 22(1), Autumn 2009, p69 – 74.

Merrell, C., Little, J. & Coe, R. (2014). Is the Attainment Gap among Primary Aged Children Decreasing? In *Harnessing what works in eliminating educational disadvantage: A tale of two classrooms,* Eds. Wood, C. and Scott, R. Pub. Demos: London.

Pagani, L. S., & Fitzpatrick, C. (2013). Children's school readiness implications for eliminating future disparities in health and education. *Health Education & Behavior*, 1090198113478818.

Ramey, C. T. & Ramey, S. L. (1998) Early intervention and early experience, *American Psychologist,*53(2), 109–120.

Raffo, C., Dyson. A., Gunter, H., Hall, D., Jones, L. & Kalambouka, A. (2007) *Education and poverty: a critical review of theory, policy and practice* Joseph Rowntree Foundation. Available online at: www.jrf.org.uk (accessed 18 December 2007.

Sammons, P., Sylva, K., Melhuish, E., Siraj, I., Taggart, B., Toth, K. & Smees, R. (2014). *Influences on students' GCSE attainment and progress at age 16. Effective pre-school, primary and secondary education project (EPPSE) research report*. Pub. London: Department for Education.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W.S. Beleld, C.R. & Nores, M. (2005). *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40.* Ypsilanti, MI:High/Scope Press.

Scheerens, J. (2016). Meta-Analyses of School and Instructional Effectiveness. In *Educational Effectiveness and Ineffectiveness* (pp. 175-223). Springer Netherlands.

Scheerens, J. (2016). Theories on educational effectiveness and ineffectiveness. In *Educational Effectiveness and Ineffectiveness* (pp. 259 - 289). Springer Netherlands.

Scheerens, J. (2017). The Routledge international handbook of educational effectiveness and improvement. Research, policy and practice.

Shonkoff, J. P., & Phillips, D.A. (2000). Neurological basis of developmental vulnerability. In Shonkoff, J.P. & S.J., Meisels (Eds.), *Handbook of early childhood intervention* (2nd Ed.)

Tymms, P., Merrell, C. and Buckley, H. (2016) *Children's development at the start of school in Scotland and the progress made during their first school year: An analysis of PIPS*

baseline and follow-up assessment data. Research report for the Scottish Government. ISBN: 9781785448942. http://www.gov.scot/Publications/2015/12/5532/0 Accessed January 2016.

Tymms, P., Merrell, C., Henderson, B., Albone, S., & Jones, P. (2012). Learning Difficulties in the Primary School Years: Predictability from On-Entry Baseline Assessment. *Online Educational Research Journal*. June 2012 , Retrieved from www.oerj.org 13/6/12

Tymms, P., Jones, P., Albone, S., & Henderson, B. (2009). The first seven years at school. *Educational Assessment and Evaluation Accountability, 21*, 67-80.

Tymms, P. & Dean, C. (2004). *Value-Added in the Primary School League Tables: A Report for the National Association of Head Teachers*. Pub. Durham, UK: CEM Centre. www.cem.org

Tymms, P. (1999). *Baseline Assessment and Monitoring in Primary Schools: Achievements, Attitudes and Value-added Indicators*. London, UK: David Fulton Publishers.

Figure 1: Distribution of ages in September 2000

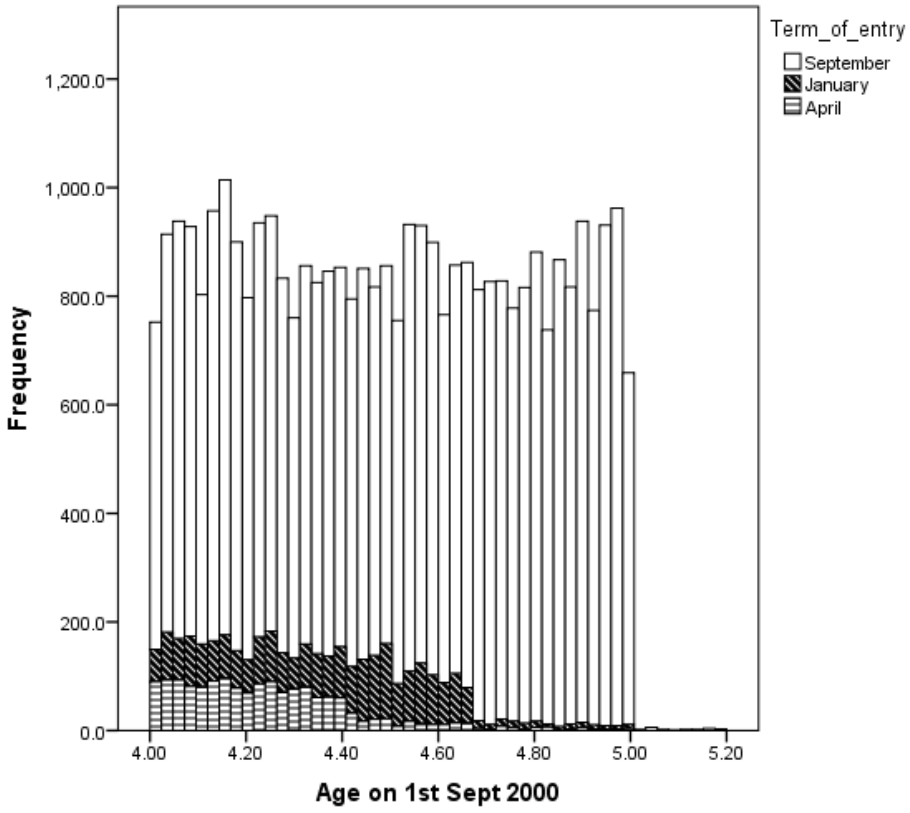

Figure 2 Mean standardised scores for students in effective Reception classes

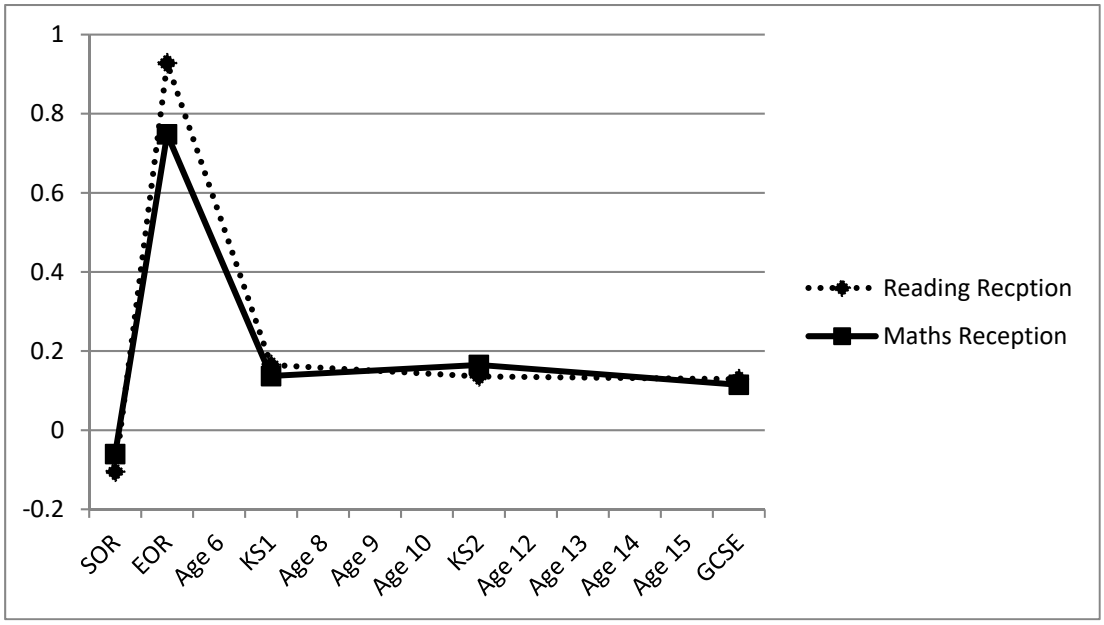

Figure 3 Mean standardised scores for students in effective KS1 classes

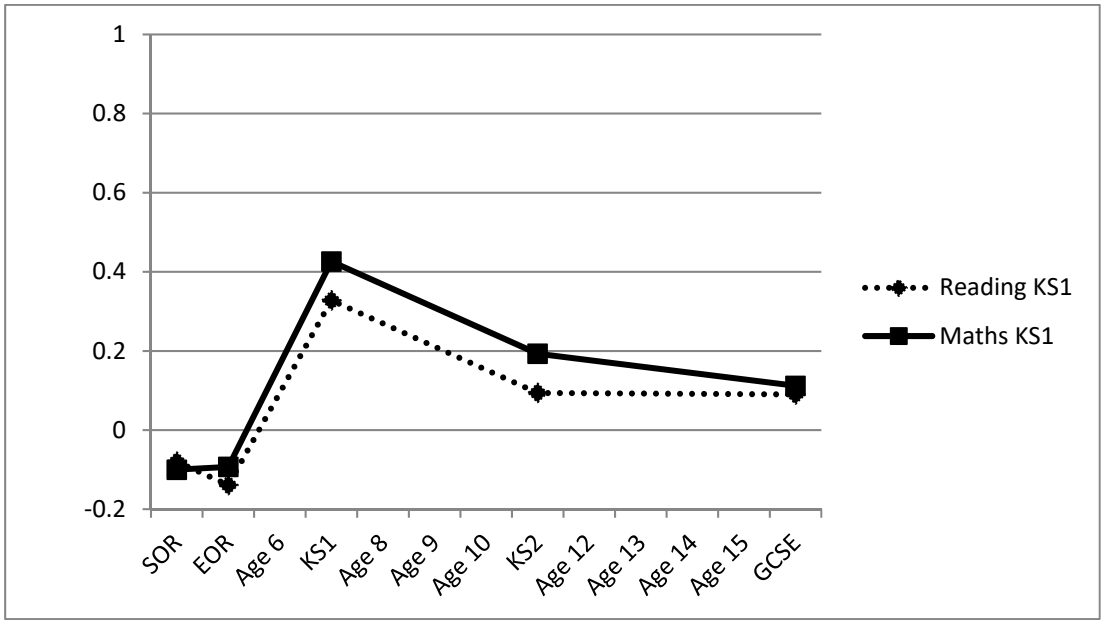

Figure 4 Mean standardised scores for students in effective KS2 classes

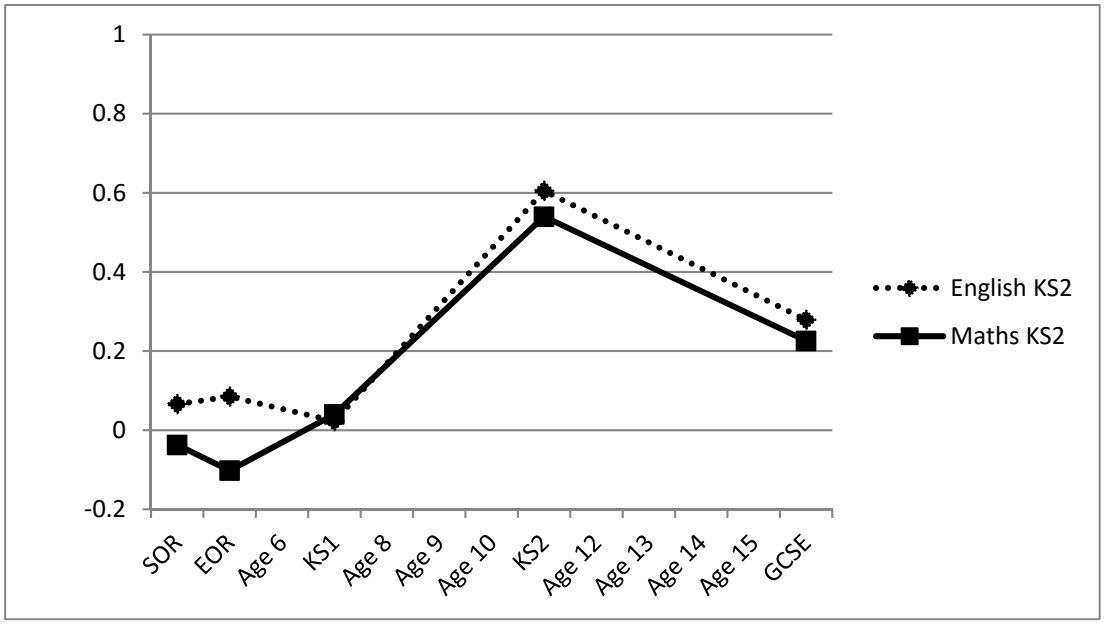

Table 1 Student characteristics

Age		Mean = 4.50 SD=0.29
IDACI	**Deprivation**	**Mean 0.209 SD=0.16**
Sex		Male=50% Female=50%
Term of entry	**Sept**	**90.2%**
	Jan	**6.8%**
	April	**3.1%**
Ethnicity	**White**	**88.9%**
	Black	**0.8%**
	Asian	**6.5%**
	Mixed	**2.3%**
	AOEG (any other)	**0.3%**
	Chinese	**0.2%**
EAL	**English**	**94.3%**
	Non-English	**5.7%**
Special Educational Needs	**School Action**	**11.5%**
	School Action Plus	**4.6%**
	Statement	**1.0%**

Table 2a Number of students scores on the attainment measures

	n	Min	Max	Mean	SD
GCSE English	42923	3.00	12.00	7.05	1.40
GCSE mathematics	43146	3.00	12.00	6.96	1.67
KS2 English	45614	2.5	5.9	4.49	0.72
KS2 mathematics	45661	2.5	5.9	4.52	0.80
KS1 Reading	44771	1	4	2.18	0.61
KS1 mathematics	45194	1	4	2.19	0.57
End of Reception reading	47920	.00	178.00	88.27	37.48
End of Reception mathematics	47920	.00	69.00	39.55	10.14
Start of Reception reading	47826	.00	99.00	29.43	14.58
Start of Reception mathematics	47796	.00	63.00	19.92	8.14

Table 2b: Missing data at GCSE for continuous variables

	GCSE English				GCSE maths			
	Mean sample	Mean Missing	P (t-test)	Effect Size	Mean sample	Mean Missing	P (t-test)	Effect Size
BLA maths	0.02	-0.23	<0.001	-0.18	0.02	-0.32	<0.001	-0.34
BLA reading	0.01	-0.16	<0.001	-0.25	0.02	-0.24	<0.001	-0.26
IDACI	0.21	0.27	<0.001	0.28	0.22	0.27	<0.001	0.27
Age	4.50	4.50	0.75	0.01	4.50	4.50	0.34	0.01
Reception class maths residual[#]	-0.023	-0.076	0.05	-0.02	-0.018	-0.136	0.07	-0.03
Reception class reading residual[#]	0.016	0.047	0.03	0.01	0.02	-0.03	0.38	-0.02

[#] The residuals are derived from the multi-level models set out later. They were designed to measure "effective" reception classes.

Table 2c: Missing data at GCSE for categoric data

		GCSE English			GCSE maths		
		Sample	Missing	P (Chisq)	Sample	Missing	P (Chisq)
Sex	Male	50.5%	55.1%	<0.001	50.8%	52.7%	0.01
	Female	49.5%	44.9%		49.2%	47.3%	
Term of entry	Sept	88.3%	88.7%	0.70	88.3%	88.6%	0.83
	Jan	7.4%	7.0%		7.4%	7.2%	
	April	4.3%	4.3%		4.3%	4.2%	
Ethnicity	White	88.5%	91.2%	<0.001	88.6%	89.1%	<0.001
	Black	0.9%	1.1%		0.8%	1.3%	
	Asian	6.9%	3.9%		6.8%	5.9%	
	Mixed	2.3%	2.7%		2.3%	2.7%	
	AOEG (any other)	0.4%	0.3%		0.4%	0.3%	
	Chinese	0.2%	0.1%		0.2%	0.0%	
EAL	English	93.8%	82.1%	<0.001	93.7%	82.3%	
	Non-English	6.2%	17.9%		6.3%	17.7%	
SEN	No data	10.7%	50.7%	<0.001	10.7%	52.2%	<0.001
	None	70.9%	25.2%		70.9%	25.8%	
	School Action	11.6%	8.8%		11.6%	8.6%	
	School Action Plus	5.4%	8.3%		5.3%	8.8%	
	Statement	1.4%	7.0%		1.3%	7.6%	

Table 3 Correlations between key variables and attainment

	Age	Deprivation level	GCSE Eng	GCSE Math	KS2 Eng	KS2 Math	KS1 Read	KS1 Math	EOR Read	EOR Math	SOR Read
GCSE Eng	0.04**	-0.27**									
GCSE Math	0.03**	-0.28**	0.73**								
KS2 Eng	0.10**	-0.27**	0.75**	0.66**							
KS2 Math	0.09**	-0.24**	0.62**	0.79**	0.73**						
KS1 Read	0.14**	-0.22**	0.56**	0.52**	0.67**	0.57**					
KS1 Math	0.17**	-0.20**	0.49**	0.56**	0.58**	0.65**	0.63**				
EOR Read	0.24**	-0.28**	0.53**	0.50**	0.64**	0.57**	0.64**	0.55**			
EOR Math	0.25**	-0.22**	0.48**	0.51**	0.58**	0.61**	0.56**	0.57**	0.77**		
SOR Read	0.24**	-0.30**	0.48**	0.45**	0.57**	0.51**	0.54**	0.48**	0.72**	0.59**	
SOR Math	0.27**	-0.23**	0.47**	0.48**	0.56**	0.57**	0.53**	0.53**	0.66**	0.71**	0.74**

** P<.01

Table 4 Common school membership

	Reception	End of KS1	End of KS2	GCSE
Different schools	0 0%	5602 11.7%	16664 35.4%	28749 60.2%
Schooled together	47038 100%	42320 88.3%	30397[1] 66.6%	1927[2] 39.8%
Number of schools	1884	3660	4859	612
Numbers of classes	2860			

[1] At least 12 students in same school
[2] At least 12 students in same school

Table 5a Multi-level Models with cognitive controls for Reading/English

		EOR Reading	***KS1 Reading***	***KS2 English***
Fixed				
Age		*0.028 (0.010)*	*-0.076 (0.010)*	*-0.183 (0.011)*
IDACI		*-0.122 (0.022)*	*-0.148 (0.022)*	*-0.31 (0.024)*
Entry	*Jan*	*-0.387 (0.016)*	*0.069 (0.013)*	*0.033 (0.014)*
	April	*-0.775 (0.023)*	*0.134 (0.020)*	*0.08 (0.021)*
Sex	*Female*	*0.049 (0.005)*	*0.06 (0.006)*	*0.108 (0.006)*
EAL		*0.06 (0.018)*	*0.057 (0.019)*	*0.058 (0.020)*
Ethnicity	*Black*	*0.087 (0.031)*	*-0.057 (0.033)*	*0.022 (0.035)*
	Asian	*0.086 (0.018)*	*0.067 (0.019)*	*0.162 (0.020)*
	Mixed	*0.036 (0.018)*	*0.034 (0.019)*	*0.061 (0.020)*
	Other	*0.001 (0.047)*	*0.118 (0.050)*	*0.208 (0.053)*
	Chinese	*0.131 (0.055)*	*0.158 (0.058)*	*0.312 (0.062)*
	Not recorded	*0.002 (0.027)*	*0.039 (0.029)*	*-0.034 (0.012)*
Special Need	*School Action*	*-0.274 (0.009)*	*-0.315 (0.010)*	*-0.349 (0.010)*
	School Action Plus	*-0.34 (0.013)*	*-0.472 (0.014)*	*-0.428 (0.015)*
	Statement	*-0.239 (0.027)*	*-0.426 (0.029)*	*-0.333 (0.031)*
SOR	*Maths*	*0.256 (0.004)*	*0.08 (0.005)*	*0.038 (0.005)*
	Reading	*0.487 (0.005)*	*0.06 (0.005)*	*0.103 (0.006)*
EOR	*Maths*		*0.105 (0.005)*	*0.081 (0.006)*
	Reading		*0.331 (0.006)*	*0.125 (0.006)*
KS1	*Maths*			*0.163 (0.005)*
	Reading			*0.369 (0.006)*
KS2	*Maths*			
	English			
Variance	*School level*	*0.139 (0.004)*	*0.042 (0.002)*	*0.059 (0.003)*
	Student level	*0.247 (0.002)*	*0.290 (0.002)*	*0.323 (0.002)*
Variance % at school level		*36.0*	*12.7*	*15.4*
Variance reduction from null (%)]	*School level*	*35.6*	*22.2*	*48.7*
	Student level	*63.1*	*48.0*	*56.2*

Table 5b Multi-Level Models with cognitive controls for mathematics

		EOR Maths	**KS1 Maths**	**KS2 Maths**
Fixed				
Age		0.138 (0.011)	-was 0.003 (0.011)	-0.281 (0.012)
IDACI		-0.060 (0.024)	-0.061 (0.024)	-0.232 (0.025)
Entry	*Jan*	-0.256 (0.017)	0.072 (0.014)	0.017 (0.015)
	April	-0.584 (0.024)	0.121 (0.021)	0.064 (0.022)
Sex	*Female*	-0.119 (0.006)	-0.146 (0.006)	-0.211 (0.006)
EAL		0.052 (0.019)	0.036 (0.019)	0.028 (0.021)
Ethnicity	*Black*	-0.011 (0.033)	-0.118 (0.034)	-0.045 (0.036)
	Asian	0.015 (0.020)	0.073 (0.020)	0.154 (0.021)
	Mixed	-0.011 (0.019)	0.005 (0.020)	0.039 (0.021)
	Other	-0.026 (0.050)	0.044 (0.052)	0.201 (0.055)
	Chinese	0.089 (0.058)	0.27 (0.060)	0.488 (0.065)
	Not recorded	-0.078 (0.011)	0.041 (0.030)	0.046 (0.012)
Special Need	*School Action*	-0.295 (0.009)	-0.241 (0.010)	-0.26 (0.011)
	School Action Plus	-0.386 (0.014)	-0.342 (0.014)	-0.29 (0.016)
	Statement	-0.431 (0.029)	-0.401 (0.030)	-0.233 (0.032)
SOR	*Maths*	0.490 (0.005)	0.118 (0.005)	0.106 (0.006)
	Reading	0.192 (0.005)	0.04 (0.006)	0.048 (0.006)
EOR	*Maths*		0.226 (0.006)	0.228 (0.006)
	Reading		0.139 (0.006)	-0.008 (0.007)
KS1	*Maths*			0.41 (0.006)
	Reading			0.141 (0.006)
KS2	*Maths*			
	English			
Variance	*School level*	0.152 (0.005)	0.052 (0.002)	0.062 (0.003)
	Student level	0.282 (0.002)	0.310 (0.002)	0.349 (0.003)
Variance % at school level		35.0	14.4	15.1
Variance reduction from null (%)]	*School level*	30.6	5.5	36.7
	Student level	58.0	40.2	54.6

Table 6 Correlations between school level MLM residuals from Tables 5a and 5b (effectiveness scores)

	Recp read	Recp maths	KS1 read	KS1 maths	KS2 reading	KS2 maths
Reception Reading	1.00					
Reception Maths	0.67	1.00				
KS1 Reading	-0.12	-0.14	1.00			
KS1 Maths	-0.09	-0.15	0.61	1.00		
KS2 Reading	-0.09	-0.12	-0.02	0.00	1.00	
KS2 Maths	-0.05	-0.19	0.02	0.02	0.55	1.00

P<.01 for all figures

Table 7a Multi-Level Model for GCSE English

		GCSE English
Fixed		
Age		*-0.289 (0.014)*
IDACI		*-0.766 (0.026)*
Entry	*Jan*	*-0.110 (0.015)*
	April	*-0.185 (0.023)*
Sex	*Female*	*0.223 (0.008)*
EAL		*0.193 (0.024)*
Ethnicity	*Black*	*0.137 (0.042)*
	Asian	*0.362 (0.023)*
	Mixed	*0.137 (0.025)*
	Other	*0.447 (0.066)*
	Chinese	*0.510 (0.077)*
	Not recorded	*-0.042 (0.039)*
Special Need	*School Action*	*-0.515 (0.012)*
	School Action Plus	*-0.643 (0.018)*
	Statement	*-0.558 (0.038)*
SOR	*Maths*	*0.184 (0.006)*
	Reading	*0.257 (0.006)*
Reception class MLM residual		*0.166 (0.011)*
Variance	*School level*	*0.032 (0.003)*
	Student level	*0.522 (0.004)*
Variance % at school level		*5.8*
Variance reduction from null (%)]	*School level*	*59.5*
	Student level	*32.5*

Table 7b Multi-Level Model for GCSE Maths

		GCSE Maths
Fixed		
Age		-0.387 (0.014)
IDACI		-0.782 (0.027)
Entry	*Jan*	-0.123 (0.016)
	April	-0.222 (0.023)
Sex	*Female*	-0.147 (0.008)
EAL		0.197 (0.025)
Ethnicity	*Black*	0.099 (0.043)
	Asian	0.413 (0.023)
	Mixed	0.049 (0.025)
	Other	0.438 (0.066)
	Chinese	0.867 (0.078)
	Not recorded	0.002 (0.040)
Special Need	*School Action*	-0.519 (0.012)
	School Action Plus	-0.623 (0.018)
	Statement	-0.529 (0.038)
SOR	*Maths*	*0.246 (0.006)*
	Reading	*0.213 (0.006)*
Reception class MLM residual		*0.070 (0.011)*
Variance	*School level*	0.029 (0.003)
	Student level	0.532 (0.004)
Variance % at school level		9.6
Variance reduction from null (%)]	*School level*	67.0
	Student level	**32.7**