
**EFFECTIVE PRE-SCHOOL AND PRIMARY EDUCATION 3-11 PROJECT (EPPE 3-11)
THE EFFECTIVENESS OF PRIMARY SCHOOLS IN ENGLAND IN KEY STAGE
2 FOR 2002, 2003 AND 2004**

*Edward Melhuish[#], Helena Romaniuk[#], Pam Sammons[§], Kathy Sylva⁺,
Iram Siraj-Blatchford^{*}, Brenda Taggart^{*}*

*[#]Birkbeck, University of London, [§]University of Nottingham, ⁺University of Oxford,
^{*}Institute of Education, University of London*

Introduction

This research brief reports the results of value added multilevel models to investigate pupil progress during Key Stage 2, controlling for prior attainment and other background factors, for all schools in England over a three year period (2002-4). These models build upon existing work on school effectiveness undertaken by DfES/Ofsted and others by incorporating further area-level variables, examining gender by ethnicity interactions and exploring differential effectiveness of primary schools for pupils with different levels of ability. The work is part of the wider Effective Pre-school and Primary Education 3-11 (EPPE 3-11) project which is studying the development and attainment of 2500 plus young children from age 3 to the end of Key Stage 2.

Key findings

- For all subjects, the prior attainment of pupils measured by Key Stage 1 (KS1) assessments is an important contributor to their performance in Key Stage 2 (KS2) assessments. KS1 performance in reading is most important for predicting KS2 English performance and KS1 Mathematics is the most powerful predictor for KS2 Mathematics, Science and average score.
- Measures of school effects on English are most variable and instable between schools and across years, while measures involving maths are most stable.
- There are marked differences in the amount of progress that different schools produce between KS1 and KS2, depending on the initial level of ability of pupils - this is termed differential effectiveness. Analysis suggests that a major differentiating feature between effective and ineffective schools (in terms of contextualised pupil progress) resides in their degree of success with low ability pupils in particular. Also while all children benefit from being in an effective school rather than an ineffective one, the consequences are markedly greater for low ability children than for high ability children.
- Pupils who are eligible for free school meals (FSM) and pupils with special educational needs (SEN) show substantially less progress across all subjects between KS1 and KS2 in all three years. As these pupils also have lower KS1 attainment, the gap is widening between them and others over time.
- In English, girls from all ethnic groups made increasingly better progress than boys from KS1 to KS2 in all subjects in all three years. Bangladeshi and Chinese boys and girls made more progress than white boys and girls.
- In maths, boys made consistently more progress than girls for all three years and in all ethnic groups. Chinese children did better than white children in all three years. Caribbean boys did worse than white boys, while progress of Caribbean girls is comparable to white girls.

Background

The White Paper "Excellence in Schools" (DfEE, 1997) highlighted the need for better information about pupils to be available to support the drive to raise standards. Specifically pupil level information was needed to track individual pupil's progress and that this information then needed to be linked to data on pupil attributes, e.g. ethnicity, special needs, free school meal (FSM) eligibility etc., to contextualise the pattern of educational performance. The National Pupil Database implemented in 1999 included for every pupil in state schools all key stage results from the summer of 2000. The Pupil Level Annual Schools Census (PLASC) provides details of ethnicity, first language, special education needs, FSM eligibility, post code, etc. Linking these databases provides a basis for the analysis of pupil progress as related to some basic demographic attributes of pupils. Where this analysis is undertaken using multilevel modelling, then the school level measures can be derived as indicators of the impact of attending a specific school having allowed for the characteristics of pupils attending that school (i.e. school effectiveness).

Aims and objectives

The aim of this work was to compare the effectiveness across Key Stage 2 of all primary schools in England for 3 successive years. The effectiveness measures were derived from the 2002, 2003 and 2004 Key Stage 2 results for English, Mathematics and Science. Factors known to influence the Key Stage 2 results, pupils' prior achievement at Key Stage 1 and certain individual pupil characteristics, were included in the analysis so that the measures of effectiveness reflect the schools' effectiveness rather than the composition of the school.

There is existing work on school effectiveness undertaken by DfES/Ofsted (<http://www.standards.dfes.gov.uk/performance/1316367/CVAinPAT2005/?versi>) and others. The models explored in this report build upon existing work through further incorporation of area-level variables, the examination of gender by ethnicity interactions and the exploration of differential effectiveness for pupils for different levels of ability in primary schools.

This effectiveness analysis is part of the wider EPPE 3-11 longitudinal study. The EPPE 3-11 project (2003-2008) builds on the work of the earlier Effective Provision of Pre-school Education (EPPE) project (1996-2003) which was the first major longitudinal study in Europe to investigate the impact of pre-school provision on a national sample of young children (Sylva, Melhuish, Sammons, Siraj-Blatchford & Taggart, 2004). The measures of primary school effectiveness derived in this report will be used in further analyses within the EPPE (3-11) project. These analyses will evaluate the contribution of primary school effectiveness in conjunction with child, family and pre-school characteristics to children's cognitive and social development within the EPPE (3-11) longitudinal study.

This report cannot distinguish what characteristics produce differences in effectiveness between schools, as it was not designed for this purpose. However other reports from the EPPE 3-11 project (Sammons et al, 2006a & b) consider this topic in terms of classroom processes associated with school effectiveness.

Method

This analysis covers children's progress during Key Stage 2 in all primary schools in England over a 3-year period. Value added multilevel models are used to investigate children's progress during Key Stage 2 by controlling for a child's prior attainment, as well as for a number of background influences. These analyses allow measurement of the extent to which children's progress can be attributed to the primary school attended. Primary schools where children make significantly greater progress than predicted (on the basis of prior attainment and intake characteristics) can be viewed as *more effective*, and schools where children make less progress than predicted can be viewed as *less effective*. The phrase 'effectiveness' throughout this brief therefore refers solely to this measure of progress, not to any other characteristics or qualities of schools.

The analyses focus on **progress**, rather than absolute attainment, in the three subject areas of English, Mathematics and Science, and in average key stage scores. The value added models controlled for pupil background characteristics such as gender, ethnic group, English as an additional language, free school meal eligibility and

special educational needs. Further development of the value added models measured the differential effects for boys and girls in different ethnic groups, as well as considering area effects. The child's postcode was used to relate the child's residence to the Index of Multiple Deprivation (IMD) (ODPM, 2004) and to variables derived from the 2001 Census. Further variables reflecting the composition of schools were also used in the analyses. From these analyses, it is possible to identify trends in effectiveness in terms of academic outcomes over the three successive years.

The analyses are designed to answer the question: *What affects pupils' progress over Key Stage 2 in primary school?* In analysing progress, the value added models include measures of a child's ability at the start of Key Stage 2, i.e. measures of their Key Stage 1 attainment as well as predictor variables that might explain progress. The consequences of this strategy are as follows:

- The inclusion of Key Stage 1 attainment in the value added models will absorb the effects of several child, parent, family, home and area factors, if their effects do not persist additively over the Key Stage 2 period. Hence the relative importance of these factors in measuring progress may appear substantially less than would be the case if Key Stage 1 scores are excluded in the models, i.e. attainment only is considered.
- Where children are not showing high levels of attainment in Key Stage 1 assessments, there is more scope for progress for such children. Hence such children may show bigger progress effects, without necessarily showing high attainment at the end of Key Stage 2.

Findings

Influence of child characteristics on Key Stage 2 attainment and progress

For all subjects in all years the **prior attainment** of the pupils, as measured by Key Stage 1 assessments, is an important contributor to their performance in Key Stage 2 assessments. Key Stage 1 performance in Reading is most important for predicting Key Stage 2 English performance (effect sizes = 1.32, 1.34 and 1.28 in 2002, 2003 and 2004 respectively), but for Mathematics, Science and the average score, Key Stage 1

performance in Mathematics is the most powerful predictor of any prior attainment or other measures (Key Stage 1 Mathematics effect sizes in 2004 = 1.74, 1.00 and 1.45 for Key Stage 2 Mathematics, Science and Average, respectively). Relative to Key Stage 1 Reading, Writing and Mathematics, the effect of Key Stage 1 Science was inconsistent across the years and tended to have lower effect sizes, even for Science itself. The Key Stage 1 Science assessment is entirely a teacher rating and possibly the unstandardised nature of the assessment contributes to its lack of consistency and predictive power. Alternatively, the Science undertaken in Key Stage 1 may be too little or fragmented to produce a more useful assessment at the end of Key Stage 1.

The powerful effects of prior attainment in predicting Key Stage 2 attainment will have consequences for the effects to be attributed to other variables such as pupil characteristics. In this report the effects for other variables can be regarded as effects on progress across the Key Stage 2 period as Key Stage 1 attainment is included in the models. As Key Stage 1 attainment will absorb much of the effects of other variables upon school attainment, the effect of other variables is likely to be substantially less than if the models focused on the contribution of other variables in predicting *attainment* at Key Stage 2 rather than *progress* across Key Stage 2.

With regard to other pupil characteristics, pupils who are **younger** in their school year consistently, across subjects and years, show slightly better progress, although the effect sizes indicate that there is only a small effect in Mathematics and a very small effect in all other subjects (2004 effect sizes = -0.07, -0.16 and -0.12 for English, Mathematics and Science respectively). It would appear that the younger pupils are slowly narrowing the gap with their older classmates.

Pupils for whom **English is an additional language (EAL)** show better progress than native speakers of English consistently for English and Mathematics, and only in 2004 for Science, although the effect sizes are small (2004 effect sizes = 0.10, 0.18 and 0.03 for English, Mathematics and Science respectively). As these pupils may well be starting from a lower base, and are not reaching higher attainment at Key Stage 2, this finding reflects a narrowing of the gap between EAL pupils and native speakers. This interpretation is congruent with

results produced by DfES (2005, 2006).

Whether pupils are eligible for **free school meals (FSM)** can be regarded as a marker for family poverty. This marker for poverty consistently predicts poorer progress in Key Stage 2 for all subjects across years. These effects are not large (range -0.10 to -0.17 in effect size) being slightly less for Mathematics than for the other subjects. The pupils eligible for free school meals are attaining lower Key Stage 2 attainment so the gap is widening over time. These results are congruent with DfES results (2005, 2006), and also consistent with Key Stage 2 attainment.

Where pupils have a **special educational need (SEN and SEN other)** they show substantially less progress across all subjects in all years, and this result is also reported by DfES (2005, 2006) for 2004 and 2005. The effect is greater for English and the average score but is very substantial for all subjects (SEN 2004 effect sizes = -0.61, -0.40, -0.27 and -0.70; SEN other effect sizes = -0.72, -0.61, -0.40 and -0.74 for English, Mathematics, Science and Average scores). Thus the gap between SEN pupils and non-SEN pupils is widening over time.

The progress girls made from Key Stage 1 to 2 varied significantly between schools for all subjects and for all years. There are consistent **gender** effects in Key Stage 2 attainment, whereby girls in all ethnic groups attain better in English, and boys do better than girls in all ethnic groups in Mathematics. In Science there is no clear pattern of gender difference in attainment.

In English, small effect sizes for the Bangladeshi and Chinese boys and girls indicate that they are progressing more than White British or Irish (WBI) boys and girls, respectively, in all years, and White Other children show small effects in 2003 and 2004. Children in the Caribbean, Black African and Black Other, Indian, Pakistani, Mixed, any other ethnic origin and ethnic origin unknown groups were, in general, comparable to the WBI children of the same gender.

In Mathematics, Chinese children do better than WBI children in all three years, children in the any other ethnic origin group do better than in 2003 and 2004 and Bangladeshi children only do better in 2003. Caribbean boys do worse than WBI boys, while Caribbean girls are comparable to WBI girls. The children in the other ethnic groups were found

to be comparable to the WBI children of the same gender.

In Science, WBI, Pakistani, Bangladeshi and girls with ethnic origin unknown had effect sizes showing that their progression was below that of the boys in the same ethnic group for all three years. In 2002, White Other, Mixed and Indian girls all did worse. The effect sizes for the other ethnic groups and years when comparing girls' progress to boys was negligible. Chinese pupils also tended to do better in science than WBI pupils, as did the Pakistani pupils in all three years. White Other girls did better than WBI girls in the last two years. Caribbean boys did worse than WBI boys in all three years.

There are consistent small effects associated with the **area** in which a pupil resides, which reflect the effects of level of deprivation. Primary schools typically have distinct catchment areas, and hence school composition effects can also be interpreted as reflecting the effect of area deprivation.

For Key Stage 2 English, prior achievement has the largest influence on outcome, with Reading and Writing having the strongest influence (effect sizes=1.28 and 0.80, respectively, in 2004). However, the effect of SEN is comparable to the effect of Writing. Aside from the joint effects of gender and ethnicity, the effects for all other child, school and area variables are either small or negligible. For Key Stage 2 Mathematics, Key Stage 1 Mathematics has the largest effect on outcome, followed by SEN and some of the gender and ethnicity combinations, with all other variables having only small or negligible effects. For Key Stage 2 Science, the Key Stage 1 subjects have the strongest relationship with Mathematics having the largest individual effect. The effect of SEN is only moderate for Science, the gender and ethnic effects are small, with all other variables only having small or negligible effects.

Stability over time

The analysis of school level measures of effectiveness across subjects and across years indicates some consistency and stability, but also that there is considerable variation and change amongst schools in their degree of effectiveness across subjects and across years. In particular, measures involving English seem open to most variation and instability, and measures of school effects upon Mathematics are most stable.

Differential effectiveness of primary schools

The analysis reveals that there are marked differences in the amount of progress that schools produce dependent upon the level of initial ability of pupils. School differences in effectiveness that are dependent upon the initial level of ability of a pupil can be termed differential effectiveness. The level of differential effectiveness is markedly different for different primary schools.

Summary and Discussion

The results with regard to the effects of prior attainment, age, FSM eligibility, EAL, SEN and gender are largely compatible with other reports such as those from DfES (2005, 2006).

The gender effects are consistent across years and in the moderate to large effect size range, being most powerful in Mathematics, indicating that they are important in understanding pupils' educational performance. Similar gender differences in progress have been reported in other research. An example is Strand (1999) who considered pupil progress for the baseline to Key Stage 1 period and found that girls showed more progress in Reading and Writing and boys more progress in Mathematics. Also, the effects reported here are consistent with the gender differences in Key Stage 2 attainment reported by DfES (2005, 2005b) for 2002 through to 2005 where girls consistently do better overall in English and related subjects, boys do better in Mathematics and the genders are equivalent for Science. However, when considering progress from Key Stage 1 to Key Stage 2 rather than attainment, the effects of gender do vary between ethnic groups, and also sometimes by subject, as indicated by the ethnic group by gender interactions, controlling for other background factors.

Overall the results for progress of ethnic groups are compatible with data on attainment. DfES (2005) summarise the educational achievement of ethnic groups in England for 2003/4 for Key Stages 1, 2, 3, and 4 (ages 7, 11, 14 and 16). Pupils of Chinese and Indian origin show high attainment relative to the average. However, pupils of Black Caribbean, Pakistani and Bangladeshi origin do worse than the average. In terms of progress across the Key Stages, the same report indicates that progress for Bangladeshi and Black African

pupils is greater than the average across Key Stage 2 and across Key Stage 4. Pakistani pupils also show greater improvement across Key Stage 4. Pupils from Chinese, Indian and "Any other ethnic origin" backgrounds show greater improvement across each Key Stage. However, pupils from White, Black Caribbean and Black Other ethnic backgrounds show lower progress than the national average.

The results associated with the interactions between gender and ethnic background lead to qualifications being placed upon the differences in progress associated with ethnic group. Note that in all ethnic groups Key Stage 2 attainment in English is better for girls than boys, whereas attainment in Mathematics and Science is very similar. This report has analysed progress rather than attainment. For some ethnic groups there are significant differences between the performance of girls and boys. DfES (2006) refers to girls consistently outperforming boys in all of the minority ethnic groups over Key Stages 1 to 4. Currently this is certainly true for GCSE overall attainment. The results reported here, which are for progress rather than attainment, partially confirm such a view but indicate that the nature of ethnic by gender interactions require a rather more nuanced approach. Also possibly the effects associated with ethnic groups may be changing with different cohorts of children working their way through school.

Most research on ethnic differences in educational attainment has focused on secondary schooling, and suggests possible reasons for the observed effects. Wilson, Burgess and Briggs (2005) find evidence that all ethnic minorities are making greater progress in secondary schools than White students. The explanations of the differences associated with ethnic background are various. Bradley & Taylor (2004) find that non-school factors may be important, e.g. the performance of Non-White pupils is more adversely affected by living in a single-parent household. Modood (2003) has argued that gender norms and cultural expectations play an important role and that many South Asians have high educational aspirations that are not constrained by social class in the way that they are in traditional White British culture. Yet other factors discussed by Cook and Ludwig (1998) and Modood (2003) refer to the fear of "acting white" that may discourage academically able black pupils from putting much effort into school work. Further exploration of the interactions between ethnicity

and area characteristics using national data on school achievement may be one way to investigate some alternative explanations for ethnic group differences.

There are consistent small effects associated with the level of deprivation of the area in which a pupil resides. Primary schools typically have distinct catchment areas hence the school composition effects may also reflect the effect of area deprivation. When considering these area/community level measures of deprivation it is not clear whether this is deprivation at the individual family level or deprivation at the community level that is influencing the individual pupil. While the data analytic models include individual pupil characteristics there is much variation in family circumstance that is not captured. Hence, the area measures may reflect aspects of the individual pupil's family as well as aspects of the area of residence because particular types of family are more likely to live in particular types of area.

School differences in effectiveness dependent upon the initial level of ability of a pupil (differential effectiveness) has been described and discussed in the literature on secondary schools (e.g. Goldstein & Thomas, 1996; Sammons, 1996; Thomas, Sammons, Mortimore & Smees, 1996) but not in the literature on primary schools. Hence the description of differential effectiveness for primary schools is a first for this report. The analysis reveals that overall measures of school effectiveness (value-added) are associated with the differential effectiveness within a school. In those schools with higher overall effectiveness (greater than expected pupil progress), low ability pupils gain a relatively greater boost than do high ability pupils. This differential boost is greater for schools where pupils make greater progress in total. It indicates that schools defined as relatively effective (in terms of contextualised pupil progress) are particularly successful with their lower ability pupils. While differences are also present between relatively effective and ineffective schools in terms of the progress of high ability pupils, these differences are less marked than for the low ability pupils. The consequences of differential effectiveness are that while all children benefit from being in an effective school rather than an ineffective one (in terms of contextualised pupil progress), the consequences are markedly greater for low ability children than

for high ability children.

Differential effectiveness, in theory, could be the consequence of a ceiling effect upon the Key Stage 2 scores which limits the amount of progress measurable for high ability pupils. However, while there is some skew in the Key Stage 2 scores suggesting lower differentiation of scores at the top end, this skew does not seem adequate to explain the differential effectiveness results. Also the differential effectiveness applies in comparisons of low ability with average ability (where ability is defined in terms of Key Stage 1 scores). The ceiling effect explanation cannot account for this. Hence this would not seem to be an adequate explanation of the differential effectiveness phenomenon.

The analysis of differential effectiveness for primary schools strongly suggests, but does not prove, that a major differentiating feature between effective and ineffective schools (in terms of contextualised pupil progress) resides in their degree of success with low ability children in particular. In a perfect world with perfectly effective primary schools, initial differences in pupil ability would be overcome by the end of primary school, i.e. initial differences would "wash out". The most effective schools are moving in this direction. While this report cannot distinguish what characteristics produce differences in effectiveness between schools, as it was not designed for this purpose, the EPPE (3-11) project will return to this topic in future reports.

Further detail of the methodology and findings is contained in Melhuish, E., Romaniuk, H., Sammons, P., Sylva, K. and Siraj-Blatchford, I. (2006), *Effective Pre-school and Primary Education 3-11 (EPPE 3-11) Tier 1: The Effectiveness of Primary Schools in England in Key Stage 2 for 2002, 2003 and 2004 Full Report* available on the EPPE website: <http://www.ioe.ac.uk/schools/ecpe/eppe/eppe3-11/eppe3-11pubs.htm>.

Acknowledgements

The authors are grateful for the advice and contribution of Dr. Alastair Leyland of the University of Glasgow.

References

Bradley, S. and Taylor, J. (2004), 'Ethnicity, educational attainment and the transition from school'. *The Manchester School*, 72, 317-346.

Cook, M. and Ludwig, J. (1998), 'The burden of "acting white": Do black adolescents disparage academic achievement?' In C. Jencks and M. Phillips (eds), *The Black-White Test Score Gap*. Washington, D.C.: The Brookings Institute.

DfEE. (1997), *Excellence in schools*. London: Department for Education and Employment.

DfES. (2005), *National Curriculum Assessment, GCSE and Equivalent Attainment and Post-16 Attainment by Pupil Characteristics in England 2004*. SFR 08/2005. www.dfes.gov.uk/rsgateway/DB/SFR/s000564/SFR08-2005v2.pdf.

DfES. (2006), *DfES (2006). National Curriculum Assessment, GCSE and Equivalent Attainment and Post-16 Attainment by Pupil Characteristics in England 2005*. SFR 09/2006. www.dfes.gov.uk/rsgateway/DB/SFR/s000640/SFR09-2006.pdf.

Goldstein, H., & Thomas, S. (1996). Using examination results as indicators of school and college performance. *Journal of the Royal Statistical Society*, 159, 149-163.

Modood, T. (2003), 'Ethnic differentials in educational performance.' In D. Mason (ed.), *Explaining Ethnic Differences: Changing Patterns of Disadvantage in Britain*. Bristol: The Policy Press.

Office of the Deputy Prime Minister. (2004), *The English Indices of Deprivation 2004 (revised)*. London: ODPM.

Sammons, P. (1996). Complexities in Judging School Effectiveness. *Educational Research and Evaluation*, 2, 113-149.

Sammons, P., Taggart, B., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Barreau, S. and Manni, L. (2006a), *Effective Pre-school and Primary Education 3-11 Project (EPPE 3-11) Variations in Teacher and Pupil Behaviours in Year 5 Classes*. DfES Research Brief 817. Nottingham: DfES.

Sammons, P., Taggart, B., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Barreau, S. and Manni, L. (2006b), *Effective Pre-school and Primary Education 3-11 Project (EPPE 3-11) Variations in Teacher and Pupil Behaviours in Year 5 Classes*.

DfES Research Report 817. Nottingham: DfES.

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, 179-202.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). *Effective Pre-school Provision*. London: Institute of Education.

Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997). Differential secondary school effectiveness: Examining the size, extent and consistency of school and departmental effects on GCSE outcomes for different groups of students over three years. *British Educational Research Journal*, 23, 451-469.

Wilson, D., Burgess, S. and Briggs, A. (2005), *The dynamics of school attainment of England's ethnic minorities. Working Paper No. 05/130*. Centre for Market and Public Organisation, University of Bristol.

Additional Information

Copies of this research brief (RBX06-06) are available free of charge by writing to DfES Publications, PO Box 5050, Sherwood Park, Annesley, Nottingham NG15 0DJ (tel: 0845 60 222 60). *Research Briefs and Research Reports can also be accessed at www.dfes.gov.uk/research/*

Further information about this research can be obtained from Jessica Dunn, W606, DfES, Moorfoot, Sheffield S1 4PQ.

Email: jessica.dunn@dfes.gsi.gov.uk

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education and Skills