Blair's Education
an international perspective

Alan Smithers
Centre for Education and Employment Research
University of Buckingham

June 2007

Contents

Foreword

Ensuring both equity and excellence in education is the aspiration that underpins the work of the Sutton Trust. Only when those from all social backgrounds are able to fulfil their academic promise will we see a more socially mobile society.

This is not impossible. During a recent trip to Finland and Sweden I was hugely impressed with the principles at the heart of their school policies. The opening statement in the Swedish Education Act says: 'All children and young people must have access to equivalent education regardless of gender, place of residence and social and financial background'. Equal opportunity is enshrined in the Finish and Swedish school systems. And international studies show that they also perform extremely well compared with other nations in the attainment of school pupils.

But how does the UK, and more specifically, England, compare on the international stage in terms of its school results? It was this question that the Centre for Education and Employment Research at the University of Buckingham sought to answer in a report published by the Trust in 2004.

The report, '*England's Education'*, by Professor Alan Smithers, concluded that English students performed above average in science, promisingly in literacy, but poorly in maths.

This latest report is an update of that review. More significantly it represents an appraisal of the Labour government's record on schools under the leadership of the Prime Minister, Tony Blair, during the last decade. It attempts to assess the Blair legacy by examining the hard figures that are available.

This is easier said than done. There are a number of national tests taken by English pupils as well as various international comparative studies. All have their own caveats; and all are open to different interpretation and presentation.

Contrary to what one might believe from reading the press, English schools perform a little above average compared with other countries, with English independent schools at the top of the international league table. However, there is no other advanced country where the gap in performance between the independent and state sectors is as large as in the UK, which is why it is so important to open up private schools to the 90% of families who cannot afford their fees.

The study also reveals that Ministers and government officials have become immersed in detailed debates about the validity of these studies as they seek to defend their policies.

It recommends that in future the educational performance of schools should be monitored independently of government. An independent and apolitical arbiter would serve the interests of the public at large, not the vested interests of a particular government. It could also perhaps monitor how equitable the school system is – how socially selective the top performing schools are, for example.

This would represent a bold move by Government. But such a body would ensure that our educational performance is objectively evaluated and that we are focused on the one target that matters to all nations: creating an education system that is both equitable and excellent.

Sir Peter Lampl
Chairman
The Sutton Trust

Executive Summary

As it became clear that Tony Blair's term as Prime Minister was coming to an end the Sutton Trust commissioned the Centre for Education and Employment Research at the University of Buckingham to make a quantitative assessment of how educational performance in England's schools had changed during his time in office, particularly from an international perspective. Unlike previous governments, Blair's had taken responsibility for 'delivery' of the performance of schools and set numerical targets to measure progress. So, ten years on, how has it done?

We summarise here our findings in three areas: trends in national tests taken by school pupils and the interpretation of these results; trends in international comparisons of educational performance, and government responses to the analyses; and broader lessons for school systems emerging from the international studies. This is followed by our overall conclusions.

National Tests

The Labour government has pointed to considerable progress on the basis of rising scores in national tests and examinations. There does appear to have been an appreciable improvement in the test scores of children both at the end of primary school and secondary school.

But there have been challenges as to whether these scores signify real improvements in education. When the scores in the national tests are compared with independently set tests, the latter do not show the same rate of progress. One likely reason for this is that since schools are themselves judged on the results of the national examinations they train up pupils specifically to take them. More controversially, some have suggested that the tests administered may have got easier over time.

In primary schools there is nonetheless evidence of improvement – more in maths than in reading.

International Comparisons

In *England's Education* (2004) we reviewed international studies through to the OECD's 2000 Programme for International Assessment (PISA). We found that while England had a generally good record in science, it usually came low down in maths, with the notable exception of the PISA study carried out in 2000. This apparent improvement was warmly welcomed by the government.

Updated Analysis

Since 2000 there have been further rounds of PISA and the Trends in Mathematics and Science Study (TIMSS). The results have been mixed for England. And this has prompted different responses to the studies from government - often on debateable grounds. We summarise the results and responses for the primary and secondary phases of education.

Secondary Pupils

PISA focuses on 15-year-olds and a particular interest for England in 2003 was whether the good maths results from 2000 study had been sustained. In fact, they had not - suggesting the 2000 results were a blip. In 2003 England slipped down the international rankings from ninth to sixteenth place.

However, the government did not acknowledge the results of the 2003 study on the grounds that the response rate for pupils taking tests had been too low. This is despite the fact that the response rate in 2003 was not very different from that in 2000. And for both studies the response rate in England was not that different from that of the USA, which was included in the main study.

The TIMSS 2003 evaluation for the secondary school pupils meanwhile showed a poor performance from England in maths.

But once again the government did not accept the findings arguing that participation was too low to be meaningful.

Primary Pupils

England rose four places in the 2003 TIMSS evaluation for primary school pupils from 10th to 6th in maths on the basis of an improvement of 47 points in test scores.

The government claimed these results were a vindication of its policies. Again this was in spite of the fact that the initial response rate was lower than that of the PISA 2003 study.

England came third in the international table in the 2001 Progress in International Reading Literacy (PIRLS) study, when again initial participation rates were lower than the PISA 2003 study.

The government argued that these good results were a consequence of its primary school literacy and numeracy strategies.

International Trends

The international studies are designed mainly to shed light on what works in education. In the PISA 2003 study, for example, it is possible to compare six methods of school admission – by residence, academic record, recommendation from feeder school, faith, special needs/interests, and siblings. None of these methods was correlated with higher maths scores for countries.

School Autonomy

But the degree of school autonomy in a country does appear to have a clear impact on pupil test scores, according to the figures produced by the PISA study.

Independent schools tended to do better than government schools across a range of countries, even when the social background of pupils is taken into account. Similarly, government-funded private schools tended to do better than government-run schools. A likely explanation is the relative freedom schools enjoy from government control.

Gender Effects

While girls do better in all countries in reading and boys do better in all countries bar one in maths, some educational systems can be characterised as 'girl-friendly' and others 'boy-friendly'. Iceland is the exemplar of a girl-friendly system with the highest difference for reading, the only country with girls ahead in maths, and with the biggest advantage to girls in science.

Conversely, Korea had the largest gaps in favour of boys in maths and science, and almost the lowest advantage to girls in reading. What makes a school system, or indeed a society, more girl- or boy-friendly?

Societal Effects

The 2003 PISA study shows that despite spending more on education wealthier countries do not necessarily see better maths results. The nature and extent of immigration though was found to have a major effect on test scores.

Personal Development

The emphasis placed by the Blair government on improving test scores in schools may have led to neglect of the less measurable, but important aspects of education associated with the personal development of pupils - variously referred to as 'soft skills', 'life skills', 'value systems', 'character', 'emotional intelligence' and 'well-being'.

This is all the more worrying given the recent United Nations Children's Fund (UNICEF) report that concluded the UK was the worst performing country among 21 nations in terms of the well being of children.

Conclusions

Our overall assessment of the impact of Blair's ten years in office on school education is - tentatively - that some progress has been made during the decade.

However, this assessment is inconclusive as it is so hard to say with certainty what the extent of any progress has been.

Test and examination scores have gone up, but the monitoring has been largely under government control and it has had a vested interest in the results.

The government would also claim that it made huge extra investment in education, increased the number of teachers and produced a more diverse secondary education system.

But there are also contrary indications. Truancy has increased, too few pupils are in education and training at age 17, demand for independent schools has remained buoyant, and the benefits of diversity *per se* are uncertain.

International comparisons meanwhile do not fully bear out the government's claims and there are also the disturbing findings from the Unicef study.

The difficulty in deciding whether education has been getting better, and if so to what extent, leads us to make the first of two proposals:

We recommend that an independent body should be set up to monitor the performance of England's school system both over time and relative to other countries.

The body's main task would to ensure accurate methods of assessment, and apply rigorous and dispassionate analysis to the results. Its remit would include both the academic and broader purposes of education. It would provide objective evidence in relation to government policies and aim to ensure confidence in school standards. In short it would be about: what works best on a long-term basis.

The second proposal arises from evidence in the international comparisons, but also supported by national exam data, that schools having autonomy over the way they are run tend to do better.

Schools should be given more genuine autonomy, but within a framework which ensures equity.

Freedom over ethos, teaching and recruitment needs to be exercised within a framework that ensures fair admissions. Otherwise pupils from poorer backgrounds are likely to lose out and segregation in society will increase. An urgent issue to be addressed is what constitutes fair admission.

1. Introduction

1.1 In 1997 Labour came to power with education at the heart of its social policies. In an attempt to drive up performance it radically changed the relationship between central government and the education system. Instead of contenting itself with legislating and leaving the implementation to schools and local authorities, it took upon itself responsibility for delivery. Learning from its friends in industry, it in effect began treating schools as the branches of a large company setting performance targets against which they would be judged. For this it needed a measurable product, and test and examination results became the equivalent of barrels of oil or cans of baked beans. In just the same way as the chief executive and senior managers of a large company pore over outputs as the ultimate test of their strategies, so Number 10, the Department for Education and Skills (as it is now) and the Treasury (though not always in concert) have come to regard test and examination results as the ultimate arbiter of their education policies.

1.2 In this report we assess how England's education has fared in terms of test and examination results in the decade of the Blair government. We will look at the national results. These, however, are internal to the system and, without external validation, not necessarily a good guide to whether or not education has been improved. Higher marks could come about in a number of ways other than genuine improvements, for example, through training in test-taking techniques. Of more interest, therefore, are the tests not directly under the government's control and chief among these are those undertaken by samples of pupils in international studies. In 2000 a major comparative programme under the auspices of the Organisation for Economic Co-operation and Development, the Programme for International Student Assessment (PISA), was added to those conducted by the International Study for the Evaluation of Educational Achievement (IEA). England, after many years as an also-ran in the international league tables, suddenly appeared to be doing rather well. Ministers and DfES officials claimed the credit and support for their education policies. This might be thought to have been premature since PISA tested 15-year-olds and the main thrust of the government's policies in the three years it had been in power was towards improving the primary education, so the age group tested will have been largely untouched.

1.3 In *England's Education* (2004), we put performance under scrutiny going back to the first IEA study in 1964. While recognising that international comparisons can be criticised from a number of points of view, we attempted a broad-brush interpretation of the various studies and concluded that in science, particularly after the introduction of a national curriculum, England had been doing relatively well. In reading, it was not possible to say since England had not taken part in the only other major study, the IEA study in 1991. But in maths we appeared to be the dunces among industrialised nations, coming well down in the various comparisons. That is, with the notable exception of P1SA 2000 where England came equal ninth with Switzerland, long cited as an example we should follow (Prais, 1995). But *pace* the cries of delight there were reasons for caution. Maths, strictly mathematical literacy, was only a minor subject in PISA 2000 and the testing was limited to a few questions (very few to none in some cases) in geometry and algebra. England's good performance could, therefore, have been just a blip. The second cycle of PISA

was conducted in 2003 and the results have been put into the public domain. In it, maths is the major area of inquiry and questions on arithmetic, and data and probability, have been added to those on geometry and algebra. It would, therefore, seem to present an ideal opportunity of seeing whether the improvements indicated in PISA 2000 are real.

1.4 Unfortunately, there is a snag. OECD/PISA, with agreement of the UK government, declined to include England in the main analysis of the 2003 results because it claimed that the response rates of schools and pupils did not meet its minimum requirements. While technically this is correct, it is also difficult to understand. Chart 1.1 shows the response rates for the UK and USA for schools and pupils in both 2000 and 2003. They look remarkably similar. But while the UK in 2000 and the USA in both 2000 and 2003 were included in the main report, the UK in 2003 was not. In part, therefore, this report undertakes the analysis which OECD/PISA should have carried out.

Chart 1.1: Percentage Response Rates in PISA 2000 and 2003

Country	2000			2003		
	School		Pupils	School		Pupils
	Before[1]	After[2]		Before[1]	After[2]	
UK	61	82	81	64	77	78
USA	56	70	85	65	68	83

Sources: PISA (2001), Table A3.2, page 235 and PISA (2004), Table A3.3, page 327.

1. Schools participating initially, before replacement.

2. Schools participating after replacing non-respondents.

1.5 Fortunately, the scores for England and Scotland (which did meet the response requirements) do appear on the PISA website. We have drawn on these data, alongside the results of other international studies including the IEA's Trends in Mathematics and Science Study (TIMSS), also in the field in 2003, England's national tests and examinations, and other sources, to assess quantitatively the impact of the Blair government's educational policies. Specifically we ask:

- How do England's results in the Programme for International Assessment (PISA) compare with those of other countries?

- What do the time courses of England's national tests and examinations show?

- How do the results obtained in PISA compare with those found in TIMSS and the Progress in International Reading Literacy (PIRLS) study conducted by the International Association for the Evaluation of Educational Achievement (IEA) based at Boston College in the United States?

- Can the differences in the performance of countries in PISA and TIMSS be traced back to differences between schools?

- To what extent do the differences between countries appear to be associated with societal factors other than education?

- Performance in tests of literacy or knowledge in reading, maths and science tell us only about particular aspects of education. How do those results compare with the findings of the United Nations Children's Fund (Unicef) in its assessment of the lives and well-being of children and adolescents in economically advanced nations?

- What do the national and international test results tell us about the success or otherwise of government policies? Is it possible to say whether England's performance has been improving?

2. Relative Performance on PISA

2.1 In Chart 2.1 we present an overview of the PISA results in 2000 and 2003 with England included. The core of the study is the countries of the Organisation for Economic Co-operation and Development (OECD). But the Slovak Republic and Turkey did not take part in PISA 2000 and the Netherlands had a response rate of only 27 per cent and so was excluded. The other 27 countries participated on both occasions (if we allow back the UK), as did Russia, Brazil, Latvia and Liechtenstein from outside the OECD. Flemish and French Belgium perform very differently and have been entered separately. Scotland's results are shown as well as England's making 33 countries and regions in all.

2.2 There was a high degree of correlation between the overall scores for the 33 countries and territories in 2000 and 2003. The Pearson coefficient, a measure of the correlation, came out at +0.930 which is significant at the one per cent level. In both years there was also a high degree of inter-correlation between maths, science and reading (the lowest coefficient was +0.889 - between science and reading in 2003) suggesting that the scales were measuring very much the same thing. PISA describes them all as tests of literacy in the sense of ability to apply knowledge – mathematical literacy, scientific literacy and reading literacy - and it could be that this underlying organising principle is responsible for the congruence. At all events, it justifies combining the scores in this part of the analysis.

2.3 Overall, England's relative standing dropped between 2000 and 2003, from eighth to eleventh place. Other movements included Poland up 6 places, Latvia up 6, Germany up 5, Luxembourg up 5, Switzerland up 4 and the Czech Republic up 4. Among the fallers were the USA down 5, Norway down 5, Spain down 4, Ireland down 4 and Japan down 4. The undoubted star was Liechtenstein, which rose 16 places from twentieth to fourth, but one can question the wisdom of comparing the 338 15-year-olds in the Principality with the hundreds of thousands in other countries. Liechtenstein put out a celebratory press release on 7 December 2004, but did not offer any explanation of the dramatic climb up the table. But one can hazard that since only twelve schools were involved the pupils in them could have been trained with relative ease to the requirements of the PISA tests. PISA itself discounts the apparent improvement in another small country, Luxembourg, as due to substantial changes in the assessment arrangements 'to reduce linguistic barriers for students' (PISA 2004, page 30), but whether any of the other apparent changes reflect modifications to educational practices awaits further investigation. The further PISA cycles planned for 2006, 2009 and subsequent years will help to reveal whether there are consistent trends.

2.4 Maths was the surprising and encouraging score for England in PISA 2000 since it was out of line with the findings of the previous forty years of TIMSS' comparisons. England's performance in maths in 2003 is, therefore, of special interest, particularly as maths is the main part of the 2003 round whereas it was only peripheral in 2000. Table 2.1 indicates that the 2000 result may have been a blip. England's score falls from the 529 obtained in 2000 to 507 in 2003, dropping it from equal 9th (with Switzerland) to 16th in the rankings.

Chart 2.1: PISA 2000 and 2003

Country or Region	Maths		Science		Reading		Total			
	2000	2003	2000	2003	2000	2003	2000		2003	
Finland	536	544	538	548	546	543	1,620	*3*	1,635	*1*
Korea	547	542	552	538	525	534	1,624	*2*	1,614	*2*
Belgium Flemish	543	553	519	529	532	530	1,594	*5=*	1,612	*3*
Liechtenstein	514	536	476	525	483	525	1,473	*20*	1,586	*4*
Japan	557	534	550	547	522	498	1,629	*1*	1,579	*5*
Canada	533	532	529	518	534	528	1,596	*4*	1,578	*6*
Australia	533	524	528	525	528	525	1,589	*7*	1,574	*7*
New Zealand	537	523	528	521	529	522	1,594	*5=*	1566	*8*
Scotland	533	524	522	514	526	516	1,581	*9*	1,554	*9*
Switzerland	529	526	496	513	494	499	1,519	*14*	1,538	*10*
England	529	507	533	519	523	506	1,585	*8*	1,531	*11*
Sweden	510	509	512	506	516	514	1,538	*12*	1,529	*12*
Czech Republic	498	516	511	523	492	489	1,501	*17*	1,528	*13*
Ireland	503	503	513	505	527	516	1,543	*10*	1,524	*14*
France	517	511	500	511	505	496	1,522	*13*	1,518	*15*
Iceland	514	515	496	495	507	492	1,517	*15*	1,502	*16*
Germany	490	503	487	502	484	491	1,461	*22*	1,496	*17*
Austria	515	506	519	491	507	491	1,541	*11*	1,488	*18*
Poland	470	490	483	498	479	497	1,432	*25*	1,485	*19*
Denmark	514	514	481	475	497	492	1,492	*19*	1,481	*20*
Norway	499	495	500	484	505	500	1,504	*16*	1,479	*21*
Hungary	488	490	496	503	480	482	1,464	*21*	1,475	*22*
USA	493	483	499	491	504	495	1,496	*18*	1,469	*23*
Latvia	463	483	460	489	458	491	1,381	*30*	1,463	*24*
Belgium French	491	498	467	483	476	477	1,434	*24*	1,458	*25*
Luxembourg	446	493	443	483	441	479	1,330	*31*	1,455	*26*
Spain	476	485	491	487	493	481	1,460	*23*	1,453	*27*
Italy	457	466	478	486	487	476	1,422	*26*	1,428	*28*
Portugal	454	466	459	468	470	478	1,383	*28*	1,412	*29*
Russia	478	468	460	489	462	442	1,400	*27*	1,399	*30*
Greece	447	445	461	481	474	472	1,382	*29*	1,398	*31*
Mexico	387	385	422	405	422	400	1,231	*32*	1,190	*32*
Brazil	334	356	375	390	396	402	1,105	*33*	1,148	*33*

Sources: PISA (2001) Tables 2.3a, 3.1 and 3.3, pages 253, 259 and 261 and PISA (2004) Tables 2.5c, 6.2 and 6.6, pages 356, 444 and 448 plus PISA (2006) spreadsheets.

2.5 The analysis can be refined, as in Chart 2.2, by comparing only the scales that were used on both occasions. These were 'space and shape' and 'change and relationships', essentially geometry and algebra. England's score on these two scales falls from 529 to 496, leaving it, in 2003, down in 19th position. Another notable finding using just these two scales is that Flemish Belgium comes top. This is discussed by De Meyer, Pauly and Van de Poele (2005), and it underlines the considerable difference from French Belgium discussed by Smithers (2004).

Chart 2.2: Maths (Geometry and Algebra) 2000-2003

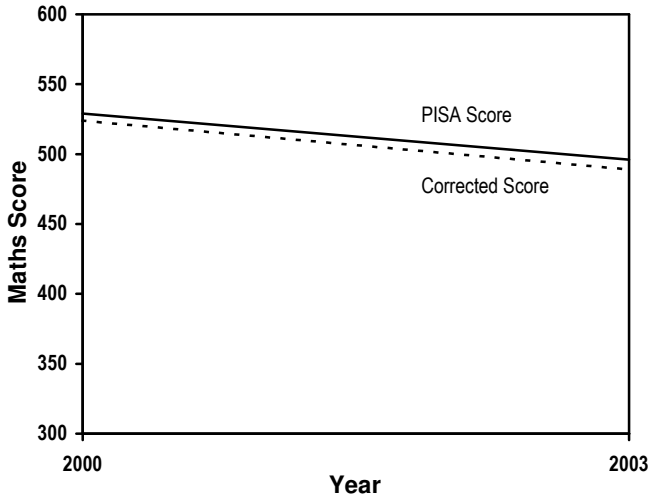
Country or Territory	2000		2003	
	Score	Rank	Score	Rank
Belgium Flemish	543	*3*	557	*1*
Korea	547	*2*	544	*2*
Finland	536	*5*	544	*3*
Japan	557	*1*	540	*4*
Switzerland	529	*=9*	536	*5*
Liechtenstein	514	*=13*	536	*6*
Czech Republic	498	*19*	528	*7*
Canada	533	*=6*	523	*8*
Australia	533	*=6*	519	*9*
New Zealand	537	*4*	518	*10*
Austria	515	*12*	514	*11*
Denmark	514	*=13*	514	*12*
Scotland	533	*=6*	513	*13*
Iceland	514	*=13*	508	*14*
France	517	*11*	507	*15*
Germany	490	*22*	507	*16*
Sweden	510	*16*	506	*17*
Belgium French	491	*21*	501	*18*
England	529	*=9*	496	*19*
Luxembourg	446	*31*	495	*20*
Poland	470	*26*	491	*21*
Ireland	503	*17*	489	*22*
Norway	499	*18*	488	*23*
Hungary	488	*23*	488	*24*
Spain	476	*25*	484	*25*
Latvia	463	*27*	484	*26*
USA	493	*20*	474	*27*
Russia	478	*24*	473	*28*
Italy	457	*28*	473	*29*
Portugal	454	*29*	458	*30*
Greece	447	*30*	442	*31*
Mexico	387	*32*	388	*32*
Brazil	334	*33*	355	*33*

Sources: PISA (2001) Tables 2.3a, 3.1 and 3.3, pages 253, 259 and 261 and PISA (2006) spreadsheets.

2.6 Against the apparent relative decline in England's maths performance in 2003 there is always the defence that the participation level was too low for the results to be meaningful. This turns on the degree and direction of bias that under-representation of particular groups may have introduced. A very careful analysis of the likely magnitude of any response bias has been conducted by Micklewright and Schnepf (2006) and they concluded that, if anything, the scores obtained were over-estimates due to under-representation of low achievers. Using the Key Stage 3 and GCSE

results as reference scores, a combination of three methods of estimating bias in the PISA means - z-scores, regression and response weights - suggests that in 2000 England's maths score was inflated by about 5 points and in 2003 by about 7 points. Similarly, the reading scores were over-estimates by around 6 and 7 points respectively, and the science scores by 6 and 8 points. Chart 3.3 portrays the effect of these corrections on England's maths scores. Applying a correction of 7 points to England's 2003 score on the geometry and algebra scales lowers it to 489, equal 21st (with Ireland) out of the 33 nations and regions.

Chart 2.3 England's Maths Scores



Sources: PISA (2001) Tables 2.3a, 3.1 and 3.3, pages 253, 259 and 261 and PISA (2006) spreadsheets, and Micklewright and Schnepf (2006) pages 59 and 82.

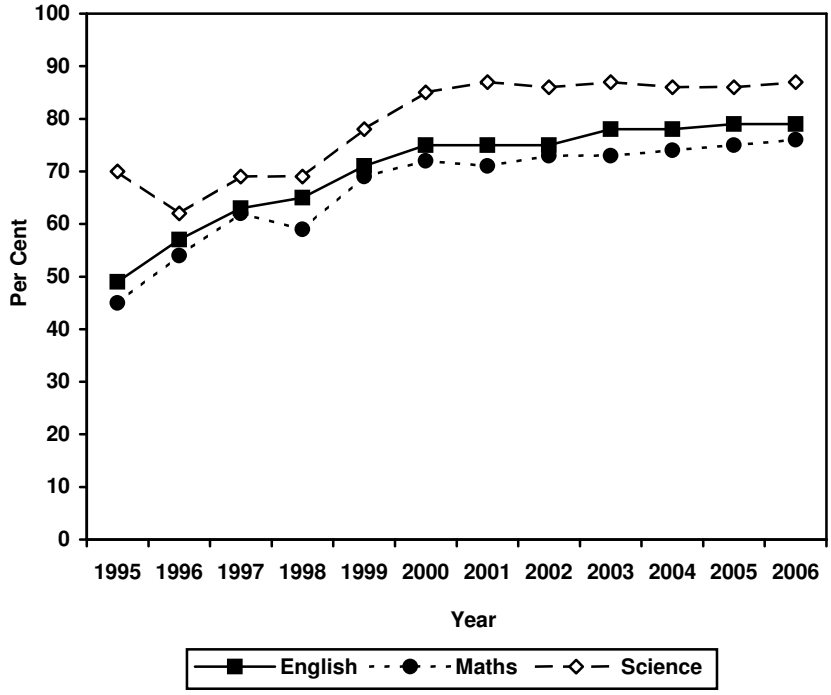
Comment

2.7 Whatever can be read into the results, it does not look as though they can be interpreted as showing that England's performance in maths has been improving relative to other OECD countries; rather the reverse. The picture in reading would seem to be similar. This appears to be at odds with the encouraging indications from the national tests and examinations and it is to these we now turn.

3. National Tests and Examinations

3.1 Buoyed by a landslide election victory in May 1997, the first Blair government came to power eager to make its mark in education. Within a week a new Standards and Effectiveness Unit (SEU) was established and, within two, ambitious targets for the literacy and numeracy of 11-year-olds had been declared. The secretary of state for education, David Blunkett, said that his head would be on the block if the targets were not met. This was reported as he would resign. His schools minister, Estelle Morris, was drawn into a categorical commitment. Asked in 1999 by the Conservatives in the House of Commons (*Hansard Debates* for 2 March), if she would commit to Blunkett's pledge she said, 'Of course I will; I have already done so. Indeed, I generously commit the under-secretary (Charles Clarke) too. We speak with one voice'. Michael Barber, the first head of the SEU, was so encouraged by the early results that he claimed in 2000, in the American magazine *Educational Week,* that 'large scale reform is not only possible but can be achieved quickly.' Ten years on from the 1997 election, how has it worked out? The thing about numerical targets is that progress against them is checkable. In this chapter we attempt a quantitative evaluation.

Chart 3.1: % Achieving Level 4 or Higher at Key Stage 2



Key Stage 2

3.2 The early results were very encouraging. In Blair's first three years in office the percentage of pupils reaching the expected levels in English and maths for 11-year-olds rose by ten percentage points and looked to be well on course to reach the targets of 80 per cent in English and 75 per cent in maths by 2002. It was against this background that the government was ready to accept that the improvement in the

PISA 2000 maths results owed something to its education policies, even though the 15-year-olds tested would not have been touched by them.

3.3 But closer scrutiny of England's national tests raises the question of whether the target-setting actually had anything to do with the rising scores. Chart 3.1 shows that the trends in the English, maths and science results were generally upward. There had been, however, a 14 percentage point rise in the English results and a 17 percentage point increase in the maths results in the three years prior to Labour taking office. In science, where no target had been set, the trajectory was similar but, if anything, more impressive. The improvements stalled from the year 2000 onwards, but such was the belief in targets, that Estelle Morris who succeeded David Blunkett as education secretary in 2001, reacted by raising them to 85 per cent for both English and maths.

3.4 Quite what underlies the shape of the trends in Chart 3.1 has become a matter of fierce dispute. A plausible description of the curves is that the feedback from the first national tests introduced by the Conservative government in 1994 (though the results were not made generally available until 1995) gave a healthy shock to the system. From the comfortable assumption that England had some of the best primary schools in the world it was plain that over half of 11-year-olds were not able to read, write and handle numbers to the expected standards. This feedback, of itself, would have been a spur to improvement. It is open to question whether pressure from the centre through targets and the associated numeracy and literacy strategies under the Labour government has added anything. The scores certainly have gone up, but to what extent do they reflect underlying improvements or could they simply show better test taking techniques?

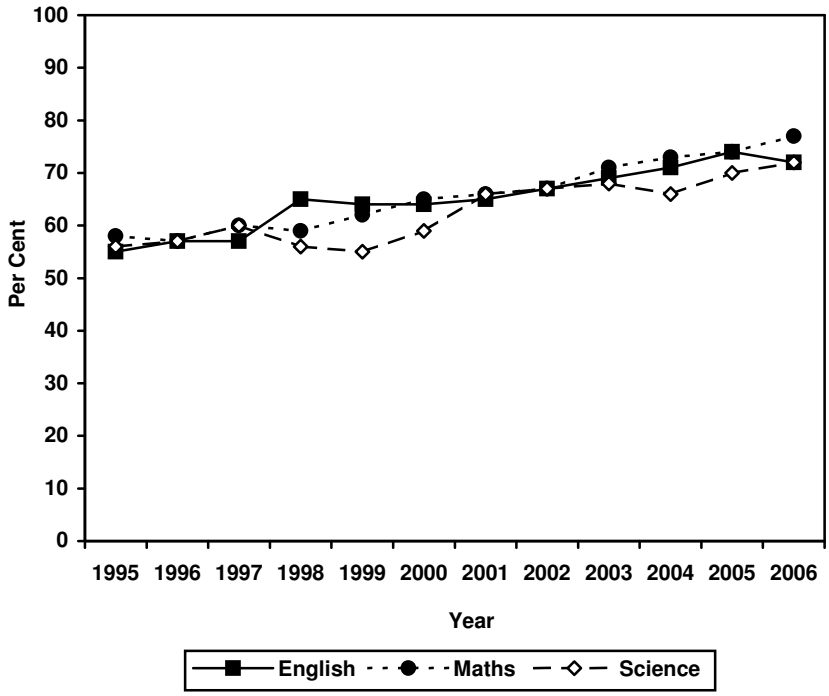
3.5 A report by the Statistics Commission on 23 February 2005, concluded: 'The Commission believes that it has been established that (a) the improvement in KS2 test scores between 1995 and 2000 substantially overstates the improvement in standards in English primary schools over that period, but (b) there was nevertheless some rise in standards.' The issue had been drawn to its attention by an article by Professor Tymms (2004) of the University of Durham which argued that the national tests were over-stating the improvements. He drew in particular on a study commissioned by the Qualifications and Curriculum Authority from the University of Cambridge Local Examinations Syndicate (Massey *et al* 2003). In the comparability-over-time project randomly assigned groups of 11-year-olds in Northern Ireland, who had not seen them before, were given the 1996 and either the 1999 or 2000 versions of the Key Stage 2 English and maths tests at the same time. In English, they did significantly better on the 1999 and 2000 tests, with Massey calculating that they were 5-7 marks easier than the 1996 version. It is part of Tymms' case that it is hard to ensure consistency over time, and some slippage had occurred contributing to the rising scores. He has also been able to compile a list of twelve other independent studies of reading which also failed to record the improvements that might have been expected from the national results. In maths Massey *et al* found that the test standards were similar in 1996 and 1999 and this was supported by standardised test data from local authorities.

3.6 The Statistics Commission's conclusions were hotly challenged by the DfES. On 3 March 2005, Sir David Normington, the permanent secretary, wrote to Professor David Rhind, Chairman of the Statistics Commission, asking the Commission to 'revisit your conclusions on the Peter Tymms article and set the record straight'. He argued that the Commission had not distinguished sufficiently between the English and maths results (most of the studies were to do with reading and there was independent evidence of improvement in maths performance). Neither did the Commission, in the DfES' view, have the evidence or the expertise to comment on whether teachers were teaching to the test. The permanent secretary further claimed that international evidence and Ofsted inspections confirmed the significant rise in performance which the Key Stage tests showed. The international evidence he had in mind was from the IEA's TIMSS and PIRLS studies, which we will look at in detail in the next chapter.

Key Stage 3

3.7 We have discussed the Key Stage 2 results at length because they raise in sharp form how difficult it is to be sure that seemingly dramatic rises in test scores reflect real improvements and even if they do whether they can be attributed to educational policies. It is, of course, the Key Stage 3 tests taken by 14-year-olds that are the nearest in age to the OECD/PISA samples of 15-year-olds, but much of what has been said about the Key Stage 2 tests applies here also. Chart 3.2 shows how the Key Stage 3 scores have changed since they first became reliably available in 1995.

Chart 3.2: % Achieving Level 5 or Higher at Key Stage 3



3.8 The trend in all three subjects is upward, although without the rapid leap that characterised the first years of the Key Stage 2 tests. This continued even in the period 2000-2003 when the PISA studies showed at best no improvement and more likely a drop relative to the other countries. What are we to make of this apparent

discrepancy? The DfES would argue (by extension from Normington, 2005) that the increases in Key Stage 3 scores represent real improvements in performance on measures specifically designed to test the attainment targets of the national curriculum, and the PISA scores have to be discounted because England's participation levels did not meet the specified requirements. The contrary view, by extension from Tymms (2004), would suggest that there may have been some improvement, but this is over-stated by the scores which have been boosted by slippage in standards and teaching to the test. Where a more general test is used, the improved performance might not be nearly so evident. It is also possible that rising performance in internal tests could show up as a decline in international comparisons if other countries had improved to an even greater extent.

3.9 Given the efforts and resources that have gone into improving education it would be astonishing if there had been no impact at all. On the other hand, politicians have perhaps put more weight on the Key Stage tests than they can really bear, in seeking to judge schools and teachers by them as much as the pupils. It is not surprising, therefore, that all the stops should have been pulled out by the schools in an attempt to maximise the scores on the particular tests and that this improved performance may not generalise fully to other tests. However to counter this kind of argument the government might also point to the progressively improving GCSE and A-level results.

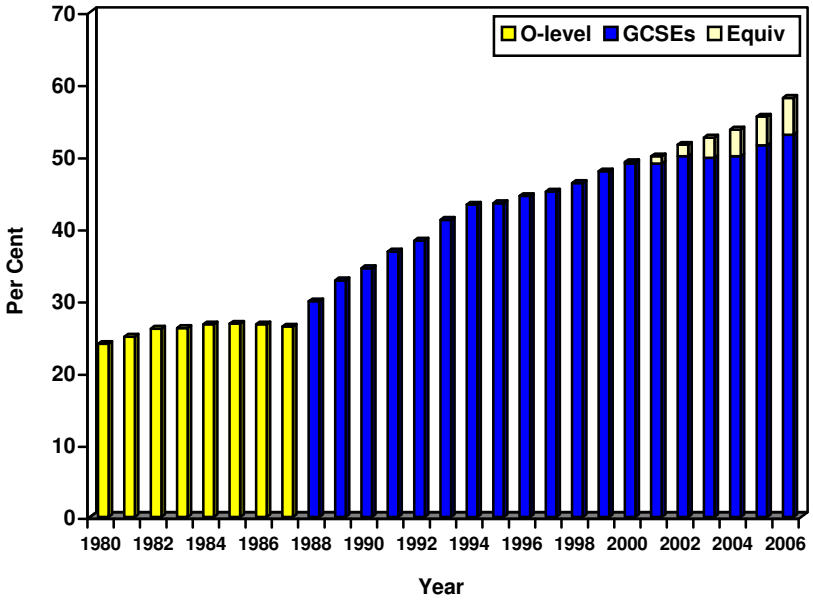
GCSE

3.10 Taken at face value the results of the GCSE and A-level examinations present a very encouraging picture. Charts 3.3 and 3.4 indicate improvements for each of the years of the Blair governments. In the case of GCSEs, Chart 3.3 shows that, in 2006, 58.1 per cent achieved five good GCSEs or their equivalent against the 45.1 per cent when they came to office. But to what extent can the increase be attributed to targets and benchmarking? Looking at the period before Blair came to power we can see that the percentage achieving five good grades has increased every year since the GCSE came on stream in 1988. From just over a quarter of the age group who attained this level on the predecessor examination, the GCE Ordinary Level, the proportion rose year by year to reach nearly 60 per cent in 2006. The year-on-year increases were highest from 1987 to 1994. It is clear that a major factor was the switch from O-level to GCSE which opened up the examination to many more candidates (only 40 per cent of the age group got to take O-level) and was more user-friendly through modularisation and course work. The early increases can be interpreted as the schools becoming more used to the new arrangements. So far there has not been much sign of the rising scores at Key Stage 2 producing further improvement in GCSE results as might have been expected if they reflected lasting gains in literacy and numeracy.

3.11 In so far as it is possible to detect an effect of the Blair governments' policies it is not entirely to the good. There has been great pressure from the centre for continually improving results, not least because the DfES signed a public service agreement with the Treasury to deliver in the period 2002-05 a two per cent increase per year in those getting five good GCSEs. This it was never quite able to achieve in spite of going with the grain of GCSE performance and using the sticks of league tables and inspections. But what did happen is that schools became increasingly

attracted to a General National Vocational Qualification in Information and Communications Technology where a pass, through a disproportionate tariff, yielded the equivalent of four good GCSEs for work being no more demanding and taking up the curriculum time of not much more than one.

Chart 3.3: Five Good GCSEs or Equivalent



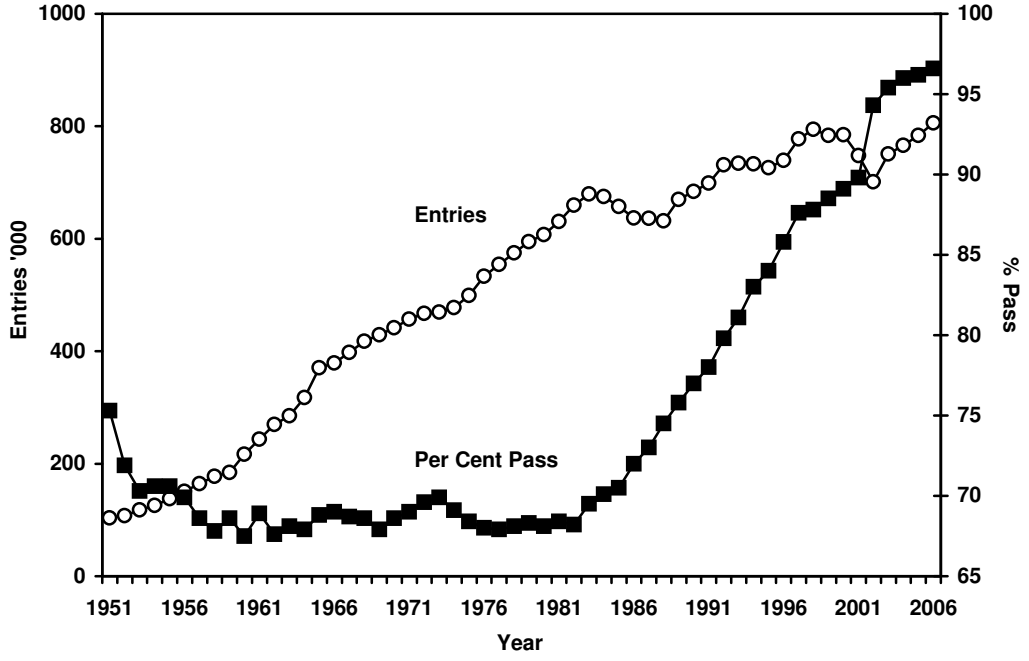
3.12 Chart 3.3 also brings out the GNVQ's growing influence on the official statistics. An hitherto unpublished breakdown obtained by BBC News (2006) showed that in 2006 actual GCSEs including those in vocational subjects accounted for only 53 percentage points of the 58.1 per cent of the secondary school pupils reported as achieving five good GCSEs. Much of the recent rise has been down to GNVQs and other qualifications. When the measure was changed to include both English and maths the percentage reaching the standard in 2006 dropped to 45.8 per cent. Whatever pressure from the centre has achieved in raising GCSE scores, it is clear that there has been an impact on the qualifications pupils are being entered for and that this may not always have been for intrinsic educational reasons.

A-Level

3.13 The story of A-levels is again one of continually improving results. In fact, Chart 3.4 shows that the pass rate has increased every year since 1982 when it was 68.2 per cent. By 2006 it had reached 96.6 per cent. In the trend line it is possible to make out the effects of major structural changes. For the first 30 years of A-levels the pass rate hovered around 70 per cent because this was set as the predetermined level. It had the unfortunate consequence of failing 30 per cent of those who had showed real promise by passing five O-levels. When the change was made in 1979 to criterion referencing - attempting to assess against a standard rather than filling quotas – it opened the way to higher pass rates, with an increase of about a percentage point a year. In 2002 there was a major fillip when the exam became completely modularised and a half-way house, the Advanced Subsidiary, was introduced.

3.14 Not only have more been passing A-level, but more have been taking it. Entries did show some downturn in the 1980s as the number of 18-year-olds began to fall sharply, but this was counteracted by the arrival of GCSE in 1988. There was also a fall around the time A-levels were restructured in 2000, but again there has been recovery. In summary, the story at A-level has been more of the more doing better. It is apparently good news in which successive governments have wished to bask only to be frustrated by the deep scepticism of the media. Each August the publication of results has been greeted with cries of dumbing down.

Chart 3.4: A-Level Entries and Passes



3.15 These suspicions may have some substance. In 2005 Tymms *et al* compared the grades awarded at A-level with a test administered every year since 1988 by the Curriculum, Evaluation and Management Centre at Durham University as part of its assessment service for schools. It found that higher A-level grades were being awarded to pupils of the same ability as shown by the independent test, and concluded that A-levels have been more leniently graded over the years. The drift was discernible year by year for the 15 years considered and not associated with any particular government.

Comment

3.16 Test and examinations are not rulers or thermometers. The numbers assigned in the assessment of educational performance carry less meaning than the measurements of length and temperature. They are at root subjective with people attaching numbers to other people, and they are open to interventions of various kinds. The test and exam scores stand as a proxy for the education that we hope is taking place. But as commentators have pointed out, 'when a surrogate measure is made a target for conducting social or economic policy, then it loses the information content that would qualify it to play such a role' (Goodhart, 1975). Or again, 'when a measure becomes a target, it ceases to be a good measure' (Strathern, 1997).

3.17 As some Ministers found to their cost, it is unwise for governments to stake their reputations on targets for students' scores. Not only can it become embarrassing for them, but it tends to distort education as schools strive to make the numbers come out right. As Campbell (1969) sagely wrote nearly 40 years ago, the danger is that those who 'have so committed themselves in advance to the efficacy of the reform … cannot afford honest evaluation'. The government's presentation of results inevitably becomes part of the narrative in which they seek to demonstrate their success. Studies outside their control like those of Tymms and his colleagues in Durham become a threat to be challenged vigorously.

4. TIMSS AND PIRLS

4.1 Instead of dwelling on the disappointing results from PISA 2003, the government focused attention on the low response rate in England through, among other things, sponsoring research and organising conferences aimed at finding what could be done to improve participation in the next round (Sturgis, Smith and Hughes, 2006). But when it came to attempting to fend off the challenge discussed in the previous chapter from Tymms, supported by the Statistics Commission, the DfES through its permanent secretary, David Normington (2005) made much of the results from two other international studies, Trends in Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy (PIRLS), conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA) based at the International Centre, Boston College, in the United States.

4.2 Normington argued in relation to mathematics:

> After Peter Tymms wrote his article, the results of the 2003 Trends in Mathematics and Science Study (TIMMS *sic*) were published. That highly respected international study showed that the rise in English primary pupils, maths scores between 1995 and 2003 was not merely statistically significant; at 47 points it was the highest of any participating country. Pupils take the TIMMS (*sic*) tests in Year 5. This corroborates the very significant rise in KS2 maths scores between 1996 and 2004.
>
> Peter Tymms' article points to a lack of a rise between 1995 and 1999 TIMMS (*sic*) results. This is irrelevant. Only secondary pupils participated in TIMMS (*sic*) in 1999.

4.3 Later he suggested the Statistics Commission should also take into consideration that:

> We would expect improvements in English to track the internationally and nationally corroborated improvements in maths over the same period, given that we have taken the same broad approach to standards improvement for both; and
>
> In PIRLS (Progress in International Reading Literacy) in 2001, our 10-year-olds were the third most able readers of 35 participating countries, doing significantly better than all other English speaking countries – our best ever performance in an international study.

4.4 We will examine these arguments in detail later and Normington is right to draw a distinction between the primary phase of education (where the government's improvement strategies were concentrated before 2002) and the secondary phase, but at the outset it is worth noting that the initial school response rates in TIMSS and PIRLS were, if anything, poorer than in PISA. It was only after replacement, with the potential for bias that that introduces, that the figures look more respectable. Chart 4.1 sets out the response rates. In all of them England comes within a place or two of the bottom if not at the very bottom. There is no discernible difference between the OECD/PISA and the IEA studies, and it is difficult to see why the DfES should attach such importance to the one and apologise for inadequate participation in the other, excepting the results from PIRLS and TIMSS appear more helpful to its case. In fact, PISA 2003, from which the government accepted the UK's exclusion, had the highest initial response from schools in England of all the studies listed.

Chart 4.1: Response Rates for Schools

Study	%Before Replacement	Relative to Other Countries	%After Replacement
Primary			
TIMSS-1995	63	15/17	88
PIRLS-2001	59	16/17	87
TIMSS-2003	53	14/14	82
Secondary			
TIMSS-1995	56	24/26	85
TIMSS-1999	51	15/15	85
PISA-2000	61	26/28	82
TIMSS-2003	39	10/11	54
PISA-2003	64	30/30	77

Sources: Data of each row in turn from Appendices of Mullis et al (1998); Mullis et al (2003); Martin et al (2004); Beaton et al (1997); Mullis et al (2000); OECD (2001); Martin et al (2004); OECD (2004).

Secondary-Age Pupils

4.5 When acronyms and jargon are being bandied about it is important to keep in mind who is being tested and on what. PISA tests 15-year-olds and, in so far as performance can be related to education policies, it can be taken as a comment on secondary education. Similarly, among the populations sampled by TIMSS is one centred on 14-year-olds. Both organisations were out in the field in 2003 and again three (PISA 2000) or four years (TIMSS 1999) earlier. Chart 4.2 shows the maths scores of the eight countries who took part in both studies on both occasions.

Chart 4.2: Maths Results 1999/2000 and 2003

Country or Region	TIMSS		PISA	
	1999	2003	2000	2003
Korea	587	589	547	542
Japan	579	570	557	534
Belgium Flemish	558	537	543	553
Hungary	532	529	488	490
Australia	525	505	533	524
USA	502	504	493	483
England	496	498	529	507
New Zealand	491	494	537	523

Sources: Martin *el al* (2004), Mullis et al (2000) and PISA (2004, 2001).

4.6 There is broad similarity across the scores with the top three places being filled by Korea, Japan and Flanders on all four occasions. The scores for England are similar in TIMSS and PISA 2003, but are higher in PISA 2000, adding further weight to the interpretation (Smithers, 2004) that the 2000 result, which caused so much government excitement, could have been a fluke. Within the context of broad similarity, Chart 4.2 shows that some countries appear to do appreciably better with one testing body or the other, notably Hungary in TIMSS and New Zealand in PISA.

This underlines the importance of keeping in mind what is being tested. TIMSS aims to test the understanding of curriculum maths and science whereas PISA is more concerned with the ability to apply what has been learned to real life situations. Some countries evidently are more comfortable with one approach then the other. School curricula differ across countries and although TIMSS focuses on common features inevitably the actual items will be more suited to some countries than others, raising the question of what part does curriculum match play in the results obtained. PISA aims to circumvent curricular differences by working back from everyday situations to test what its calls 'literacy', which raises the interesting question of whether since it eschews the curriculum the results shed any light at all on educational policies. It is also true that whatever the intention the PISA questions will bear more relation to the curricula of some countries than others.

4.7 In Chart 4.3 we broaden out the frame of TIMSS maths comparisons in the secondary phase to include all the countries who participated in 1995, 1999 and 2003. Again, the main impression is of stability with little discernible impact of government policies. Across the three sweeps, England is down in 13th or 14th, position with only Romania, Cyprus and Iran consistently below it. Although, as we shall see, the British government has been keen to publicise the results for the primary age range, it has said little about these secondary scores, taking refuge in the low response rates. But England's relatively poor performance in maths goes back to the earliest international comparisons with the exception of PISA 2000.

Chart 4.3: Year 9 Maths Results in TIMSS 1995, 1999 and 2003

Country or Region	1995		1999		2003	
	Score	Rank	Score	Rank	Score	Rank
Singapore	643	*1*	604	*1*	605	*1*
Korea	607	*2*	587	*2*	589	*2*
Hong Kong	588	*4*	582	*3*	586	*3*
Japan	605	*3*	579	*4*	585	*4*
Belgium Flemish	565	*5*	558	*5*	537	*5*
Netherlands	541	*7*	540	*6*	536	*6*
Hungary	537	*8*	532	*8*	529	*7*
Slovak Republic	547	*6*	534	*7*	508	*8=*
Russian Federation	535	*9*	526	*9*	508	*8=*
Latvia	493	*16*	505	*12*	508	*8=*
Australia	530	*10*	525	*10*	505	*11*
USA	500	*15*	502	*13*	504	*12*
Lithuania	477	*17*	482	*16*	502	*13*
England	506	*13*	496	*14*	498	*14*
Israel	522	*12*	466	*18*	496	*15*
New Zealand	508	*13*	491	*15*	494	*16*
Bulgaria	527	*11*	511	*11*	476	*17*
Romania	474	*18*	472	*19*	475	*18*
Cyprus	468	*19*	476	*17*	459	*19*
Iran	418	*20*	422	*20*	411	*20*

Sources: Martin *el al* (2004); Mullis et al (2000) and Beaton et al (1997).

Primary-Age Pupils

4.8 While the British government has been silent on the results from PISA 2003 and TIMSS regarding the secondary phase it has embraced those for the primary age range with gusto, and indeed there is support to be derived here. Comparison of the TIMSS maths results for Year 5, mainly ten-year-olds, in 1995 and 2003, as in Chart 4.4, shows England's score was up by 47 points in 2003 compared to 1995 (as Normington pointed out to the Statistics Commission), with its position improving from 10th to 6th among the countries who participated on both occasions. Scotland, which in 1995 was somewhat above England, fell back to 12th place.

Chart 4.4: Year 5 Maths TIMSS 1995 and 2003

Country or Region	1995		2003	
	Score	Rank	Score	Rank
Singapore	590	*1*	594	*1*
Hong Kong	557	*3*	575	*2*
Japan	567	*2*	565	*3*
Netherlands	549	*4*	540	*4*
Latvia	499	*7*	533	*5*
England	484	*10*	531	*6*
Hungary	521	*5*	529	*7*
USA	518	*6*	518	*8*
Cyprus	475	*12*	510	*9*
Australia	495	*8*	499	*10*
New Zealand	469	*13*	496	*11*
Scotland	493	*9*	490	*12*
Slovenia	462	*14*	479	*13*
Norway	476	*11*	451	*14*
Iran	387	*15*	389	*15*

Source: Martin *el al* (2004)

4.9 It could be that the signs of improving maths performance among primary age children, as would be hoped for from the intensive numeracy strategy, are beginning to show through. Even the Tymms' (2004) study which made a good case for England's national Key Stage tests over-estimating the extent of the actual improvement in reading did find more of a genuine increase in maths. Applying a correction he argued that the increase in maths scores from 1995 to 2004 was closer to 22 points (44 per cent to 66 per cent) than the reported 28 points. The main thrust of his attack was against the reported results in English where the increase he calculated was only from 48 per cent to 58 per cent instead of to the 75 per cent claimed.

4.10 Nevertheless, in both the national and international results, there are indications that primary-age children in England are doing better at maths. But more than half the increase in adjusted Key Stage 2 maths scores occurred between 1995 and 1998 (12 points) before the national maths strategy was put in place. The rising scores could, therefore, be due more to increasing familiarity with a testing regime introduced by the Conservatives than the Blair government's policies. Tymms' analyses found an

increase in English scores of only 6 per cent from 1998 to 2004. But the government has been keen to point to other international comparisons in support of its policies. As we have seen Normington's letter of complaint (2005) to the Statistics Commission cited the PIRLS study of reading, 'In PIRLS (Progress in International Reading Literacy) in 2001, our 10-year-olds were the third most able readers of 35 participating countries, doing significantly better than all other English speaking countries – our best ever performance in an international study.' Chart 4.5 sets out those results.

Chart 4.5: Reading Scores for Year 5 in PIRLS 2001

Country	Score	Rank	Years Schooling	Age
Sweden	561	*1*	4	10.8
Netherlands	554	*2*	4	10.3
England	553	*3*	5	10.2
Bulgaria	550	*4*	4	10.9
Latvia	545	*5*	4	11.0
Lithuania	543	*6=*	4	10.9
Hungary	543	*6=*	4	10.7
United States	542	*8*	4	10.2
Italy	541	*9*	4	9.8
Germany	539	*10*	4	10.5
Czech Republic	537	*11*	4	10.5
New Zealand	529	*12*	5	10.1
Scotland	528	*13=*	5	9.8
Singapore	528	*13=*	4	10.1
Russia	528	*13=*	3 or 4	10.3
Hong Kong	528	*13=*	4	10.2
France	525	*17*	4	10.1
Greece	524	*18*	4	9.9
Slovak Republic	518	*19*	4	10.3
Iceland	512	*20=*	4	9.7
Romania	512	*20=*	4	11.1
Israel	509	*22*	4	10.0
Slovenia	502	*23*	3	9.8
Norway	499	*24*	4	10.0
Cyprus	494	*25*	4	9.7
Moldova	492	*26*	4	10.8
Turkey	449	*27*	4	10.2
Macedonia	442	*28*	4	10.7
Colombia	422	*29*	4	10.5
Argentina	420	*30*	4	10.2
Iran	414	*31*	4	10.4
Kuwait	396	*32*	4	9.9
Morocco	350	*33*	4	11.2
Belize	327	*34*	4	9.8

Source: Mullis el al (2003).

4.11 PIRLS comes from the same stable as TIMSS. Its second major study was conducted in 2001 ten years after the first in 1991. The 2001 results are indeed encouraging for England putting it in third place behind Sweden and the Netherlands. But because it did not take part in the first study we cannot say whether that relative position has changed. Ministers, however, have few public doubts. In an interview with *The Guardian* (5 May 2003) David Miliband, then schools minister, said before we came to office England was average: 'I can't remember sixteenth or seventeenth. It was just redone and it's published last week: we are third best achievers for primary schools achievement at age 10 out of 35 countries.' What he had in mind was a study by Brooks, Pugh and Schagen in 1996 which attempted to show where England would have been placed if it had participated in 1991. It concluded that England would have been among a group of 13 countries with average scores. But this is hardly a sound basis for comparison.

4.12 Without in any way wishing to explain away England's seemingly impressive performance it is necessary to set it in context. Hilton (2007) has drawn attention to three important points. First, the PIRLS test was largely compiled by a team of researchers at the NFER who were also responsible for the Key Stage 2 tests so children in England would have had the advantage of being familiar with the form and content of the tests. Secondly, England had a high level of exclusions not testing children with disabilities of various kinds, nor children judged by the teachers as being unable to cope with the conditions, nor immigrants lacking fluency in reading and writing the English language. Thirdly, as Table 4.5 shows, 10-year-olds in England have received an extra year of formal schooling compared with most other countries. The scores in the primary phase are also strongly related to the age of those tested. Chart 4.6 shows the impact on the reading scores of those countries who participated in both the 1991 and 2001 studies.

Chart 4.6: Reading Scores in 2001 and 1991

Country	IEA 1991				PIRLS 2001			
	Score	Rank	Years School	Age	Score	Rank	Years School	Age
Italy	513	*1=*	4	9.9	500	*3*	4	9.8
Iceland	513	*1=*	4	9.8	486	*5*	4	9.8
United States	511	*3*	4	10.0	521	*1*	4	10.0
Greece	507	*4*	4	10.0	466	*7*	4	9.3
New Zealand	502	*5*	5	10.0	498	*4*	5	10.0
Sweden	498	*6*	3	9.8	513	*2*	3	9.8
Slovenia	493	*7*	3	9.8	458	*9*	3	9.7
Singapore	489	*8*	3	9.1	481	*6*	3	9.3
Hungary	475	*9*	3	9.7	459	*8*	3	9.3

Sources: Martin *el al* (2003)

4.13 With the exception of New Zealand, which does not show up well in primary-age reading tests, the countries at the top were those with more years of formal schooling. Sweden, which tops the ranking in Chart 4.5, is well down the list in Chart 4.6 because a younger age group are here taken for comparison. If this had been used in Chart 4.5 Sweden would have come 24th not first. Because there is as

yet no reliable comparative data for England, it will not be until the results of PIRLS 2006 become available that we shall be able to see whether it ranks consistently among the best performers in the world. But even then, as we have seen, interpretation will not be straightforward.

Comment

4.14 Looked at dispassionately the results of the PISA, TIMSS and PIRLS international data would appear to show that the standing of England relative to other countries in maths for 14 and 15-year-olds has not changed during the decade of the Blair governments. On the other hand, there are indications that the maths performance of primary-age pupils is improving and in the one reading study in which England took part it did very well. This would be consistent with Blair's first education policies concentrating on raising literacy and numeracy in the primary phase. Interpretation, however, is not easy and ministers have been too quick to claim credit when results have been good and to attempt to brush them aside when they have appeared less so. Ministers are wrong to lock themselves into test results in this way. As we shall see, many things besides education can influence the test scores obtained.

5. School Organisation

5.1 Although media interest has tended to focus on a country's position in international 'league tables', the studies are designed mainly with a view to discovering whether educational performance can be related to particular policies. One of the ways countries do differ is in how the places in secondary education are allocated. As part of PISA 2003 headteachers were asked what was a 'prerequisite' or 'a high priority' in school admission. Table 5.1 shows the replies alongside the country's score in the maths tests.

Chart 5.1: Maths Scores in PISA 2003 and Mode of School Admission

OECD Country[1]	PISA Maths 2003		Per Cent Main Method of School Admission[1]					
	Score	Rank	Residence	Acad Record	Recom Feeder School	Faith	Special Needs	Siblings
Finland	544	*1*	67.2	3.3	3.7	5.2	10.2	5.5
Korea	542	*2*	30.6	56.6	12.3	6.9	14.3	0.8
Netherlands	538	*3*	6.5	70.0	84.2	17.3	16.7	3.2
Japan	534	*4*	29.7	88.1	36.7	10.6	38.1	3.3
Canada	532	*5*	75.3	13.1	18.2	16.5	30.9	12.2
Belgium	529	*6*	0.8	27.0	6.6	38.3	37.9	4.6
Switzerland	527	*7*	78.2	53.0	41.1	1.4	23.1	5.0
Australia	524	*8*	40.8	7.0	16.0	28.3	27.5	43.5
New Zealand	523	*9*	42.5	12.5	14.0	14.2	22.5	32.9
Czech Republic	516	*10*	22.8	51.3	2.4	12.6	12.9	5.5
Iceland	515	*11*	93.1	0.0	3.0	1.8	1.1	0.2
Denmark	514	*12*	59.3	4.0	4.6	17.0	14.0	19.6
Sweden	509	*14*	62.3	4.5	2.5	8.5	11.7	9.5
UK	506	*15*	62.2	9.2	7.1	17.2	6.8	40.3
Austria	506	*16*	26.2	52.5	5.0	9.8	39.3	13.7
Germany	503	*17*	54.9	23.9	49.6	6.0	19.6	11.6
Ireland	503	*18*	41.7	4.1	7.0	25.7	10.0	38.0
Slovak Republic	498	*19*	11.1	49.7	2.6	6.8	23.7	3.3
Norway	495	*20*	73.8	0.0	2.6	0.9	2.2	2.5
Luxembourg	493	*21*	15.0	49.6	13.1	14.2	11.9	24.0
Poland	490	*22*	82.2	26.6	12.7	12.4	12.4	8.2
Hungary	490	*23*	8.8	75.0	8.0	23.3	64.2	15.5
USA	483	*24*	79.3	15.9	16.6	8.8	23.7	7.9
Spain	485	*25*	71.5	2.6	1.7	11.2	12.4	37.6
Portugal	466	*26*	58.7	0.7	1.3	8.9	41.7	29.9
Italy	466	*27*	7.2	7.7	9.1	6.9	51.2	13.1
Greece	445	*28*	64.6	1.3	1.7	2.5	12.9	15.1
Turkey	423	*29*	27.3	11.8	3.4	1.1	7.2	3.5
Mexico	385	*30*	14.5	36.0	12.1	10.3	13.4	12.5

Sources: PISA (2004) Tables 2.5c and 5.8

1. France did not return this information.

5.2 There is wide variation. Ninety-three per cent of the headteachers in Iceland said that school admissions were based mainly on proximity compared to only 0.8 per cent in Belgium. Among the criteria which could cut across attending the local school were academic selection, choice of faith school, or going to the same school as siblings. Again there was wide variation. Eighty-eight per cent of headteachers in Japan said the pupils were mainly admitted on the basis of academic record; 84.2 per cent of the heads in the Netherlands said the main criterion was the recommendation of feeder schools; and 64.2 per cent of the heads in Italy said it was the special needs/wishes of the pupils that was decisive. Admission by faith was highest in Belgium at 37.9 per cent and by siblings in Australia (43.5 percent) and the UK (40.3 per cent).

5.3 However, although interrelated among themselves, none of the main methods of school admission was significantly associated with a country's performance in the PISA 2003 maths tests. Whether school admission was by proximity, by previous performance, or by one of the other methods did not make a sufficient difference to show up in the national results. There was a consistency to the headteachers' responses which suggests that they were authentic. 'Student's academic records', 'recommendation from feeder schools' and 'students' needs or desires for a special programme' were significantly related as were 'parents' endorsement of the instructional or religious philosophy of the school' and 'attendance of other family members at the school'. And as to be expected academic selection, religion and special needs were inversely associated with going to the nearest school. But none correlated with maths performance. The highest relationship was with admission on the basis of recommendation by feeder schools where a Pearson coefficient of +0.309 was obtained, but for 29 countries (France did not supply this information) significance barely reaches the 10 per cent level. It could be, therefore, that there was, in fact, little difference, or that the PISA maths tests did not capture the differences which do exist.

5.4 The method of school admission was found to be related to the extent to which the variance in student performance in a country is within schools or between schools. Countries with the highest between-school variance, like Hungary, the Netherlands and Japan tend to admit on academic ability and those with the least, like Iceland, Finland and Norway tend to admit on proximity. The split for the UK (essentially England) is something of a surprise with over three-quarters of the variance (77.7 per cent) being found within schools. This challenges the common impression of a very diverse and uneven system with students being privileged. As we shall be seeing, there is a marked gap in the performance of the independent and state schools so presumably the overall figure comes out as it does because in the UK the independent (and also the grammar schools) comprise only a small part of the whole.

5.5 The OECD has been inclined to interpret the results summarised in Chart 5.2 as showing the superiority of comprehensive systems over selective systems, but the data do not really bear this out. It is true that Finland with the second lowest between-school variance obtained the highest mean score in PISA 2003 maths, but it was closely followed by the Netherlands with the second highest between-school variance. The countries with low between-school variance and within-school

variance are scattered through the maths ranking, and not surprisingly there is no statistical association.

Chart 5.2: Maths Performance and School Diversity

OECD Country[1]	Maths in PISA 2003		Per Cent Variation in Student Performance	
	Score	Rank	Between Schools	Within Schools
Finland	544	*1*	4.8	95.2
Korea	542	*2*	42.0	58.0
Netherlands	538	*3*	58.0	42.0
Japan	534	*4*	53.0	47.0
Canada	532	*5*	17.3	82.7
Belgium	529	*6*	46.0	54.0
Switzerland	527	*7*	34.2	65.8
Australia	524	*8*	21.1	78.9
New Zealand	523	*9*	18.1	71.9
Czech Republic	516	*10*	47.8	52.2
Iceland	515	*11*	3.8	96.2
Denmark	514	*12*	13.4	86.8
Sweden	509	*14*	10.5	89.5
UK	506	*15*	22.3	77.7
Austria	506	*16*	51.7	48.3
Germany	503	*17*	15.9	84.1
Ireland	503	*18*	19.7	80.3
Slovak Republic	498	*19*	41.7	58.3
Norway	495	*20*	6.6	93.4
Luxembourg	493	*21*	31.6	68.4
Poland	490	*22*	12.6	87.4
Hungary	490	*23*	58.3	41.7
USA	483	*24*	25.7	74.3
Spain	485	*25*	19.7	80.3
Portugal	466	*26*	33.6	66.4
Italy	466	*27*	52.2	47.8
Greece	445	*28*	36.3	63.7
Turkey	423	*29*	54.9	45.1
Mexico	385	*30*	39.4	60.6

Sources: PISA (2004) Tables 2.5c and 4.1a.

1. France did not provide the information for these calculations to be made.

5.6 While school admissions, and within- and between-school variance in student performance, were not significantly linked to the PISA 2003 maths results, one aspect of school organisation does stand out. OECD/PISA provides data on maths performance by school type. It categorises schools as independent; government-funded private; and government-run. In Chart 5.3 we compare the performance of the independent and government schools showing also the proportions of pupils within them (the percentages do not necessarily sum to 100 since there is also the

category of government-funded private schools to which we turn in Chart 5.4). Although there are some notable exceptions (which need to be looked at in more detail) the independent schools tended to do better than the government schools. In some cases the differences were considerable. The UK's independent schools, which obtained the highest overall score for a school group on reading in PISA 2000, were second only to Korea in maths in PISA 2003, 86 points ahead of the state schools.

Chart 5.3: Comparison Independent and Government Run Schools

Country	Independent Schools		Government Schools		Diff Ind Govt Schools	Difference[1] Adjusting for Social Background
	Per Cent Students	Score	Per Cent Students	Score		
Korea	21.7	593	42.3	527	66	14
UK	5.3	589	93.8	503	86	50
New Zealand	4.6	579	95.4	522	57	25
Macao-China	45.8	529	5.0	483	46	40
Spain	7.7	520	64.2	472	48	19
Japan	26.4	513	73.0	544	-31	-40
USA	5.7	507	94.3	483	24	4
Uruguay	14.1	501	85.9	409	92	54
Switzerland	3.8	497	95.3	528	-31	-39
Brazil	12.6	454	87.4	342	112	74
Italy	3.5	452	96.1	468	-16	-32
Mexico	13.2	430	86.7	375	55	26
Thailand	6.0	428	88.0	416	12	-3
Indonesia	44.5	345	51.4	373	-28	-27

Sources: PISA (2004), Table 5.19, pages 436-37.

1. Comparison adjusted for economic, social and cultural status (last columnl) includes as independent both private and government dependent private schools. Korea (36.0%), Macao China (49.3%), Spain (28.1%), Thailand (6.0%) and Indonesia (4.1%) have the percentages of government-dependent private schools shown in brackets in addition to the percentages of independent schools shown in the chart.

5.7 There were also large differences in favour of independent schools in Korea, New Zealand, Mexico and Spain. Some of the differences in non-OECD countries were even greater. In Brazil the independent schools were 112 points ahead, in Uruguay 92, and in Macao-China 46. There would thus seem to be an interesting school effect to be explained. A first port of call would be the background the intakes. OECD/PISA has compiled an Index of Economic, Social and Cultural Status (ESCS) from such variables as parental occupational status, parents' education and home background including the number of books in the home. The final column in Chart 5.3 takes account of these differences in the intakes. Pupils in independent schools in most countries do tend to come from higher socio-economic backgrounds, though not invariably so. Korea and Luxembourg are interesting exceptions. When the maths scores are adjusted to take account of ESCS differences the gap between private and state schools is reduced, but does not disappear. In the UK, for example, the 86 point raw difference reduces, but only to 50 points.

5.8 There do, therefore, seem to be effects beyond social background. This is borne out by the comparisons between the government-dependent private schools and the government-run schools shown in Chart 5.4. The private schools tended to be ahead, notably in Germany, Canada, Ireland, the Slovak Republic and the Netherlands, even after adjustments for background. Macao-China, Spain, Korea and Thailand have both independent and government-dependent private schools and there was a progression in the PISA 2003 maths scores from the fully independent doing best to the government schools doing worse. But that is not always the case. In Indonesia, for example, the government schools do best, particularly in relation to the government-funded private schools.

Chart 5.4: Comparison of Govt Dependent Private Schools and Govt Schools

Country	Govt-Dependent Private Schools		Government Schools		Diff Govt Dependent and Govt Schools	Difference Adjusting[1] for Social Background
	Per Cent Students	Score	Per Cent Students	Score		
Canada	3.8	573	94.2	529	44	26
Germany	7.5	566	92.2	497	69	30
Netherlands	76.7	541	23.3	516	25	10
Finland	6.7	539	93.3	545	-6	-13
Korea	36.0	532	42.3	527	5	14
Macao-China	49.3	528	5.0	483	45	40
Slovak Republic	12.6	523	87.4	495	28	16
Austria	6.7	518	92.0	504	14	5
Hong Kong	6.5	518	93.1	552	-34	-41
Sweden	4.3	516	95.7	509	7	-7
Ireland	57.6	516	41.6	486	30	16
Denmark	21.7	511	77.8	515	-4	-5
Spain	28.1	505	64.2	472	33	19
Czech Republic	5.8	505	93.3	517	-12	-14
Hungary	9.8	504	88.9	489	15	5
Thailand	6.0	419	88.0	416	3	-3
Indonesia	4.1	326	51.4	373	-47	-27
Luxembourg	14.1	463	85.9	498	-35	-28
Portugal	4.2	459	93.7	465	-6	11

Sources: PISA (2004), Table 5.19, pages 436 –37.

1. Comparison adjusted for economic, social and cultural status (last columnl) includes as independent both private and government dependent private schools. Korea (21.7%), Macao China (45.8%), Spain (7.7%), Thailand (6.0%) and Indonesia (44.5%) have the percentages of government-dependent private schools shown in brackets in addition to the percentages of independent schools shown in the chart.

Comment

5.9 In the UK, independent schools are only a small part of the overall system and they have history, tradition, income and academic selection on their side, as well as social background. But private schools also do better in some countries where they are in the majority, as in the Netherlands, Ireland and Macao-China, suggesting that what it is they have will in the right circumstance scale up. What can it be? One possible candidate is that independent schools by their very nature are less subject to

government interference than state schools. Smithers and Robinson (2007) suggested that an important element in the success of UK independent schools is that they have been largely free of the welter of government initiatives that have befallen their state counterparts. The case of the Netherlands is especially interesting because three-quarters of its pupils go to government-funded private schools and it is the third highest performer overall. An unpalatable lesson for governments from the OECD/PISA studies may be that schools are better off without their close attention. School autonomy may be the key to the independent schools' success.

6. Societal and Personal Factors

6.1 The OECD/PISA studies are intended to shed light on the performance of education systems with a view to effecting improvements. But at their heart is a logical inconsistency. The designers have deliberately sought to devise tests that reach beyond the school curriculum to the capacity to apply knowledge. They distinguish this approach by referring to 'mathematical literacy', 'scientific literacy' and 'reading literacy' rather than just maths, science and reading. While seeking to capture the use to which knowledge is put in everyday situations is commendable, it does increase the likelihood that any differences found between countries may derive aspects of a country's life other than its school system. Performance could, for example, be related to, among other things, the wealth of the country, commitment to education, personal characteristics, or the extent and nature of new arrivals.

Wealth

6.2 Chart 6.1 shows the 30 OECD countries ranked on earning power, spend on education, staying on rates in education and performance in PISA maths. The data are all referenced to 2003. There is a broad correlation with the poorer countries doing less well in PISA. But wealth was not guarantee of success. The three countries with the highest GDP per capita - Luxembourg, USA and Norway - come two-thirds of the way down on PISA scores. And, conversely, Finland Japan and Korea, occupying three of the four top places on PISA maths fall in the lower half on income. Where wealth does come into play is through countries which can afford it spending more on education. The strongest association in Chart 6.1 is between GDP per capita and spend on education. The Pearson coefficient comes out at +0.894 and even some of the discrepancies may be more apparent than real. Both Ireland and Netherlands, for example, appear to spend less on education than their income would permit, but as we have seen both have distinctive secondary education systems with substantial independent sectors. There is an association between public spend on education and PISA, but it is relatively weak at +0.367, significant at the 5 per cent level.

Commitment to Education

6.3 Much more important than a country's wealth in predicting the PISA 2003 results was participation in upper secondary education, which can be taken as an index of individual commitment to education when participation is voluntary and a country's when it is compulsory. Chart 6.1 shows that the proportion of 25-34 year-olds participating in upper secondary education varied between 97 per cent in Korea and 25 per cent in Mexico. The United Kingdom, with 70 per cent, comes well down this list in 23rd place. Completing upper secondary education is highly correlated with PISA maths performance at age 15, yielding a Pearson coefficient of +0.796. This is exemplified by Korea which is 23rd on wealth, but top for staying on and second in PISA maths. Similarly, Japan, 17th on wealth, is equal third for participation and fourth in PISA. At the other end of the scale Luxembourg emerges top on wealth, but comes only 21st for completing upper secondary education and is 21st in PISA. The correlation is not perfect. Norway has high participation but is low on PISA and the Netherlands and Australia are low on participation and high on PISA, but overall there is a strong association.

Chart 6.1: GDP, Education Spend, Participation and PISA Performance

Country	GDP Per Capita[1]		Cumulative Spend on School Educ[2]		%25-34 Completed Upper Sec Education[3]		PISA 2003 Maths Score[4]	
	$ (PPP)	Rank	$(PPP)K	Rank	%	Rank	Score	Rank
Luxembourg	55,571	*1*	188.6	*1*	74	*21*	493	*21*
USA	37,510	*2*	107.6	*4*	87	*10=*	483	*25*
Norway	37,237	*3*	120.6	*3*	96	*2*	495	*20*
Ireland	34,171	*4*	66.9	*21*	79	*19*	503	*18*
Switzerland	33,217	*5*	130.0	*2*	89	*8=*	527	*7*
Netherlands	31,792	*6*	69.0	*18*	80	*16=*	538	*3*
Australia	31,100	*7*	84.9	*11*	77	*20*	524	*8*
Austria	30,797	*8*	100.2	*8*	87	*10=*	506	*16*
Iceland	30,774	*9*	102.5	*6*	68	*24*	515	*11*
Denmark	30,677	*10*	103.9	*5*	86	*12*	514	*12*
Canada	30,403	*11*	77.8	*16*	91	*6=*	532	*5*
Belgium	30,089	*12*	83.3	*12*	80	*16*	529	*6*
United Kingdom	29,609	*13*	82.5	*14*	70	*23*	507	*15*
Sweden	29,522	*14*	89.6	*9*	91	*6=*	509	*14*
France	28,373	*15*	85.1	*10*	80	*16=*	511	*13*
Finland	28,334	*16*	77.7	*17*	89	*8=*	544	*1*
Japan	28,071	*17*	81.7	*15*	94	*3=*	534	*4*
Germany	27,619	*18*	83.1	*13*	85	*13=*	503	*17*
Italy	26,561	*19*	100.4	*7*	64	*25*	466	*27*
Spain	24,812	*20*	67.4	*20*	61	*26*	485	*24*
New Zealand	23,551	*21*	68.4	*19*	85	*13=*	523	*9*
Greece	20,479	*22*	55.0	*24*	73	*22*	445	*28*
Korea	19,317	*23*	63.2	*23*	97	*1*	542	*2*
Portugal	17,617	*24*	63.6	*22*	40	*28*	466	*26*
Czech Republic	17,284	*25*	44.1	*26*	94	*3=*	516	*10*
Hungary	15,112	*26*	44.7	*25*	84	*15*	490	*23*
Slovak Republic	13,114	*27*	29.6	*28*	94	*3=*	498	*19*
Poland	11,583	*28*	38.0	*27*	60	*27*	490	*22*
Mexico	9,585	*29*	22.8	*29*	25	*30*	385	*30*
Turkey	6,762	*30*	11.2	*30*	33	*29*	423	*29*

1. Adapted from OECD (2006) *Education at a Glance,* OECD Indicators. Paris: OECD, Table X2.1, page 430.

2. Adapted from OECD (2006) *Education at a Glance,* OECD Indicators. Paris: OECD, Table B1.3a, page 190.

3. Adapted from OECD (2006) *Education at a Glance,* OECD Indicators. Paris: OECD, Table A1.2a, page 38.

4. Adapted from OECD (2004) *Learning for Tomorrow's World,* First Results from PISA 2003. Paris: OECD, Tables 2.5c, page 356.

Personal Characteristics

National

6.4 It is striking how well the south-east Asian countries do in the international comparisons, particularly in mathematics. In Chart 4.3, page 17, in which the performance of 20 countries in TIMSS 1995, 1999 and 2003 is compared, Singapore, Korea, Hong Kong and Japan consistently occupy the top four places.

Chart 2.1, page 5, shows that Korea came second overall and Japan was in the top five out of 33 countries and regions in both PISA 2000 and 2003. When it participated in PISA 2003 Hong Kong-China achieved the highest maths score. Neither is this impressive performance confined to country comparisons. The DfES (2006) compared performance by ethnicity in England's 2006 national tests and examinations. It concluded that 'Chinese, pupils of Mixed White and Asian heritage, Irish and Indian pupils consistently achieve above the national average across Key Stage 1, Key Stage 2 and Key Stage 4'.

6.5 We thus have to ask what is about the south east Asian countries and children that enables them to do so well? It is likely to be some combination of natural talent and cultural attitudes to education, and that is about as far as we can take it here. But it does underline the difficulty in interpreting the results from international comparisons. A country's score may have less to do with its education system than the people whom it serves and their attitudes to the provision. It makes little sense, therefore, for England, for example, to attempt to import Korean methods in the hope that they would improve maths performance.

Gender

6.6 That is not to say the nature of the education system is without impact. One way of determining whether a difference is genetic or otherwise deep-seated is the consistency across countries. One of the most striking sets of data emerging from PISA 2003 is the gender differences. These are set out in Chart 6.2. In all 30 countries the female students are ahead in reading, on average, by at least 20 points, rising to 58 points in the case of Iceland. Conversely, in maths male students obtain the better average score in all countries bar one – Iceland (where girls score higher by 15 points) – though the range is less – from 5 points in the Netherlands and Australia to 23 points in Korea. The picture in science is more mixed with male students doing better in 21 countries with the advantage to males being greatest in Korea, 18 points, and to girls in Iceland, by 10 points.

6.7 The same countries keeping recurring at the top and bottom of the lists and this raises the possibility that out of the 30 OECD countries some are more female-friendly or have more female-friendly education systems, and conversely others are more male-friendly. Iceland is the prime example of a country that seems to get the best out of girls or, alternatively, let down boys (in maths girls came seventh and boys equal 17th). Of all the countries, it had the highest advantage to girls in reading, was the only one where they were ahead in maths, and had the biggest difference in favour of girls in science. In stark contrast, in Korea, the reading gap for boys was one of the smallest and the advantages to boys in maths and science were the highest (in maths, the Korean boys came top and the girls were fifth). Norway and Finland are similar to Iceland, but another Scandinavian country, Denmark, is closer to Korea. Other countries similar to Korea in this respect are New Zealand and Canada, while Austria and Germany are more like Iceland. What makes a country or educational system more female-friendly or male-friendly would be an interesting topic for further inquiry. But these results do show that the scores obtained are not just a product of genetic differences. The consistency with which performance in maths, science and reading can be shifted towards girls or boys points to effects of the culture and/or the education system in a country.

Chart 6.2: Gender Difference in Performance in PISA 2003

Country	Maths		Science		Reading		Difference M-F		
	Male	Female	Male	Female	Male	Female	Maths	Science	Reading
Finland	548	541	545	551	521	565	7	-6	-44
Korea	552	528	546	527	525	547	23	18	-21
Netherlands	540	535	527	522	503	524	5	5	-21
Japan	539	530	550	546	487	509	9	4	-22
Canada	541	530	527	516	514	546	11	11	-32
Belgium	533	525	509	509	489	526	8	0	-37
Switzerland	535	518	518	508	482	517	17	10	-35
Australia	527	522	525	525	506	545	5	0	-39
New Zealand	531	516	529	513	508	535	14	16	-28
Czech Republic	524	509	526	520	473	504	15	6	-31
Iceland	508	523	490	500	464	522	-15	-10	-58
Denmark	523	506	484	467	479	505	17	17	-25
France	515	507	511	511	476	514	9	0	-38
Sweden	512	506	509	504	496	533	6	5	-37
England	511	504	521	517	491	511	7	4	-20
Austria	509	502	490	492	467	514	7	-3	-47
Germany	508	499	506	500	471	513	9	6	-42
Ireland	510	495	506	504	501	530	15	2	-29
Slovak Republic	507	489	502	487	453	486	19	15	-33
Norway	498	492	485	483	475	525	6	2	-49
Luxembourg	502	485	489	477	463	496	17	13	-33
Poland	493	487	501	494	477	516	6	7	-40
Hungary	494	486	503	504	467	498	8	-1	-31
Spain	490	481	489	485	461	500	9	4	-39
USA	486	480	494	489	479	511	6	5	-32
Portugal	472	460	471	465	459	495	12	6	-36
Italy	475	457	490	484	455	495	18	6	-39
Greece	455	436	487	475	453	490	19	12	-37
Turkey	430	415	434	434	426	459	15	0	-33
Mexico	391	380	410	400	389	410	11	9	-21

Sources: PISA (2004) Table 2.5c page 356, Table 6.6 page 448, Table 6.2 page 444 plus PISA (2006) spreadsheets.

Immigration

6.8 A country's score on the tests in OECD/PISA represents not only the children who have been born there, but also those who have arrived and were in the educational system at age 15. Chart 6.3 examines the effect on overall performance of the new arrivals. It compares the PISA maths scores overall with the scores of the native born children only. It further shows the percentage of native born, the difference between the scores of the native born and arrivals, both as raw scores and taking into account socio-economic status.

Chart 6.3: Maths Scores in PISA 2003 Contrasting Native and Immigrant Children

OECD Country	Maths PISA 2003		Maths Scores of Native Born[1]		%Native Born	Diff Native and Non Native	Diff taking into account ESCS
	Score	Rank	Score	Rank			
Finland	544	*1*	546	*2*	98.1	-	-
Korea	542	*2*	543	*4*	100.0	-	-
Netherlands	538	*3*	551	*1*	89.0	66	37
Japan	534	*4*	543	*5*	99.9	-	-
Canada	532	*5*	537	*7*	79.9	1	1
Belgium	529	*6*	545	*3*	88.2	100	60
Switzerland	527	*7*	543	*6*	80.0	76	49
Australia	524	*8*	527	*9*	77.3	3	-1
New Zealand	523	*9*	528	*8*	80.2	14	11
Czech Republic	516	*10*	523	*11*	98.7	-	-
Iceland	515	*11*	517	*14*	99.0	-	-
Denmark	514	*12*	520	*12*	93.5	-	-
France	511	*13*	520	*13*	85.7	54	21
Sweden	509	*14*	517	*15*	88.5	64	41
UK	506	*15*	510	*17*	92.0	16	5
Austria	506	*16*	515	*16*	86.7	61	36
Germany	503	*17*	525	*10*	84.6	81	35
Ireland	503	*18*	503	*19*	96.5	4	18
Slovak Republic	498	*19*	499	*20*	99.1	-	-
Norway	495	*20*	499	*21*	94.4	52	34
Luxembourg	493	*21*	507	*18*	66.7	38	13
Poland	490	*22*	491	*22*	100.0	-	-
Hungary	490	*23*	491	*23*	97.7	-	-
USA	483	*24*	490	*24*	85.6	28	4
Spain	485	*25*	487	*25*	96.6	45	36
Portugal	466	*26*	470	*26*	95.0	61	62
Italy	466	*27*	468	*27*	97.9	-	-
Greece	445	*28*	449	*28*	92.6	43	27
Turkey	423	*29*	425	*29*	99.0	-	-
Mexico	385	*30*	392	*30*	97.7	-	-

Sources: PISA (2004) Tables 2.5c, 4.2f and 4.2h.

1.Native students are defined as born in the country of assessment with one at least of their parents born in that country.

2.ESCS is an index derived from occupational status of parents, educational background and access to home educational and cultural resources including books in the home.

6.9 Chart 6.3 shows that immigration can have a considerable bearing on a country's PISA result. Germany would have come tenth in PISA 2003 maths if only native born students had been included as against seventeenth in the overall standing, and the Netherlands not Finland would have come top. The extent of the difference in maths performance between the native-born and non-native students ranged from 100 points in Belgium, 81 in Germany and 76 in Switzerland to just few points in

Canada and Australia. The latter are two of the countries with the highest proportions of immigrants, so the effect on PISA score depends not only on the extent of immigration but also on who it involves.

6.10 OECD (2006c), itself, has conducted an extensive analysis of the PISA results in relation to immigration and despite a determinedly positive tone exemplified in the title, *Where Immigrant Students Succeed*, comes to some blunt conclusions. It finds that 'despite strong learning dispositions immigrant students often perform at levels significantly lower than their native peers…In the majority of countries at least 25% of immigrant students could face considerable challenges in their future professional and personal lives as they do not demonstrate basic mathematics skills in the PISA 2003 assessment'. Its analysis reveals that background characteristics, including whether the students speak the language of instruction at home, only partly explain the differences in performance. It suggests there are lessons to be learned from those countries where there are relatively small performance differences between immigrant and native students, or the performance gaps for second-generation students are significantly less than for the first.

Comment

6.11 The OECD/PISA studies are designed to shed light on educational systems. But, especially as they are intended to look beyond the curriculum, they could owe more to other aspects of a nation's life. The association between wealth and PISA performance was found to be comparatively weak, but there was a relationship with commitment to education. Natural talent also played a part. There was a gender effect, but also evidence that educational systems could be more 'girl friendly' or more 'boy friendly'. The nature and extent of immigration was also found to have a major impact on the results obtained in PISA. These societal and personal factors underline the importance of not assuming that because a study has been set up to study the performance of educational systems the results are necessarily best interpreted in these terms.

7. Personal Development

7.1 Even if the OECD/PISA studies were based on perfect measures and the results could be interpreted in terms of the performance of educational systems, it could be argued that the picture they provide is incomplete. Their focus is academic and they pay little attention to the broader aspects of personal development. The same is true of the national tests and examinations by which the British government seeks to monitor its education system. The broader aspects of education are less easily defined and measured, but the United Nations Children's Fund (Unicef, 2007) has attempted to provide a comprehensive assessment of the lives of young people in 21 prosperous nations by drawing together data from a wide variety of sources. Chart 7.1 summarises this data, including the somewhat disturbing finding that the UK comes at the very bottom.

Chart 7.1: Well-Being

	Overall Rank	Material Well-Being	Health & Safety	Educational Well-being	Family & Peer Relationships	Behviours & Risks	Subjective Well-Being
Netherlands	*1*	*10*	*2*	*6*	*3*	*3*	*1*
Sweden	*2*	*1*	*1*	*5*	*15*	*1*	*7*
Denmark	*3*	*4*	*4*	*8*	*9*	*6*	*12*
Finland	*4*	*3*	*3*	*4*	*17*	*7*	*11*
Spain	*5*	*12*	*6*	*15*	*8*	*5*	*2*
Switzerland	*6*	*5*	*9*	*14*	*4*	*12*	*6*
Norway	*7*	*2*	*8*	*11*	*10*	*13*	*8*
Italy	*8*	*14*	*5*	*20*	*1*	*10*	*10*
Ireland	*9*	*19*	*19*	*7*	*7*	*4*	*5*
Belgium	*10*	*7*	*16*	*1*	*5*	*19*	*16*
Germany	*11*	*13*	*11*	*10*	*13*	*11*	*9*
Canada	*12*	*6*	*13*	*2*	*18*	*17*	*15*
Greece	*13*	*15*	*18*	*16*	*11*	*8*	*3*
Poland	*14*	*21*	*15*	*3*	*14*	*2*	*19*
Czech Republic	*15*	*11*	*10*	*9*	*19*	*9*	*17*
France	*16*	*9*	*7*	*18*	*12*	*14*	*18*
Portugal	*17*	*16*	*14*	*21*	*2*	*15*	*14*
Austria	*18*	*8*	*20*	*19*	*16*	*16*	*4*
Hungary	*19*	*20*	*17*	*13*	*6*	*18*	*13*
USA	*20*	*17*	*21*	*12*	*20*	*20*	-
UK	*21*	*18*	*12*	*17*	*21*	*21*	*20*

Source: Unicef (2007)

7.2 The reasons for the UK's low placing are drawn out in Chart 7.2. It falls in the bottom quarter on 25 of the 35 dimensions considered. The picture is particularly gloomy when it comes to behaviour where the UK compares poorly on 10 of the 12 indicators including drunkenness, taking drugs, smoking and teenage pregnancy. Among the few dimensions on which the UK comes out relatively well are several to do with education. In PISA 2003, which is incorporated into educational well-being,

the UK ranks 9th out of 25, but its overall rank is lowered by the low staying on rate and low aspirations. It also emerges respectably on having 'six educational possessions' (7th out of 24). The UK's best position is 2nd out of 24 on 'deaths from injuries or accidents during the first 12 months of life' and it is also placed in the top half on 'physical activity' (6th out of 21), 'reporting low family affluence' (8th out of 20), and 'talking with parents' (12th out of 25). But the overall impression is of a country in which the teenagers do not have a secure moral compass.

Chart 7.2: UK Ranks

Dimension	UK Rank	Countries Compared
Material Well Being		
Household income less than 50 per cent of the median	23	24
Parents Unemployed	20	24
Health and Safety		
Deaths before 12 months	20	25
Low birth weight	21	25
Vaccination	21	25
Educational Well Being		
15-19 year olds not in education, training or employment	19	23
Expecting work with low skills	21	25
Relationships		
Living in single parent families	20	21
Living with step families	20	21
Eating meals with parents	22	25
Peers not 'kind and helpful'	21	21
Behaviours and Risks		
Eat breakfast every day	16	21
Eat fruit each day	19	21
Overweight	17	21
Smoke	17	21
Drunk two or more times	21	21
Used cannabis	18	20
Had sexual intercourse	17	17
Used condom during intercourse	12	15
Teenage pregnancy	22	24
Involved in fighting	17	21
Bullied	16	21
Subjective Well-being		
Rating of health	21	21
Liking school	16	21
Satisfaction	16	21

Source: Unicef (2007)

7.3 Chart 7.1 is chiefly notable for the UK and the USA being anchored at the bottom. Across the other countries there is wide variation with only three of the dimensions being correlated – material well-being and health and safety (+0.594, P<0.01), health and safety and behaviours and risks (+0. 513, P< 0.05), and behaviours and risks and subjective well-being (+0.441, P=0.05). Sweden tops three categories – material well-being, health and safety and behaviours and risks – Belgium is first on education, Italy on relationships and the Netherlands on subjective well-being. Unicef (2007) does not discuss these findings and is content to offer them as a first attempt at a comprehensive assessment, warning against 'leaning too hard on limited data'.

7.4 There are reasons for being wary of some of the measurements. An important element in material well-being, for example, is 'income poverty' but this is a relative measure defined as below 50 per cent of the national median. It does not mean that the UK's children, or those in the United States who came below the UK, are poor in absolute terms, only that the incomes in those countries are more spread out than in some others. Neither since this is the first attempt at producing such a ranking can we know whether significant changes have taken place overall in the Blair years. But Unicef's 2007 Report Card is their seventh and they have considered particular dimensions in their previous reports. In the sixth they focused on child poverty. Consistent with the UK's low standing in the 2007 report it is placed 20th of 25 countries. In comparison with ten years previously, covered in the first report, the UK shows the biggest increase in the measure of child poverty. Other report cards have looked at child deaths by injury, teenage births, educational disadvantage, and child maltreatment. In the study of educational disadvantage based on TIMSS 1995 and 1999 and PISA 2000 the UK fared reasonably well coming 7th of 25 nations.

Comment

7.5 Unicef's data pose more questions than they answer and they do not tell us what aspects of particular cultures and communities are associated with the rankings that emerge. Whatever the limitations of the data, however, the poor behaviour of children in the UK compared to other countries does stand out and has to be a major concern. It can probably be best interpreted as a breakdown in values. Kitwood (1980) studied how value systems develop and concluded that those of children and adolescents are only loosely attached to them. The values are less theirs than a reflection of those of their parents and their schools. This suggests that while there is no evidence that schools are responsible for the current state of children's values, society's best hope of intervening is through them. The thrust of Blair's policies for education has been to raise academic performance and it is possible that value development has been somewhat neglected. How best to enable children to develop soundly-based value systems should be an important area for further inquiry. It may be that values can be taught directly through the curriculum, but more likely they will be acquired through the ethos of the school. This makes it important that schooling should not be exclusively geared to test and exam scores, but also embrace personal development.

8. Is England's Education Improving?

8.1 The crunch questions towards which we have been working are: has education in England got better under Tony Blair and, if so, is it better compared to other countries? The short answer is that it much harder to tell than it should be. Test and examination scores have certainly gone up and, in so far as they capture the underlying quality of education, then it has been getting better. But are they an accurate representation? There have to be two main areas of doubt. The first is prompted by the failure in many instances of independent tests, including international comparisons, to bear out the improvements claimed on the basis of the national tests and examinations. This raises the possibility that extraneous factors such as teaching to the test and juggling with exam entries have pushed up the scores without improving education itself.

8.2 The second area of doubt has to do with the broader purposes of education. Let us assume for a moment that the tests and examinations are perfect measures and there have been the improvements that the scores appear to show. Is this enough? What of the less measurable aspects of education? One of the problems here is that while we are instinctively aware that education is not just about academic success we do not have an accepted language in which to address the issue. Employers talk about 'soft skills' (eg Chartered Institute of Personnel Development, 2006); Unicef (2007) homes in on 'well-being'; Gardner (2006) speaks of seven forms of intelligence, only two of which are traditionally taught in schools; Seldon (Wilby 2007) encapsulate it as 'happiness'; more generally it is usual to refer to a well-rounded education. Let us examine both these areas of doubt – the meaning of the scores and the broader purposes of education - in more detail and see if there are any policy pointers to be drawn from them.

Test and Examination Scores

8.3 Scores in the national tests and examinations have risen impressively in the past decade. As we saw in Chapter 3, performance in both primary and secondary schools appears to have gone up sharply. Chart 8.1 summarises the main changes.

Chart 8.1: Test and Exam Scores

Benchmark	Per Cent	
	1997	2006
KS 2 English Level 4	63 0	79.0
KS 2 Maths Level 4	62.0	76.0
KS 2 Science Level 4	69.0	87.0
5 A*-C GCSE	45.1	59.0
A-level Passes	87.8	96.6
A-level A Grades	15.7	24.1

Sources: DfES (2006b, 2007a); Annual publications of InterBoard Statistics, compiled by CEER.

8.4 The government would claim these as real improvements and attribute them to its policies and strategies, the extra money that it had found for education and additional teachers being recruited. But the difficulty in accepting these data at face value, as

we found in Chapters 2 and 4, is that independent tests do not show similar increases. Chart 8.2 summarises the results from PISA and TIMSS for the secondary-age pupils from 1999-2003. No improvement is evident. If anything England's performance seems to have slipped back relative to the comparison countries.

Chart 8.2: 14/15 Year-Olds in International Context

Study	1999/2000		2003	
	Score	Rank	Score	Rank
PISA		(33)[1]		(33)[1]
Reading	523	*8*	506	*11*
Maths	529	*=9*	507	*15*
Science	533	*4*	519	*9*
TIMSS		(33)[2]		(33)[2]
Maths	496	*17*	498	*16*
Science	538	*8*	544	*6*

Sources: PISA (2004, 2001); Mullis *et al* (2000, 2004a,b); Martin *et al* (2000)
1.33 countries and regions as listed in Chart 2.2
2.Overlapping but different 33 countries from PISA including South Africa, Philippines, Morocco, Tunisia, Chile, Indonesia, Cyprus, Iran etc

8.5 But of course there have to be doubts too about what these numbers from the international studies actually mean. The British government has refused to acknowledge the PISA and TIMSS (secondary age) 2003 results on the grounds that England's response rates were too low. That is in spite of welcoming the PISA 2000, the PIRLS 2001 and the TIMSS (primary-age) 2003 results when the response rates were no better (see Chart 4.1, page 16) but were more favourable to its case. Data for the primary age range are summarised in Chart 8.3.

Chart 8.3: 10 Year-Olds in International Context

Study	1995		2001/2003	
	Score	Rank	Score	Rank
PIRLS				(34)
Reading			553	*3*
TIMSS		(15)		(15)
Maths	513	*11*	531	*6*
Science	551	*5*	540	*4*

Sources: Mullis *el al* (2003); Mullis *et al* (1998); Martin et al (1998); Mullis *et al* (2004a,b).

8.6 There are other difficulties in interpreting these findings. The first is that they are relative scores affected by the performance of other countries. It is perfectly possible, therefore, that England's pupils did better, but appeared to decline because other countries had improved even more. It is also not clear to what extent a nation's scores on these tests reflects its educational system. This is assumed to be the case and the results are interpreted accordingly. But, as we showed in Chapter 6, the scores are affected by both societal factors (for example, extent and nature of immigration) and personal factors (for example, talent). Interpretation of PISA scores in relation to schooling is especially problematic because the approach is to

deliberately eschew the curriculum in favour of testing ability-to-apply, which it calls 'literacy'.

8.7 The doubts from the international comparisons are reinforced by the findings of Peter Tymms and his team at Durham University, which we discussed in detail in Chapter 2. The backbone of their work is providing a testing service to schools. As a result they have administered the same tests each year in quite a number of the schools in England. But they do not find the rapid rises recorded in the Key Stage 2 tests. There has been some improvement in reading and rather more in maths, but nothing like the sharp increases found in the national tests. These findings are borne out by other independent studies, and were supported by the Statistics Commission. There are several possible reasons for the discrepancy. The national tests are specific to the literacy and numeracy strategies and a lot hangs on the results. They are not taken just as a measure of the progress of pupils, but are used to judge schools and teachers. Schools with poor results can find themselves in special measures or closed even. It is not surprising that the schools should pull out all the stops to optimise the scores with children be trained up to do as well as they possibly can. The Durham tests are more general and are used for assessing pupils not schools. While it easy to see how the different results may have come about, the worrying implication is that the apparent improvements in literacy and numeracy may be only skin deep.

8.8 Another difference between the national and the Durham tests is that the national tests depend on different versions each year while the Durham tests remain the same. It is difficult to hold standards constant over time and, as we saw, there is evidence from Massey from the Cambridge Local Examinations Syndicate that the tests may have got easier. The Durham group made a similar finding for A-levels. It is hard, therefore, to be sure whether academic performance is improving, certainly at the rate the national tests and examinations appear to show. The tests and examinations are not perfect measures, but beyond that governments have a vested interest in the outcomes. This is particularly the case with the Blair government which put its reputation on the line by declaring targets and taking responsibility for delivery. It was made clear to education ministers that they themselves would be judged on pupil performance (Pollard, 2005). In the circumstances, it becomes very difficult for ministers and policy makers to look at the results of national and international tests dispassionately. Favourable results tend to become part of the government's presentation of itself to the electorate and unfavourable ones tend to be sidelined.

8.9 If there is a genuine desire to improve the education system on the basis of evidence then the monitoring of progress must be tackled in another way. This is not the place to go into details and logistics, but the crucial point is that it must be carried out by a body that is genuinely independent. The Bank of England's Monetary Committee comes to mind as an example. The prime functions of the body would be to ensure authentic assessment and to apply rigorous analysis to the results. It is possible that consistency over time could be improved by a lighter touch approach in which questions from the same test were administered to representative samples of pupils year by year. (National tests for all pupils could still be held at the end of primary education and exams at the end of secondary education as at present.) An independent national monitoring body would engender confidence in the standards

being reported and provide an objective assessment of the impact of government policies. It would take a brave government to do this, and the fate of the Schools Council liquidated in 1984 during Margaret Thatcher's premiership, serves as a warning of the dislike of politicians for monitoring that is not under their control. So any body set up or given the responsibilities outlined would have to be sufficiently powerful to withstand government pressure. One possibility would be to give the task to the National Audit Office which has demonstrated its independence on a number of occasions. Another would be for it to be part of Ofsted so that the overall statistical picture could be looked at hand-in-hand with direct observation of the classroom, though its ties to the DfES could be a problem. A third would be for a free-standing body, but its funding would have to be guaranteed so that it could not be taken away if the government did not like what it was saying.

The difficulty in deciding whether the performance of pupils is getting better, and if so to what extent, leads us to propose that an independent monitoring body should be established to monitor the performance of England's educational system both over time and in relation to other countries.

Personal Development

8.10 Even if the test and examination scores are a true representation of academic progress, which is contestable, they would be a one-dimensional view of education. The relentless pressure to improve academic performance and the numerous initiatives handed down to schools with that end in view could be leading to neglect of other equally important, perhaps even more important, aspects of education that are not so readily defined or measured. The strongest evidence that is all is not well comes in the Unicef (2007) report which we discussed in Chapter 8. Drawing together 40 indicators into six dimensions of well-being in children and young people it places the United Kingdom at the very bottom of the 21 rich countries on which it had sufficient and comparable data. The UK is last on two dimensions 'family and peer relationships' (eg finding peers kind and helpful) and 'behaviours and risks' (eg getting drunk, taking cannabis and early sexual intercourse) and next to last on 'subjective well-being' (eg ratings of satisfaction and health).

8.11 It is not clear to what extent this evident malaise is connected with schooling. On 'educational well-being', in the Unicef study, derived from the 2003 PISA scores, staying on rates and aspirations, the UK comes 17th. In terms of the percentage of young people aged 11, 13 and 15 'liking schools a lot' the UK came 16th (a curiosity here is that Finland which does extremely well in PISA came at the foot of this table). The relatively low participation in education and training in the years immediately following compulsory schooling also emerges in the DfES figures extracted in Chart 8.4, which shows that somewhat fewer are now engaged than ten years ago. Whether making it compulsory to continue as has been proposed (Demopoulos, 2006), is debatable since, as Chart 8.4, shows there is already perceptible and growing truancy. The number of secondary schools pupils with at least one unauthorised absence increased by 19.4 per cent from 1997 to 2006. Of those, the DfES (2007b) found, on data collected for the first time in 2005/06, that there were 217,390 persistent absentees, or 7.1 per cent of the 11-16 school population. Chart 8.4 also provides an insight in the way parents view what has

been happening to education. Given the improvements in test and examination scores and the extra money put in, it might have been expected that fewer would be wanting to pay high fees to send their children to independent schools. In fact, comparing the intakes of day pupils (overseas entrants tend to board) we can see that there has been an increase of 12 per cent over the past decade. This figure has to be treated with some caution since it refers only to Independent Schools Council (ISC) schools, taking over 80 per cent of the pupils. But it is fair to conclude at least that interest remains undiminished.

Chart 8.4: Other Changes 1997-2006

Measure	1997	2006
Unauthorised Pupil Absences from Secondary Schools	530,110	633,050
Not in Education or Training Age 17	121,700	132,200
Day Pupils in Independent Schools	395,940	441,758

Sources: DfES (1998, 2006 a, 2006d and 2007b); Independent Schools Council 2007).

8.12 There are a number of reasons for the continuing popularity of independent schools. They enjoy considerable academic success which shines through, as we saw in Chapter 5, even in the measures used by PISA. Many have history and tradition on their side and their income enables them to recruit good teachers, provide good facilities and have small classes. But many parents also instinctively feel that independent schools provide a more rounded education. This is not easily articulated or quantified, but it is variously expressed as self-discipline, character, and 'life skills'. Essentially, it means providing children with an inner pilot (a value system) with which to navigate their lives. Whether this is better attempted through the conventional curriculum and associated extra-curricular activities or by the direct teaching of 'well-being' and personal skills is not clear, but there has to be a risk that the intense focus on test and exam scores over the past decade will have caused these less tangible aspects of education to have been neglected.

8.13 In a study of headship (Smithers and Robinson, 2007) the headteachers of independent schools were inclined to attribute the success of their schools to their relative freedom from government interference. In comparing themselves with their state school counterparts, among other things, they said: '*They are totally driven by results and league tables. They are so, so answerable. I think it is more cold-bloodied. It ceases to be about people and starts being about statistics. It is not why people are in education I am sure*.' '*Independent school heads feel they can fulfil a vision because they have the scope to make their decisions and not have government breathing down their necks.*' '*Politicians and civil servants with no background or understanding of teaching try to impose policies that certainly seem to work on paper but take no account of the realities on the floor.* Ironically, from the outset, the Blair government saw itself as giving more freedom to schools. It saw itself as engaged in 'strategic management' (Pollard, 2005) in which it held 'the big picture' and 'the authority to intervene when things go wrong', while headteachers were 'free to determine means and proximate ends.' But, in practice, it led to a massive flow of directives from the centre (322 in 1998 alone). This was pithily described to us by one headteacher as '*independence with a big thick collar and chain'*.

8.14 The implication we would draw from this is that it is right to give schools more autonomy over the way they are run. In Chapter 5 we saw that independent schools tend to do better than state-run schools across a number of countries, even when social background is taken into account. A likely explanation that comes through strongly in qualitative studies is that freed from government micro-management schools are able to concentrate on education both academic and personal. It seems as if political thinking has been rapidly moving in that direction. Towards the end of his tenure Blair became very keen to develop 'independent state schools' through academies and trusts, and in this was supported by the Conservatives (Willetts, 2007).

8.15 It is important to emphasise, however, that the freedoms given over ethos, teaching and recruitment of staff and pupils need to be exercised in a framework that ensures fair admissions. Otherwise it is likely that children from poorer backgrounds would lose out as headteachers pick the easy to teach. The nub of that framework would be the admissions process. If parental preference is to be at the heart of it, then popular schools will receive more applications than places. This means there will have to be selection of some kind. In principle, there are five ways of selecting: academically; house location; faith; first-come-first-served, and ballot (even independent schools do not compete on price). We saw in Chapter 5 no differences showed up in maths scores in PISA 2003 associated with the preferred method of determining entry, which could mean there were no differences or the tests did not capture them. Of the possible methods the fairest is by ballot. But logic and psychology do not always correspond. Such was the outcry over Brighton's plans to go down this route (BBC News, 2007) that the council changed hands in the May elections. Recent research by the Sutton Trust (2007), however, found less resistance to the idea when set against the alternatives. The present admissions code is flexible, but also open to manipulation. Any government wanting to put a shape on the education system and ensure fair access needs to spell out its preferred means of selection. A second policy pointer we would offer is therefore:

Schools should be given genuine autonomy, but within a framework which ensures equity.

Conclusion

8.16 In reviewing the national and international data on the development of the education system under Tony Blair we conclude that it is harder than it should be to tell whether there has been real progress, in comparison with past results or the results of other countries. This is because the monitoring is under government control and government has a vested interest in the outcomes. We recommend that an independent body be set up to monitor the performance of the school system. We also found that while the government saw itself as striking a balance between control from the centre and freedom in schools, in practice it was perceived as freedom on a ball and chain. International studies show that independent schools and government-funded private schools generally obtained better results, irrespective of the social background of the intakes, than government-run schools. A likely explanation is that these schools do not continually have government on their backs. It would seem right, therefore, to give schools more autonomy in the way they are run, but this must be within a framework which secures fair admissions.

References

Barber, M (2000). 'Large-scale reform is possible.' *Education Week*, 15 November.

BBC News (2006). 'Five good GCSEs obtained by 59%', 19 October, http://news.bbc.co.uk/1/hi/education/6065436.stm

BBC News (2007). 'Schools to give places by lottery', 28 February, http://news.bbc.co.uk/1/hi/education/6403017.stm

Beaton, A.E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D.L. and Smith, T.A. (1997). *Mathematics Achievement in the Middle School Years*. Boston MA, International Study Centre, Boston College.

Brooks, G., Pugh, A.K. and Schagen, I (1996). *Reading Performance at Nine*. Slough: NFER.

Campbell, D.T. (1969). 'Reforms as experiments.' *American Psychologist*, 24, 409-429.

Chartered Institute of Personnel Development (2006) 'Employers are prioritising school leavers 'soft skills' says survey', 25 August, http://www.trainingreference.co.uk/news/gn060825.htm

De Meyer, I., Pauly, J. and Van de Poele, L. (2005). *Learning for Tomorrow's Problems: First Results from PISA*. Gent: University of Gent Department of Education.

Demopoulos, K. (2006). 'Brown call to raise school leaving age'. *Education Guardian*, 25 September.

Department for Education and Skills (1998) *Pupil Absence and Truancy from Schools in England: 1993/94-1997/98,* Statistical Bulletin B14/1998, London: the Stationery Office

Department for Education and Skills. (2004). 'OECD programme for international student assessment (PISA 2003): treatment of UK data.' *Press Notice* 2004/0197, 19 November.

Department for Education and Skills (2006a). *Participation in Education, Training and Employment by 16-18 Year Olds in England, 2004 and 2005.* National Statistics First Release SFR21/2006.

Department for Education and Skills (2006b). *National Curriculum Assessments at Key Stage 2 in England, 2006.* National Statistics First Release SFR 31/2006. London: DfES.

Department for Education and Skills (2006c). *National Curriculum Assessments at Key Stage 3 in England, 2006 (Provisional).* National Statistics First Release SFR 34/2006. London: DfES.

Department for Education and Skills (2006d). *Pupil Absence in Schools in England 2005/06 (Provisional).* National Statistics First Release SFR 35/2006. London: DfES.

Department for Education and Skills (2006e). *National Curriculum Assessment, GCSE and Equivalent Attainment and Post-16 Attainment by Pupil characteristics in England 2005/06 (Provisional).* National Statistics First Release SFR 46/2006. London: DfES.

Department for Education and Skills (2007a). 'Jim Knight puts English and maths at the heart of driving up GCSE results', *Press Notice* 11 January 2007.

Department for Education and Skills (2007b). *Pupil Absence in Maintained Secondary Schools in England in 2005/06,* National Statistics First Release SFR11/2007, 29 March.

Gardner, H. (2006). *Multiple Intelligences: New Horizons in Theory and Practice*. New York: Basic Books.

Goodhart, C.A.E. (1975). 'Monetary Relationships: A view from Threadneedle Street'. *Papers in Monetary Economics,* Volume 1, Reserve Bank of Australia.

Hilton, M. (unpublished). The measuring of standards in primary English since 1996, the PIRLS Report and the NC tests: issues of validity and accountability.

House of Commons (1999). Hansard Debates for 2 March 1999, Column 948, http://www.publications.parliament.uk

Independent Schools Council (2007). *Annual Census*, http://www.isc.co.uk/Publications_ISCCensus.htm.

Kitwood, T (1980). *Disclosures to a Stranger*. London: Routledge and Kegan Paul

Martin, M. O., Mullis, I. V. S., Beaton, A.E., Gonzalez, E. J., Smith, T.A. and Kelly, D.L. (1998). *Science Achievement in the Primary School Years*: *IEA's Third International Mathematics and Science Report.* Boston MA, International Study Centre, Boston College.

Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Gregory, K.D., Smith, T.A., Chrostowski, S.J., Garden, R.A., O'Connor, K.M. (2000). *TIMSS 1999 International Science Report.* Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade. Boston MA, International Study Centre, Boston College.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., & Kennedy, A.M. (2003). *Trends in Children's Reading Literacy Achievement 1991-2001: IEA's Repeat in Nine Countries of the 1991 Reading Literacy Study*, Chestnut Hill, MA: Boston College.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., & Chrostowski, S.J. (2004a). *TIMSS 2003 International Mathematics Report*. Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades. Chestnut Hill, MA: TIMSS & PIRLS International Study Centre, Boston College.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., & Chrostowski, S.J. (2004b). *TIMSS 2003 International Science Report*. Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades. Chestnut Hill, MA: TIMSS & PIRLS International Study Centre, Boston College.

Massey, A., Green, S., Dexter, T. and Hammet, L. (2003). *Comparability of National Tests Over Time: KS1, KS2 and KS3 Standards between 1996 and 2001*. Final Report to QCA of the Comparability Over Time Project, Research and Evaluation Division of the University of Cambridge Local Examinations Syndicate.

Miliband, D. (2003b). Interview in *The Guardian*, 5 May 2003.

Micklewright, J and Schnepf, S.V. (2006). *Response Bias in England in PISA 2000 and 2003*. London: DfES Research Report 771.

Mullis, I. V. S., Martin, M. O., Beaton, A.E.,Gonzalez, E. J., Kelly, D.L. and Smith, T.A. (1998). *Mathematics Achievement in the Primary School Years*: *IEA's Third International Mathematics and Science Report.* Boston MA, International Study Centre, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K.D., Garden, R.A., O'Connor, K.M., Chrostowski, S.J., Smith, T.A. (2000). *TIMSS 1999 International Mathematics Report.* Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade. Boston MA, International Study Centre, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kennedy, A. M. (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools.* Boston MA, International Study Centre, Boston College.

Normington, D. (3 March 2005). Letter headed *Measuring Standards in English Primary Schools: Report by the Statistics Commission on an article by Peter Tymms*, 21/02/05. London: DfES.

OECD (2001). *Knowledge and Skills for Life: First Results from PISA 2000*. Paris: OECD.

OECD (2003a). *Literacy Skills for the World of Tomorrow: Further Results from PISA 2000*. Paris: OECD.

OECD (2004). *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.

OECD (2006a) *Education at a Glance, OECD Indicators*. Paris: OECD.

OECD (2006b). *PISA 2003 Country Profiles (21 December 2006)* www.pisa.oecd.org

OECD. (2006c). *Where Immigrant Students Succeed: A Comparative Review of performance in Engagement in PISA 2003*. Paris: OECD.

Pollard, S. (2005). *David Blunkett.* London: Hodder and Stoughton.

Portal of the Principality of Liechtenstein (2004). 'Excellent PISA results for Liechtenstein', News 7 December 2004, www.**liechtenstein**.li/en/fl-portal-aktuell?newsid=10973

Prais, S.J. (1995). *Productivity, Education and Training*. Cambridge: The University Press.

Rhind, D. (16 March 2005). Reply to Letter from Normington on *Measuring Standards in English Primary Schools: Report by the Statistics Commission on an article by Peter Tymms*. London: Statistics Commission.

Smithers, A. (2001). Education policy. In *The Blair Effect*: *The Blair Government 1997-2001*. Ed Anthony Seldon, pp 405-426. London: Little, Brown.

Smithers, A. (2004). *England's Education*. London: Sutton Trust.

Smithers, A. (2005a). 'Education.' In *The Blair Effect 2001-05*. Ed Anthony Seldon and Dennis Kavanagh, pp 256-282. Cambridge: University Press.

Smithers, A. (2005b). 'Do better results mean worse exams?' *Managing Schools Today,* 22-27*,* September/October issue.

Smithers, A. and Robinson, P. (2007). *School Headship: Present and Future*. London: NUT.

Statistics Commission (2005). *Measuring Standards in English Primary Schools: Report by the Statistics Commission on an article by Peter Tymms*. London: Statistics Commission.

Strathern, M (1997). 'Improving Ratings: Audit in the British University System'. *European Review*, 5, 305-321.

Sturgis, P., Smith, P and Hughes, G. (2006). *A Study of Suitable Methods for Raising Response Rates in School Surveys*. London: DfES Research Report 721.

Sutton Trust (2007). *Ballots in School Admissions*. London: Sutton Trust.

Twist, L., Sainsbury, M., Woodthorpe, A. and Whetton, C. (2003). *Reading All Over The World.* National Report for England on PIRLS, Progress in International Reading Literacy Study. Slough: NFER.

Tymms, P. (2004). 'Are standards rising in English primary schools?' *British Educational Research Journal*, 30, 477-494.

Tymms, P., Coe, R. and Merrill, C. (2005). *Standards in English Schools*. CEM Centre, University of Durham. (Available at http://www.times-archive.co.uk/onlinespecials/english_in_schools.html)

United Nations Children's Fund Innocenti Research Centre (2007) *An Overview of Child Well-Being in Rich Countries*. Florence: Unicef.

Wilby, P. (2007). 'Every so often, public schools need and iconoclastic leader. Step up Anthony Seldon, headteacher, reformer and Blair biographer, who wants to help children be themselves', *Education Guardian*, 29 May.

Willetts, D. (2007). 'Better schools and more social mobility', *Speech to CBI Conference on Public Service Reform*, 16 May.