English Education: World Class in Primary?

John Jerrim, Natalie Perera and Peter Sellen

December 2017

EDUCATION POLICY INSTITUTE

UCL Institute of Education

Research Area:
Benchmarking English Education

About the authors

John Jerrim is Professor of Social Statistics at UCL Institute of Education. He conducted all PISA-related statistical analysis for this publication, and contributed to the writing of the report.

Natalie Perera is Executive Director and Head of Research at the Education Policy Institute. Natalie was formerly a civil servant, working in the Department for Education and the Cabinet Office until 2015.

Peter Sellen is Chief Economist at the Education Policy Institute. Peter was formerly a civil servant, working in the Department for Education, HM Treasury and Communities and Local Government.

Acknowledgements

The authors are grateful to EPI colleagues, including Jon Andrews and David Laws, as well as to colleagues in the Department for Education for their advice on the interpretation and use of the TIMSS data.

About the Education Policy Institute

The Education Policy Institute is an independent, impartial, and evidence-based research institute that aims to promote high quality education outcomes, regardless of social background.

Education can have a transformational effect on the lives of young people. Through our research, we provide insights, commentary, and critiques about education policy in England – shedding light on what is working and where further progress needs to be made. Our research and analysis spans a young person's journey from the early years through to higher education and entry to the labour market. Because good mental health is vital to learning, we also have a dedicated mental health team which considers the challenges, interventions, and opportunities for supporting young people's wellbeing.

Our core research areas include:

- Accountability and Inspection
- Benchmarking English Education
- Curriculum and Qualifications
- Disadvantaged, SEND, and Vulnerable Children
- Early Years Development
- School Funding
- School Performance and Leadership
- Teacher Supply and Quality
- Children and Young People's Mental Health

Our experienced and dedicated team works closely with academics, think tanks, and other research foundations and charities to shape the policy agenda.

This publication includes analysis of the National Pupil Database (NPD):

www.gov.uk/government/collections/national-pupil-database

The Department for Education is responsible for the collation and management of the NPD and is the Data Controller of NPD data. Any inferences or conclusions derived from the NPD in this publication are the responsibility of the Education Policy Institute and not the Department for Education.

Published in December 2017, Education Policy Institute.

ISBN: 978-1-909274-50-1

Contents

Foreword

The Education Policy Institute is an independent, impartial and evidence-based research institute which aims to promote high quality education outcomes for all, through analysis that both informs and influences the policy debate in England and internationally.

While EPI's main focus is on education in England, we are concerned to ensure both that English standards are accurately benchmarked against those of other education systems around the world and that we are able to learn from experience overseas.

We are delighted to again be working with John Jerrim of the IOE – one of the leading analysts working to compare and understand how education outcomes in England differ from those in other countries.

In this report, we look at education outcomes at the end of the primary phase of education, compare those with other OECD and world leading nations, and seek to describe the international differences using English measures of attainment.

The results are interesting, and highlight both the relatively strong performance of the most able students in England, but also the considerable gap which separates the outcomes for lower performing students in England compared with those in other countries. For top attaining pupils in primary education, England is delivering close to "world class standards" – but the results are far less impressive for our tail of lower performers.

As ever, we welcome comments on this report, and this will help inform our future work.

Rt. Hon. David Laws
Executive Chairman, Education Policy Institute

Executive summary

In August 2017, the Education Policy Institute, in partnership with the Institute of Education, published new analysis which showed how far the performance of secondary aged pupils in England would need to improve in order to match that of top performing countries, as measured by the OECD's Programme for International Student Assessment (PISA). We found that less than 40 per cent of pupils were meeting our secondary 'world-class benchmark' and that, in order to match the top-performing countries, this would need to increase to 50 per cent.

In this report, we apply a similar approach to measure the performance of primary-aged pupils. We use the Trends in Mathematics and Science Study (TIMSS) to compare England's mathematics performance with that of the top-performing countries. Based on the results of pupils in England who participated in both, we translate the TIMSS results into equivalent Key Stage 2 assessment scores to enable us to assess how well primary schools compare with the top-performing countries at a national, regional and local level.

Our main finding is that, while England compares reasonably well with other nations at primary, this hides a long tail of underperformance amongst low attaining pupils. **According to the TIMSS data, only New Zealand and Turkey have a significantly greater variation in performance than England amongst developed jurisdictions.**

Our detailed findings

Fifty-six countries and jurisdictions participated in TIMSS in 2015, with an average score of 546 points in mathematics. The five top-performing countries (Singapore, Hong Kong, South Korea, Taiwan and Japan) scored an overall average of 606 points. England scored an average of 546 points.

When we convert the 2015 TIMSS scores into equivalent 2016 Key Stage 2 assessment scores, we estimate that the average Key Stage 2 scaled score of the top-performing nations is 107 (compared to 104 in England). The Key Stage 2 assessments require pupils to achieve a scaled score of between 100 and 109 points in order to meet the expected standard in mathematics, while a score of 110 or higher is required to meet a 'high score'. **This means that the expected standard in Key Stage 2 mathematics is broadly in line with the average performance of the top-performing countries.**

In our top-performing nations, we estimate that an average of 90 per cent of pupils would have achieved the expected standard, compared to 75 per cent in England. **This means that around an additional 90,000 primary pupils in England would need to achieve the expected standards in mathematics in order for our system to be considered word-class on this basis.**

We also find that Northern Ireland performs particularly well in mathematics. In the TIMSS study, Northern Ireland scored an average of 570 points (significantly higher than England). **According to our estimated Key Stage 2 analysis, around 80 per cent of pupils in Northern Ireland are reaching a world-class standard.**

These findings are based on a cohort of pupils in England who took the new Key Stage 2 assessments in the first year that they were introduced in 2016. These results, therefore, may be subject to change as the new assessment arrangements settle in.

Nevertheless, another important finding for policy-makers in this report is that England has a long tail of low-performing pupils in mathematics. This is particularly stark when we compare the distribution of scores in England with other top-performing nations. **The variation in TIMSS scores in England is significantly higher than that of many other countries included in the study, and only significantly lower than that of two developed jurisdictions (those which are members of the OECD) - New Zealand and Turkey.**

When we convert TIMSS performance into estimated Key Stage 2 scores, we find that, **in the top-performing nations, the difference between the highest and lowest attaining pupils is around 16.2 points. In England, this increases to 18.6 points.**

When we consider how well different parts of the country have performed against our primary world-class benchmark of 90 per cent of pupils meeting the expected standard in mathematics, we find that there are no local authorities that are yet meeting this. We do, however, find considerable variation in performance across the country – consistent with our earlier findings on performance at the end of secondary school. In Kensington and Chelsea, 83 per cent of pupils met the expected standard in mathematics, compared to only 60 per cent of pupils in Bedford.

Of the top-performing 20 local authorities, 17 were in London. Conversely, there are no local authorities in London or the North East in the bottom 20 local authorities.

We also find that some authorities in the North East perform reasonably well at primary but then decline rapidly by the end of secondary. Most notable of these is Redcar and Cleveland, which ranks 16th in our primary table but then falls to 122nd out of 150 authorities by the end of secondary school (as measured by performance against our world class benchmark for secondary performance).

Areas in the East of England and the East and West Midlands perform consistently poorly in both phases of education. These include Northamptonshire, Norfolk, Peterborough and Stoke-on-Trent.

Conclusion

This new analysis finds that England performs relatively well on international rankings for primary aged pupils. However, this masks considerable variation. There is a long tail of low performing pupils – indicating that more needs to be done to support vulnerable pupils throughout the course of primary school or, indeed, earlier - and there is significant variation across the country.

The next phase of our work on benchmarking England against international standards will look at the socio-economic gap in attainment. We will have then undertaken a comprehensive analysis of the available data but we will continue to monitor England's performance across all of these phases and metrics as new data become available.

The results from the Progress in International Reading Study (PIRLS) were also published on the 5th December 2017. This assessment, also conducted by the IEA, monitors trends in pupils reading in Year 5 (Grade 4). The PIRLS 2017 report finds that England's average score increased by seven points since 2011, from 552 to 559 and that the distribution of scores appears to be narrowing slightly compared to previous cycles – with a smaller percentage of pupils failing to meet or just meeting the lower benchmark. Because this data has only just been released, we are not able to include detailed analysis of its findings in this report.

Introduction

This report looks at how well pupils in England perform in the Trends in Mathematics and Science Study (TIMSS) and compares their performance against that of the top-performing nations and jurisdictions.

What is TIMSS?

The Trends in Mathematics and Science Study is conducted by the International Association for the Evaluation of Educational Achievement (IEA). It takes place every four years and attempts to measure the knowledge and skills relative to an internationally-determined mathematics and science curriculum for pupils in both Year 5 (4th grade) and Year 9 (8th grade).[1]

TIMSS was last conducted during May and June 2015. For the purposes of this report, our focus is on the mathematics results of the Year 5 cohort who took the TIMSS test in 2015. We are interested in how they then went on to perform one year later, when they took their Key Stage 2 assessment at the end of primary school. [2]

In 2015, 49 countries and 7 jurisdictions participated in TIMSS. Pupils sit a 72 minute paper-based test covering both science and mathematics. This tests participants' knowledge, understanding and application of topics such as 'number', 'geometric shapes and measures' and 'data display'.

Because the TIMSS test is based on an internationally determined curriculum, not all of the questions covered in the test are taught within any given country's national curriculum. However, this presents less of an issue in the case of England, with around 90 per cent of TIMSS mathematics questions covered by the national curriculum.[3]

International, standardised assessments, such as TIMSS, PISA and PIRLS, are not without imperfections in their design and have been subject to criticism. The technical methods applied, including the sampling of pupils in participating countries and the wide variation in confidence levels raise questions about the reliability of such tests. There are also concerns about the impact of culture (including the prevalence of private tuition) and the breadth of the test questions. Analysis from international, standardised tests should therefore be part of a range of contextual factors and evidence which policy-makers should take into account when reviewing and improving education systems.

Chapter 1 of this report looks at how England compares with other participants in the TIMSS test. It then estimates how well pupils in the top performing nations would have performed, had they taken the Key Stage 2 assessments one year later as the England sample did. This provides us with a 'benchmark' against which we assess pupil performance in England at age 11.

---

[1] Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016), 'TIMSS 2015 International Results in Mathematics'. Available from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/

[2] England's results for 2015 are described in detail in Greany, T., Barnes, I., Mostafa, T., Pensiero, N. and Swensson, C. (2016), 'Trends in Maths and Science Study (TIMSS): National Report for England', available at: https://www.gov.uk/government/publications/timss-2015-national-report-for-england

[3] Mullis, I. V. S., Martin, M. O., Goh, S., & Cotter, K. (Eds.) (2016), 'TIMSS 2015 Encyclopedia: Education Policy and Curriculum in Mathematics and Science', available at: http://timssandpirls.bc.edu/timss2015/Encyclopedia/

Chapter 2 looks at how well England is performing against our identified benchmark. It also compares performance in different regions, local areas and the government's identified 'Opportunity Areas'.

The full methodology used to estimate performance in Key Stage 2 is set out in Annex A.

Part 1: Calculating a world class benchmark

In this Chapter, we look at how well Year 5 pupils in England performed in the 2015 TIMSS tests. We then estimate how well the leading nations would have performed had they sat the Key Stage 2 national tests, in order to create a 'world class primary benchmark'.

Identifying the world-leading countries

We have defined the 'world-leading countries' as the top five highest-attaining in mathematics, as measured by TIMSS 2015. We use mathematics rather than science because standardised assessments are only used in the former's case at Key Stage 2.

As we stated in our report on world-class standards in secondary education, it is important to recognise that, while we look at nations at the very top of international performance tables, those nations may have some undesirable features in terms of pupils' performance.[4] For instance, although a country may have high average scores, there may also be large disparities between the highest and lowest attaining pupils.

Figure 1.1 compares England to a group of countries across a range of mathematics performance indicators in TIMSS and PISA to establish the strengths and weaknesses of each education system in terms of pupils' outcomes. The red shading indicates a statistically significant worse performance than England, while blue shading indicates a significantly better performance. The '*'s indicate significant differences from England at the five percent level.

[4] Andrews, J., Jerrim, J., Perera, N., (2017), 'English education: World Class?', Education Policy Institute. available from: https://epi.org.uk/wp-content/uploads/2017/08/English-education-world-class.pdf

Figure 1.1: Key indicators across OECD and participant economies[5]

Country	Average TIMSS maths score	Average PISA maths score	10th percentile of TIMSS scores	90th percentile of TIMSS scores	Standard deviation of TIMSS scores
Singapore	618*	564*	502*	722*	86
Hong Kong	615*	548*	531*	696*	66*
South Korea	608*	524*	522*	691*	67*
Taiwan	597*	542*	505*	685*	71*
Japan	593*	532*	505*	680*	69*
Northern Ireland	570*	493	456*	675*	86
Russia	564*	494	470*	656	73*
Ireland	547	504*	451	636*	73*
Belgium (Flemish)	546	507*	468*	624*	61*
England	546	493	438	651	84
Kazakhstan	544	-	440	650	82
Portugal	541	492	447	632	72*
United States	539	470*	432	640*	81
Denmark	539	511*	440	633*	75*
Lithuania	535*	478*	441	624*	71*
Finland	535*	511*	448	619*	67*
Poland	535*	504*	441	624*	71*
Netherlands	530*	512*	457*	601*	56*
Hungary	529*	477*	412*	635*	88
Czech Republic	528*	492	437	616*	70*
Bulgaria	524*	441*	413*	624*	83
Cyprus	523*	437*	415*	623*	81
Germany	522*	506*	437	604*	65*
Slovenia	520*	510*	430	605*	69*
Sweden	519*	494	428	604*	69*
Serbia	518*	-	403*	625*	87
Australia	517*	494	408*	622*	83
Canada	511*	516*	413*	604*	75*
Italy	507*	490	413*	596*	72*
Spain	505*	486*	414*	592*	69*
Croatia	502*	464*	415*	584*	66*
Slovak Republic	498*	475*	391*	593*	80
Norway	493*	502*	400*	583*	72*
New Zealand	491*	495	371*	602*	90*
France	488*	493	390*	584*	74*
Turkey	483*	420*	354*	598*	95*
Chile	459*	423*	363*	551*	73*
Georgia	463*	404*	347*	570*	87
UAE	452*	427*	312*	587*	105*
Bahrain	451*	-	335*	561*	-
Qatar	439*	402*	312*	563*	97*
Iran	431*	-	290*	555*	102*
Oman	425*	-	293*	553*	101*
Indonesia	397*	386*	280*	509*	89
Saudi Arabia	383*	-	264*	502*	92*

[5] This table does not include Kuwait, South Africa, Morocco and Jordan

The five top-performing countries in TIMSS are therefore:

- Singapore (average TIMSS score = 618)
- Hong Kong (615)
- South Korea (608)
- Taiwan (597)
- Japan (593)

Estimating Key Stage 2 results

In order to convert this into a meaningful world-class benchmark for primary pupils, we estimate how well pupils from the top performing countries would have performed had they taken the Key Stage 2 assessments in 2016.

To do this, we look first at how well the pupils in England who took the TIMSS 2015 test then went on to perform in their Key Stage 2 assessments one year later.

This cohort of pupils who took the new Key Stage 2 assessments in 2016 was the first to be assessed against the new national curriculum and 'without levels'. Results for the first cohort of any new test can be lower than in subsequent years as teachers and schools adjust to both new content and new approaches to assessment. This has been reflected in results at a national level where the proportion of pupils achieving the expected standard in mathematics increased from 70 per cent in 2016 to 75 per cent in 2017. This was a larger year-on-year increase than was seen in the old expected standard (level 4 or above) at any point in the past decade.

Because of this, it is not unreasonable to conclude that had TIMSS been carried out in 2016, and hence matched to the 2017 Key Stage 2 data, the absolute value of the world class benchmark is likely to have been higher. Our approach to comparison means that all countries are calibrated against the same dataset and they would also see increases (as England's TIMSS results would not be subject to such volatility). So, it would not change the relative position of England in comparison to other countries, nor would it affect the relative positions of areas within England unless an area had been significantly affected by the reforms.

A total of 147 primary schools from England were randomly selected to participate in TIMSS, with one class then randomly chosen from within each school.[6] The final response rates were 98 per cent (147 schools) at the school level and 98 per cent (4,006 pupils) at the pupil level. This is fully compliant with the sample quality requirements set by the survey organisers (the IEA).

The TIMSS 2015 data has been linked by the Department for Education (DfE) to the National Pupil Database (NPD) – administrative data held for every state school pupil in the country. The NPD includes information on pupil's performance in Key Stage 2 examinations.[7] Of the 4,006 pupils that participated in TIMSS 2015, matched Key Stage 2 scores are available for 3,545 pupils (88 per cent of the final participating sample).

To estimate how well pupils from leading nations would have performed had they taken the Key Stage 2 tests, we use a 'multiple imputation' method. In other words, we predict how well pupils in other

[6] On a few occasions, two classes were chosen from the same school

[7] National Key Stage 2 results for 2016 can be found at: https://www.gov.uk/government/statistics/national-curriculum-assessments-key-stage-2-2016-revised

countries would have done had they taken the Key Stage 2 exams, based upon how well they did in the TIMSS 2015 test and the relationship between TIMSS performance and Key Stage 2 results for pupils in England who took both assessments.

A full explanation of the method used to simulate the distribution of Key Stage 2 scores in high performing nations is included in **Annex A**.

Results

Table 1.2 below shows the percentage of pupils estimated to reach the 'expected standard' of 100 test points were they to sit the Key Stage 2 mathematics exam. The right hand column shows the estimated average Key Stage 2 maths score of pupils in each country.

In England, three-quarters of children reached the expected standard and the average score is 104.[8] This puts England towards the upper part of Table 1, and compares quite well relative to many other countries. In terms of the leading East Asian countries, we predict that more than 90 percent of children in Hong Kong, Korea and Singapore would reach the expected standard in Key Stage 2 mathematics, with an average scaled score of around 107/108 points. Under both measures, England ranks 11th in the table. Interestingly, we predict that around 80 percent of pupils in Northern Ireland would reach the expected standard in Key Stage 2 mathematics, compared to 75 percent in England.

[8] This figure is different to the national proportion published by the Department for Education for maths in 2016 (70 per cent) because we are only counting those who took the TIMSS test. This difference could be explained by higher-performing schools disproportionately entering into the TIMSS test.

Table 1.2 The estimated Key Stage 2 mathematics scores of participants in TIMSS 2015

Country	Percentage reaching the expected standard	Country	Mean scaled score
Hong Kong	92%	Hong Kong	108
Korea	90%	Singapore	107
Singapore	90%	Korea	107
Taiwan	88%	Taiwan	107
Japan	88%	Japan	106
Russia	81%	Northern Ireland	105
Northern Ireland	80%	Russia	105
Belgium - Flemish	78%	Ireland	104
Ireland	77%	Belgium - Flemish	104
England	75%	England	104
Portugal	75%	Kazakhstan	104
Kazakhstan	75%	Portugal	103
Lithuania	74%	USA	103
Denmark	74%	Denmark	103
Finland	74%	Poland	103
USA	74%	Finland	103
Poland	73%	Lithuania	103
Netherlands	73%	Hungary	103
Czech Republic	72%	Netherlands	103
Hungary	71%	Czech Republic	103
Bulgaria	70%	Bulgaria	102
Germany	70%	Cyprus	102
Slovenia	70%	Germany	102
Cyprus	69%	Slovenia	102
Sweden	69%	Serbia	102
Serbia	68%	Sweden	102
Australia	67%	Australia	102
Canada	66%	Canada	102
Italy	65%	Italy	101
Spain	65%	Spain	101
Croatia	63%	Croatia	101
Slovak Republic	62%	Slovak Republic	101
Norway	60%	New Zealand	101
New Zealand	60%	Norway	100
France	58%	France	100
Turkey	58%	Turkey	100
Georgia	52%	Georgia	99
Chile	49%	UAE	99
UAE	48%	Chile	99
Bahrain	47%	Bahrain	98
Qatar	44%	Qatar	98
Oman	42%	Oman	98
Iran	42%	Iran	97
Indonesia	33%	Indonesia	96
Saudi Arabia	31%	Saudi Arabia	96

For each country, we have also looked at the simulated distribution of Key Stage 2 maths scores, and how this compares to the distribution in England. We are particularly interested in the gap between the lowest and highest performing pupils in each country. In England, we find that the difference in the average point score between those in the bottom and top deciles is 18.6 – one of the largest differences amongst the developed world.

In our five top-performing countries, the average difference between pupils in the top and bottom deciles is 16.2 points. The difference between the top and bottom performing pupils in Hong Kong is 15.7 points.

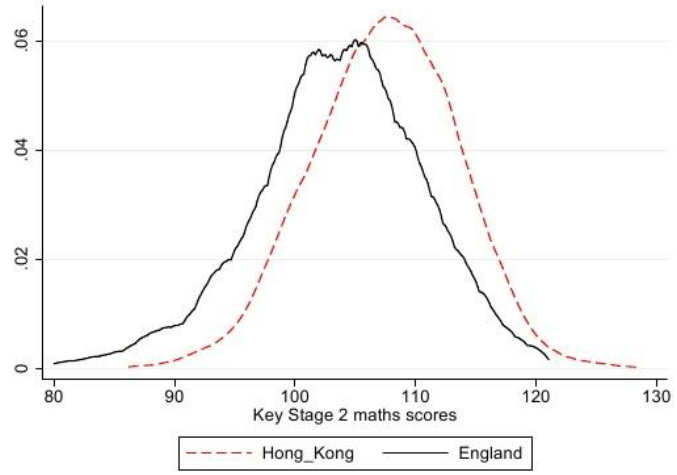
Figure 1.3 below shows the estimated Key Stage 2 attainment gap amongst all participating countries in 2015. **Taken together with the standard deviation shown in Figure 1.1, we find that the gap between the highest and lowest performing pupils in England is considerably larger than in most other participating countries. Based on performance in TIMSS, only two other developed jurisdictions (those which are members of the OECD) have a variation in pupil performance that is significantly wider than England, while most others have a significantly smaller variation.**

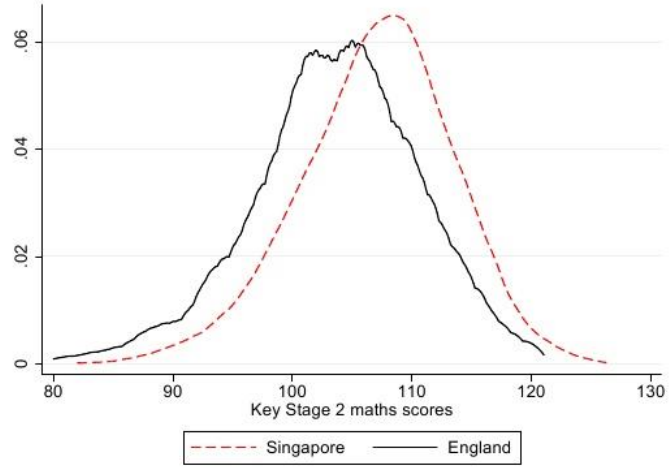
Figure 1.4 shows the distribution of scores in England compared to our top five comparator countries.

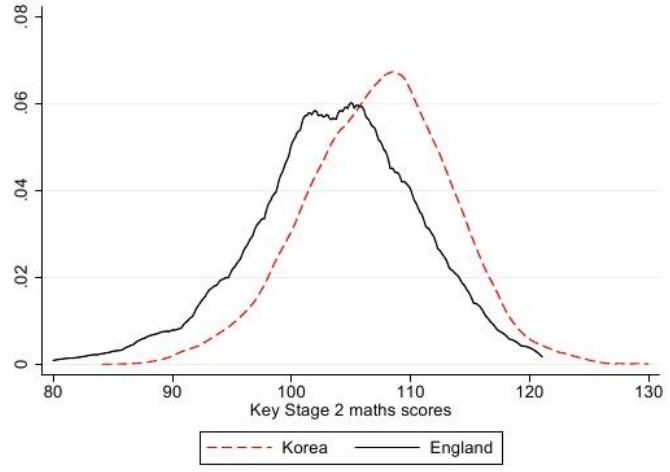
Figure 1.3: Difference in estimated Key Stage 2 mathematics scaled scores

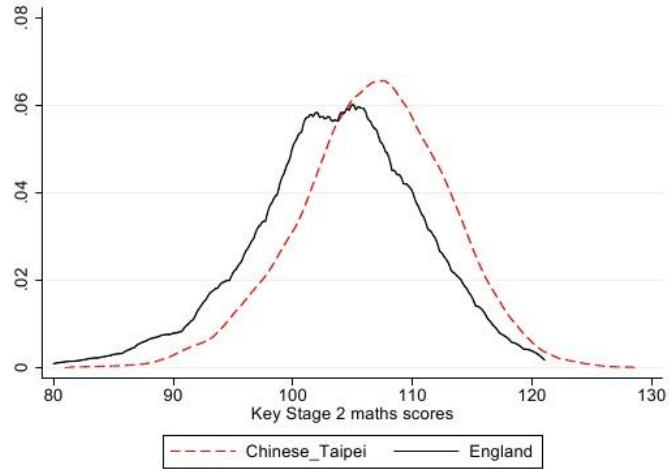
Country	Difference between lowest and highest decile (scaled score points)
UAE	18.8
England	18.6
Hungary	18.5
Turkey	18.4
Australia	18.2
New Zealand	18.2
Bulgaria	18.2
Serbia	18.1
Kazakstan	18.0
Northern Ireland	18.0
USA	18.0
Cyprus	18.0
Qatar	17.7
Slovak Republic	17.6
Denmark	17.5
Oman	17.5
Canada	17.4
Georgia	17.4
Poland	17.4
Iran	17.3
Ireland	17.3
France	17.3
Portugal	17.2
Bahrain	17.1
Russia	17.1
Lithuania	17.1
Czech Republic	17.1
Italy	17.0
Sweden	17.0
Norway	16.9
Finland	16.9
Slovenia	16.9
Spain	16.9
Germany	16.9
Singapore	16.8
Chile	16.8
Croatia	16.7
Japan	16.3
Tawain	16.2
Belguim Flemish	16.1
Korea	15.9
Netherlands	15.7
Hong Kong	15.7
Indonesia	15.6
Morocco	15.4

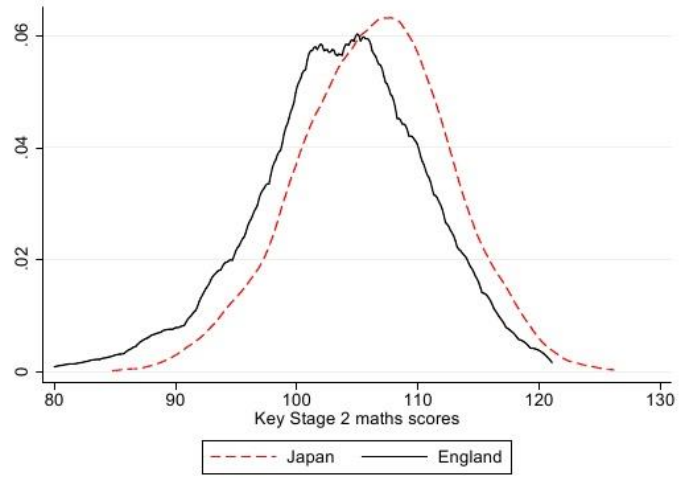
Figure 1.4: The distribution of Key State 2 assessment scores in England compared to the estimated scores in the top-performing nations











In conclusion, we find that England's maths performance by the end of primary school is relatively good when we consider the average attainment of pupils compared to the top five performing nations. However, only 75 per cent of pupils in England are achieving the expected standard in mathematics, compared to 90 per cent of pupils in the highest performing nations.

Worryingly, the gap between the highest and lowest attaining pupils in England is greater than in most other developed countries. England therefore needs to address its long tail of under-performing children if educational standards are to be comparable with the highest performing nations.

Part 2: England's performance against a world class benchmark

This section examines how well different parts of the country are performing against our world class benchmark of 90 per cent of pupils meeting the expected standard in mathematics, using data from 2016 – the year the cohort of pupils included in TIMSS 2015 took Key Stage 2 assessments.[9] The expected standard is a scaled score of 100 or above.

As well as looking at England overall, we also consider different geographical areas, including: a breakdown of the 150 local authorities;[10] areas covered by the eight different Regional Schools Commissioners; and the 12 Opportunity Areas selected by the Department for Education for targeted intervention.

Overall performance in England

In 2016, 75 per cent of pupils in England who took the TIMSS assessment achieved the expected standard in mathematics at the end of Key Stage 2.[11]

In order to be on a par with the highest performing countries, the proportion of pupils achieving the expected standard in maths needs to increase by around 15 percentage points. This equates to around 90,000 additional pupils.

Performance by local authority area

Figure 2.1 shows the proportion of pupils that achieved the expected standard (of a score of 100 or higher) and a high score (of 110 or higher) by local authority area, and the map in Figure 2.2 plots relative performance across areas.[12] We find that:

- There were no local authorities in England in which 90 per cent of pupils or more were achieving the expected standard (i.e. no local authority is currently meeting our world-class benchmark).
- The highest performing local authority was Kensington and Chelsea, where 83 per cent of primary pupils met the expected standard.
- 17 of the top performing 20 local authorities were in London. The remaining three were in the North West (Trafford), the North East (Redcar and Cleveland), and the South East (Wokingham).

[9] We have used the 2016 Key Stage 2 test data as these were the tests taken by the TIMSS 2015 cohort. A later cohort of pupils took Key Stage 2 tests in the summer of 2017 and these results are not reflected in this report.

[10] Excludes the Isles of Scilly and City of London.

[11] As noted in Footnote 8, a 5 percentage-point greater proportion of England's TIMSS sample achieved the expected standard at Key Stage 2 than did the national population. It is similarly possible that the estimated, hypothetical Key Stage 2 results of the TIMSS samples of other jurisdictions are different from those we would obtain based on their whole populations. If the deviation were in the same direction as England's, a World Class Standard for performance would be set lower than that estimated here. However, as an illustration, if we lowered our benchmark from 90 to 85 percent (assuming for example that all jurisdictions TIMSS samples would have had the same 5 percentage-point attainment difference relative to their wider populations), we would still find that no local authority in England met that standard in 2016.

[12] This analysis is based on pupils that are resident in each local authority area rather than pupils that attend schools in each local authority area.

- The three lowest performing local authorities were Bedford (60 per cent), Peterborough (61 per cent) and Norfolk (62 per cent). Of 20 local authorities with the lowest proportions reaching the expected level, six were in the East of England and six in the South East, whilst there were none in the North East or London.

This analysis also highlights the variation in some local authorities between primary and secondary performance. When we compare primary results with those set out in our analysis of the PISA data, we find that:

- Some areas in the North East demonstrated reasonably strong performance in the primary phase, but slip significantly by the end of secondary. These include Darlington, Hartlepool, Sunderland and, most notably, Redcar and Cleveland (which ranks 16th in our primary tables but falls to 122nd in our secondary tables).

- Conversely, there are a small number of other authorities (more disparately spread across the country) where relative performance improves by the time pupils leave secondary school. This includes Bedford and Rutland.

- Areas which show poor performance at both primary and secondary phases are largely concentrated in the East of England and East and West Midlands. They include Northamptonshire, Peterborough, Norfolk and Stoke on Trent.

Figure 2.1: Performance at the end of primary school in maths by local authority area - part 1 of 2

Rank	Local Authority	Percentage reaching expected standard
27	Barking and Dagenham	75.8
14	Barnet	77.8
64	Barnsley	71.6
105	Bath and North East Somerse	68.1
150	Bedford	59.6
26	Bexley	75.9
129	Birmingham	66.1
53	Blackburn with Darwen	72.5
99	Blackpool	68.7
40	Bolton	74.3
81	Bournemouth	70.2
134	Bracknell Forest	65.0
130	Bradford	65.6
28	Brent	75.4
75	Brighton and Hove	70.4
107	Bristol, City of	68.1
3	Bromley	81.2
51	Buckinghamshire	72.5
49	Bury	72.7
121	Calderdale	66.8
118	Cambridgeshire	67.0
13	Camden	77.9
139	Central Bedfordshire	64.2
52	Cheshire East	72.5
92	Cheshire West and Chester	69.3
132	Cornwall	65.6
113	Coventry	67.6
67	Croydon	70.9
117	Cumbria	67.1
39	Darlington	74.4
140	Derby	64.2
87	Derbyshire	69.6
88	Devon	69.5
142	Doncaster	63.8
147	Dorset	62.9
125	Dudley	66.2
34	Durham	74.9
30	Ealing	75.3

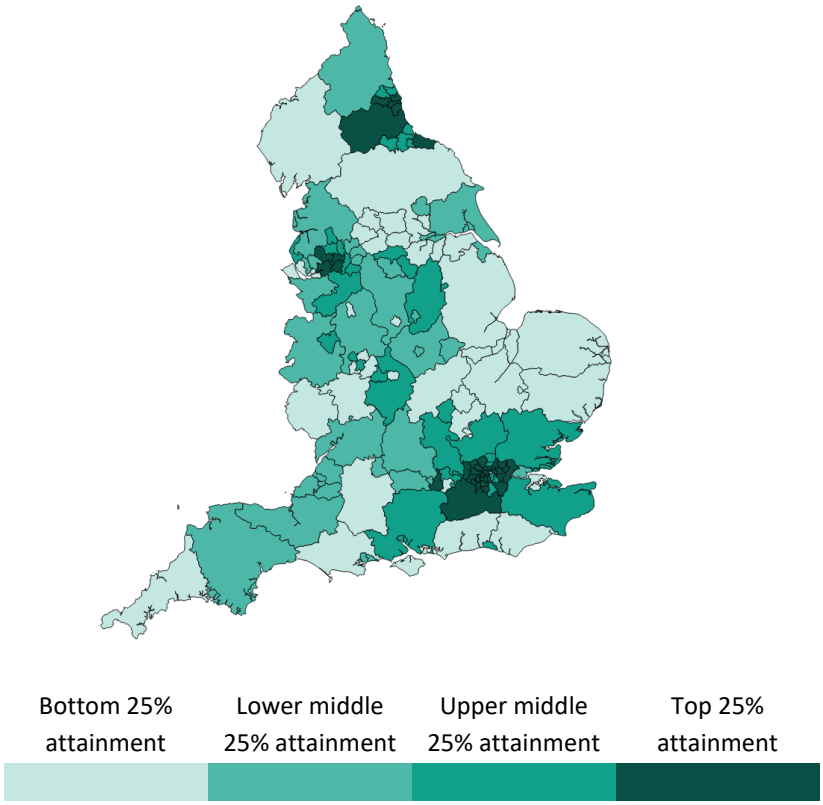
Rank	Local Authority	Percentage reaching expected standard
109	East Riding of Yorkshire	67.9
131	East Sussex	65.6
60	Enfield	71.9
63	Essex	71.6
24	Gateshead	76.1
80	Gloucestershire	70.2
17	Greenwich	77.1
12	Hackney	77.9
126	Halton	66.2
5	Hammersmith and Fulham	80.6
62	Hampshire	71.7
45	Haringey	73.8
10	Harrow	78.5
59	Hartlepool	72.0
18	Havering	77.1
116	Herefordshire	67.3
46	Hertfordshire	73.2
23	Hillingdon	76.2
20	Hounslow	76.6
146	Isle of Wight	63.1
33	Islington	75.1
1	Kensington and Chelsea	83.0
61	Kent	71.8
97	Kingston upon Hull, City of	69.0
8	Kingston upon Thames	79.0
114	Kirklees	67.5
110	Knowsley	67.8
15	Lambeth	77.8
77	Lancashire	70.3
127	Leeds	66.1
78	Leicester	70.3
101	Leicestershire	68.5
48	Lewisham	73.0
119	Lincolnshire	67.0
133	Liverpool	65.5
144	Luton	63.7
66	Manchester	71.0
135	Medway	64.7

Figure 2.1: Performance at the end of primary school in maths by local authority area - part 2 of 2

Rank	Local Authority	Percentage reaching expected standard
35	Merton	74.8
68	Middlesbrough	70.8
56	Milton Keynes	72.3
55	Newcastle upon Tyne	72.3
6	Newham	79.9
148	Norfolk	62.4
96	North East Lincolnshire	69.1
115	North Lincolnshire	67.4
82	North Somerset	70.1
47	North Tyneside	73.2
124	North Yorkshire	66.7
128	Northamptonshire	66.1
102	Northumberland	68.4
91	Nottingham	69.4
73	Nottinghamshire	70.6
111	Oldham	67.7
95	Oxfordshire	69.1
149	Peterborough	60.9
93	Plymouth	69.2
94	Poole	69.2
141	Portsmouth	64.0
108	Reading	68.1
21	Redbridge	76.5
16	Redcar and Cleveland	77.7
2	Richmond upon Thames	81.9
76	Rochdale	70.3
58	Rotherham	72.2
112	Rutland	67.7
36	Salford	74.8
74	Sandwell	70.5
44	Sefton	73.9
98	Sheffield	68.9
89	Shropshire	69.5
50	Slough	72.6
54	Solihull	72.4
106	Somerset	68.1
83	South Gloucestershire	70.1
32	South Tyneside	75.1

Rank	Local Authority	Percentage reaching expected standard
72	Southampton	70.7
70	Southend-on-Sea	70.8
42	Southwark	74.0
79	St. Helens	70.2
100	Staffordshire	68.5
41	Stockport	74.3
57	Stockton-on-Tees	72.2
143	Stoke-on-Trent	63.8
137	Suffolk	64.5
31	Sunderland	75.2
37	Surrey	74.5
7	Sutton	79.4
85	Swindon	69.8
84	Tameside	69.8
65	Telford and Wrekin	71.3
104	Thurrock	68.2
103	Torbay	68.3
9	Tower Hamlets	78.9
4	Trafford	81.2
122	Wakefield	66.8
123	Walsall	66.7
19	Waltham Forest	76.8
25	Wandsworth	76.0
29	Warrington	75.4
71	Warwickshire	70.7
86	West Berkshire	69.6
145	West Sussex	63.4
22	Westminster	76.5
38	Wigan	74.5
120	Wiltshire	66.8
43	Windsor and Maidenhead	73.9
136	Wirral	64.5
11	Wokingham	78.0
69	Wolverhampton	70.8
138	Worcestershire	64.3
90	York	69.4

Figure 2.2 Performance in maths at the end of primary school by local authority



Performance by Regional Schools Commissioner region

In 2014 the Department for Education introduced eight Regional Schools Commissioners (RSCs), primarily to provide an additional layer of oversight as part of the academies and free schools programmes. The RSCs are split across eight regions, as shown in Figure 2.3 below.

Figure 2.3: Coverage of the RSCs

The RSCs have a range of responsibilities, including intervening in under-performing academies and free schools; supporting the development of academy sponsors and taking action to improve poorly performing sponsors; considering applications from local authority schools to convert to academy status; advising on new free schools; and brokering support for underperforming local authority schools.

As Figure 2.4 demonstrates, differences between RSC regions were modest compared to the variation seen at the local authority level, with around 4 percentage points separating all eight. The two regions containing London had the highest proportions reaching the expected standard, with 72 per cent, whilst the West Midlands had 68 per cent doing so.

Figure 2.4: Performance in maths at the end of primary school by RSC region

Rank	RSC region	Percentage of pupils reaching the expected standard
1	North West London & South Central	72.0
2	South London & South East	72.0
3	North	71.6
4	North East London & East	70.8
5	Lancashire & West Yorkshire	69.7
6	East Midlands & Humber	68.8
7	South West	68.2
8	West Midlands	68.1

Performance in Opportunity Areas

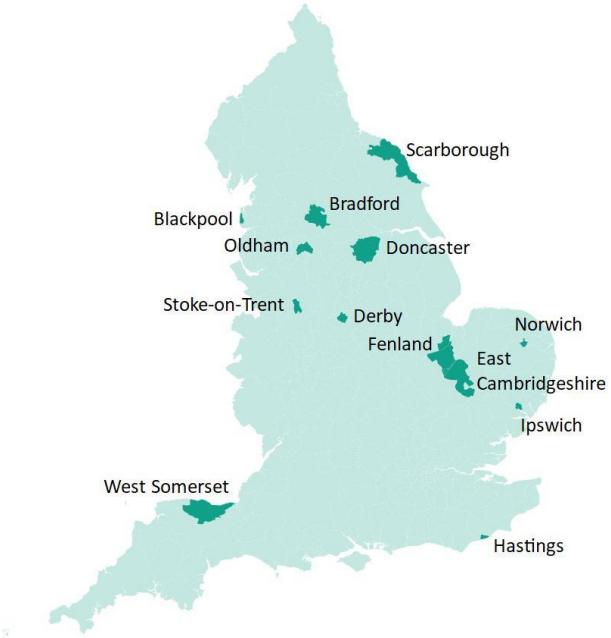
In October 2016, the Department for Education announced that it had identified six areas in the country which were 'the most challenged when it comes to social mobility'. These areas were: West Somerset, Norwich, Blackpool, Scarborough, Derby, and Oldham. They were identified using the Social

Mobility Index, published by the Social Mobility Commission in January 2016.[13] The six identified areas were not the six worst-performing areas identified by the Commission – West Somerset and Norwich were the worst two, but the remaining four areas fell within the 9th and 30th worst performing areas under this measure.

In January 2017, the Secretary of State announced a further six Opportunity Areas: Bradford, Doncaster, Fenland & East Cambridgeshire, Hastings, Ipswich, and Stoke-on-Trent, along with a further investment of £3.5 million (£2m from the Education Endowment Foundation and £1.5m from the Department for Education) to establish a research school for each of the 12 Opportunity Areas.

Figure 2.5 below shows the location of Opportunity Areas and Figure 2.6 shows the proportion of pupils achieving the expected standard in each of those areas (the Department for Education does not publish data on the proportion of pupils achieving a higher score in these areas).[14]

Figure 2.5: Locations of the Opportunity Areas



We find that, overall, the proportion of pupils in Opportunity Areas achieving the expected standard was lower than elsewhere (65 per cent compared to 71 per cent). Opportunity Areas did, however, demonstrate a wide range in performance: **69 per cent of pupils in Blackpool reached the expected standard, while that figure falls by 10 percentage points to 59 per cent in Norwich.**

[13] See: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/496103/Social_Mobility_Index.pdf.
[14] Figures for Fenland and East Cambridgeshire are shown separately.

Figure 2.6: Performance at the end of secondary school by Opportunity Area

Local authority district	Percentage of pupils reaching the expected standard
Blackpool	68.7
Oldham	67.7
Bradford	65.6
East Cambridgeshire	65.1
Ipswich	65.0
Derby	64.2
Doncaster	63.8
Stoke-on-Trent	63.8
Fenland	62.8
Hastings	61.3
West Somerset	60.4
Scarborough	59.8
Norwich	58.6
Pupils in all opportunity areas	**64.6**
Pupils in other areas	**70.5**

Part 3: Conclusions and policy implications

Whilst not yet reaching a 'world-class standard', primary aged pupils in England, on average, perform well in mathematics when compared to other developed countries.

This report, however, uncovers three important issues for the government and school leaders to address.

The first is that, in order to match the highest-performing countries, around an additional 90,000 pupils will need to meet the expected standard in mathematics by the end of primary school. This would move current national performance levels in England from 75 per cent to a world-class standard of 90 per cent.

The second is that England has one of the largest gaps between the highest and lowest performing pupils in participant countries. The overall variation in attainment in England, its standard deviation, is significantly higher than many other countries included in the study, and only lower than that of two other developed jurisdictions.

The third is the regional and local variation in performance amongst primary-aged pupils. This issue is a recurring finding from EPI and wider research and is acutely visible across all phases of education.

The 12 Opportunity Areas highlighted by the government for targeted intervention are, indeed, demonstrating poor performance using our world-class benchmark for primary-aged pupils. But, as our local authority performance table shows, there are many other areas where performance is as poor or, indeed, worse than that of those 12 Opportunity Areas.

Left too long, the country's most vulnerable children are at risk of falling even further behind. In 2018, we will be publishing two further pieces of research that will help to further our understanding of this gap. The first will look at the relationship in England between the attainment gap and the socio-economic gap, using international data as our benchmark. The second will look at the composition of the country's lowest attaining pupils and begin to ask what policies need to be put in place in order to improve their life-chances.

Annex A: Methodology

Our empirical methodology is based on multiple imputation. The TIMSS-NPD file for England includes both children's TIMSS test scores (plausible values) and their scores in the Key Stage 2 mathematics test. We append to this the public use TIMSS datafiles for all other comparator countries. Hence we have a set of variables (TIMSS scores) which are observed for all participating pupils in all countries, and another set of variables (Key Stage 2 mathematics scores) which are only observed for pupils in England. The fact that Key Stage 2 grades are not observed in other countries is treated as a missing data problem, which we attempt to solve via multiple imputation. In other words, we predict how well children in other countries would have done had they taken Key Stage 2 exams, based upon how they performed on the TIMSS 2015 test. This prediction is based upon the relationship between TIMSS and Key Stage 2 scores in England. One advantage of using multiple imputation by chained equations is that we are able to retain in our analysis even those pupils in England whose Key Stage 2 data could not be matched. Hence all pupils who participated in TIMSS 2015 in England are included in our results. This includes pupils in independent schools, who do not typically sit the Key Stage 2 tests.

The imputation model

Our imputation model applies multiple imputation by chained equations (MICE). A simple Ordinary Least Squares regression model underpins the prediction of Key Stage 2 scores, and this is of the form:

$$\mathrm{KS2\_Scores}_i = \alpha_i + \beta.TIMSS_i + \varepsilon_i$$

Where:

$\mathrm{KS2\_Scores}_i$ = The Key Stage 2 mathematics scores of child i

TIMSS = Children TIMSS scores, using the first plausible value15

In our main specification, TIMSS scores are entered as dummy variables in terms of the international benchmarks, using the first plausible value only. The final pupil senate weight is applied, with the imputation models run separately for England in combination with each comparator country. In our companion report using the PISA data, we have run further robustness tests, and found that results do not change substantially if a more complex imputation model is estimated, or if raw PISA scores (percentage correct) rather than scaled

Once the multiple imputation stage is complete, we follow Rubin's Rules (Rubin 1987) to predict the likely distribution of performance on TIMSS examinations in each comparator country. From these results, we can infer how England's Key Stage 2 achievement distribution needs to change in order for it to become one of the leading education systems in the world.

[15] We have experimented with using all ten plausible values and find very little change to our substantive results.