HR 01
Ymchwiliad i hawliau dynol yng Nghymru
Inquiry into Human Rights in Wales
Ymateb gan: J. Jones
Response from: J. Jones

I would like to bring this opinion to the attention of the Committee.

It has long been the case in Wales that public service employers can discriminate in favour of Welsh speakers.

It has long been the case that the Welsh Government has encouraged the increase of Welsh medium schools by making it mandatory on LAs to survey parental preference for Welsh medium schooling and obliging those LAs to provide Welsh medium schools.

The Welsh government has never obliged any LA to survey parental preference for English medium schooling and has never insisted to LAs that they provide English medium schooling for parents who would like their children to be schooled in their home language if that language is English.

The Welsh government has therefore consistently discriminated against the rights of the child and they should in future enshrine in law the right of parents to demand that their child is educated in EITHER of the two legal languages of Wales.

Since 2011 it has been the law that:-

"In Wales, the Welsh language should be treated no less favourably than the English language.
Persons in Wales should be able to live their lives through the medium of the Welsh language if they choose to do so."

The term "no less favourably..." essentially means that at no time can the English language or the interests of someone speaking English or wishing to use only English be given any kind of preference.

A sign on a public building can be in Welsh only but not in English only.

A Welsh sign on the same line as an English sign must be to the left of the English so that it is read first.
A sign on a public building written in Welsh cannot be below one written in English and the font size of the Welsh sign can be bigger than the English version but not smaller.
All recorded messages on phones in public bodies must be in Welsh first.
An employee can not be disciplined for demanding the right to speak Welsh but an employer can discipline an employee for insisting that he has a right to speak only English.

So far the extent to which the 2011 Welsh language act can be taken has not been fully explored. Already LAs are complaining at the cost of providing Welsh language services where no measurable demand exists...Blaenau Gwent, Torfaen, Newport, Merthyr and Monmouthshire have so few people who, according to the Welsh language use survey, are "more comfortable speaking Welsh", that when I asked for a percentage of all those LAs taken together the number preferring Welsh was a statistical zero.
Jobs are given to fluent (and usually first language) Welsh speakers in many LAs and denied to monoglot English employees for no discernible reason.

When it comes to "human rights" in wales there is a massive and state generated discrimination against all those people who are most comfortable using English and want to live their lives through English. Only 4.4% of the Welsh population is most comfortable using Welsh.

The 2011 Welsh language act must be re-drafted to give English language rights to the 95% majority.

The present state of affairs is a monstrous injustice and it is a proven fact that pupils who have no Welsh speaking Parent do not reach their full potential in Welsh medium schools. See attachments.

Links

- Achievement of 15-Year-Olds in Wales: PISA 2015 National Report – December 2016 (John Jerrim and Nikki Shure)
- If you don’t understand, how can you learn? Global Education Monitoring Report

HR 01a
Ymchwiliad i hawliau dynol yng Nghymru
Inquiry into Human Rights in Wales
Ymateb gan: J. Jones
Response from: J. Jones

Achievement of 15-Year-Olds in Wales: PISA 2015 National Report

December 2016

John Jerrim and Nikki Shure.

UCL Institute of Education.

Contents

List of figures

Chapter 5

Chapter 6

Chapter 7

List of tables

Chapter 8

Chapter 9

Chapter 10

Executive summary

Introduction

The Programme for International Student Assessment (PISA), led by the Organisation for Economic Co-operation and Development (OECD), provides new evidence on how the achievement and abilities of 15-year-olds varies across countries. PISA has been conducted every three years since 2000, with Wales having participated in each round since 2006. Over 70 countries participated in the 2015 edition of PISA, including all members of the OECD and all four countries within the UK. Pupils were tested in four subjects (science, mathematics, reading and collaborative problem solving), while contextual information was also gathered from all participating pupils and schools. Each time PISA is conducted, one subject is the focus. In 2015, it was the turn of science, having last been the focus of PISA in 2006. A major change was made to how PISA was conducted in 2015, with computer-based assessment (CBA) used in the main study for the first time.

This national report for Wales is published simultaneously with the OECD's international report on PISA 2015. It complements the OECD's report by (i) providing a more focused comparison of Wales with other countries and (ii) providing analysis of differences within Wales across school and pupil characteristics.

International comparisons of Wales in the national report include contrasts with a number of different groups. This includes the average across industrialised countries (the 'OECD average') and the average across the 10 countries with the highest average PISA scores (usually in reference to the science domain). The 10 'high-performing' countries in PISA science are Singapore, Japan, Estonia, Taiwan, Finland, Macao, Canada, Vietnam, Hong Kong and China. The report reveals that pupil attitudes and outcomes, along with headteachers' views, often vary widely among these high-performing countries.

Analysis of differences within Wales is enhanced by linking PISA to administrative records about pupils and schools. This allows us to consider for the first time how PISA scores differ between different school types (e.g. Welsh versus English medium schools), by the Welsh National School Categorisation System, and by various pupil characteristics such as Welsh language and eligibility for Free School Meals (FSM).

While the analysis in each chapter uncovers correlations, it does not establish cause and effect. Changes in PISA 2015 results from previous cycles should not be taken as providing evidence as to the impact of any previous or ongoing educational reform.

Achievement in science

The average PISA science score for Wales in 2015 was 485. This is 20 points lower than the average in 2006 (505). There are 29 countries where the mean science score is at least 10 points ahead of Wales, and 31 countries where the mean science score is at least 10 points lower. Wales, along with Finland, Australia and New

Zealand, is an example of a country in which there has also been a sustained fall in average science scores since 2006. Portugal and Macao are two of the few countries where there has been a statistically significant and sustained improvement in science achievement over the last decade.

The top-performing 10 per cent of pupils in Wales achieved a PISA score of at least 602 points. There are 33 countries where the top 10 per cent of pupils are more than 10 points ahead of their peers in Wales. In comparison, the lowest achieving 10 per cent of pupils in Wales score below 368 on the PISA science test. However, there were only 18 countries where the lowest 10 per cent of pupils were ahead of Welsh pupils by more than 10 PISA test points. The comparatively low performance of Wales' high achievers in science is therefore a notable weakness of the Welsh educational system.

Achievement in different aspects of scientific literacy

Pupils in Wales achieve similar scores in what PISA defines as the 'living' scientific system (which roughly equates to biology), the 'physical system' (which measures knowledge about matter, motion and forces), and 'earth and space sciences' (looking at earth's history, the earth in space, and the universe). The PISA 2015 test also examines skills in three core scientific competencies: 'interpreting data and evidence scientifically', 'evaluating and designing scientific enquiry' and 'explaining phenomena scientifically'. Pupils in Wales are, on average, slightly stronger at explaining phenomena scientifically than they are at evaluating and designing scientific enquiry. This pattern of results is reasonably uncommon, and is not found in many of the highest performing countries.

Achievement in mathematics

The average PISA mathematics score for Wales in 2015 was 478. The average score has fluctuated over the last decade, but is at a similar level in 2015 as it was in 2006 (484). There are 33 countries where the mean mathematics score is at least 10 test points ahead of Wales, and 28 countries where the mean mathematics score is at least 10 test points lower. The top seven ranked jurisdictions in PISA mathematics are all within East Asia.

A number of countries have caught up or overtaken Wales in mathematics over the last decade, including Italy, Portugal and Russia. On the other hand, average scores in the Czech Republic, Australia, New Zealand and Iceland have all declined since 2006.

The lowest-performing 10 per cent of pupils in Wales achieved a PISA mathematics score below 377 points. There are 21 countries where the bottom 10 per cent of pupils in mathematics are more than 10 test points above their peers in Wales. In comparison, the highest achieving 10 per cent of pupils in Wales score above 578 points on the PISA mathematics test. There are 40 countries where the highest achieving pupils are at least 10 test points ahead of the highest achieving pupils in Wales. In only three OECD countries (Turkey, Mexico and Chile) is the mathematics performance of the highest achievers lower than in Wales. Due to this comparatively low performance of high achieving pupils, inequality in 15-year-olds mathematics scores is lower in Wales than almost anywhere else in the industrialised world. Nevertheless, the comparatively low performance of Wales' high achieving pupils in mathematics is a significant weakness of the Welsh education system.

Achievement in reading

The average PISA reading score for Wales in 2015 was 477. This has remained stable since 2006 (481). There are 31 countries where the mean reading score is at least 10 test points ahead of Wales, and 29 countries where the mean reading score is at least 10 test points lower. Countries with a similar average reading score to Wales include Lithuania, Israel and Luxembourg.

Although Wales' average reading score has remained stable, there have been changes in the performance of a number of other countries over the last decade. Some of the higher-performing countries in 2006 have experienced a decline in PISA reading scores, including South Korea (556 to 517), Finland (547 to 526) and New Zealand (521 to 509). Meanwhile, other countries have caught up or overtaken Wales in reading, including Russia (440 in 2006 to 495 in 2015), Spain (461 to 496) and Portugal (472 to 498).

The lowest-performing 10 per cent of pupils in Wales achieved a PISA reading score below 368 points. There are 17 countries where the bottom 10 per cent of pupils in reading are more than 10 test points above their peers in Wales. In comparison, the top 10 per cent of pupils in Wales achieve a PISA reading score of more than 588 points. There are 37 countries where the reading scores of the top 10 per cent of pupils are at least a quarter of a school year higher. Turkey, Mexico and Chile are the only members of the OECD where the PISA reading scores of the top 10 per cent are significantly lower than in Wales. Consequently, the gap between the highest and lowest achieving pupils in reading in Wales stands at 219 test points; this is amongst the lowest anywhere in the industrialised world (OECD average 249 points). Nevertheless, this again highlights the comparatively low skills of the top 10 per cent of pupils in Wales.

Variation in scores by pupil characteristics

In Wales there is no evidence of a gender difference in pupils' science scores. The mathematics skills of boys in Wales are, on average, around 10 test points ahead of girls. This is not an unusual finding; there is a similar gender gap in mathematics skills in many other OECD countries, and has also been present in Wales in previous PISA cycles. Girls in Wales achieve higher average reading scores than boys. However, the same also holds true in every other developed country, and at 10 test points, the gender gap in reading skills in Wales is actually among the smallest anywhere in the world.

Although there are clear socio-economic differences in 15-year-olds' PISA scores, socio-economic inequality is actually much lower in Wales than the rest of the UK, and compared to most other countries across the world. In Wales, the gap between pupils from the most and least advantaged 25 per cent of families in Wales is around

50 test points in science. This is much smaller than the average across industrialised countries (88 points). However, this small gap is at least partly driven by the comparatively weak academic performance of pupils from the most advantaged socio-economic backgrounds in Wales relative to their equally advantaged socio-economic peers in other industrialised countries.

Around 29 per cent of 15-year-olds from disadvantaged socio-economic backgrounds in Wales manage to achieve a PISA science score that puts them in the top 25 per cent of test takers internationally. When looking across countries, it is apparent that there is little association between the use of academic selection to assign pupils into different secondary schools and the proportion of disadvantaged pupils who manage to succeed academically against the odds.

Pupils who took the PISA test in Welsh achieved PISA science and reading scores more than 20 test points behind pupils who took the test in English. In fact, even pupils who study Welsh as their first language did better on the PISA science test if they decided to take the assessment in English rather than Welsh.

Differences in achievement between schools

In Wales, there are bigger differences in achievement amongst 15-year-olds who attend the same school than there are differences in achievement between pupils who attend different schools. This is not unusual for a country with a mainly comprehensive schooling system, with a similar finding occurring across a diverse set of countries within the OECD (e.g. Finland, South Korea, United States). The same does not hold true in many countries where academic selection into secondary schools is used, such as the Netherlands and Germany, where differences in achievement are just as big between schools as they are within schools.

Pupils who attend a Welsh medium school achieve similar average PISA scores to pupils who attend an English medium school. There are clear differences in achievement depending upon the National School Categorisation System category of the school. In science, pupils in green coded schools achieve an average PISA score of 497, compared to 485 for pupils in yellow category schools and 471 for those in the amber category. Differences of a similar magnitude across the support categories also occur in reading and mathematics.

School management and resources

A lack of good quality school infrastructure stands out as a particular concern of headteachers in Wales. This is especially true for headteachers who are leading schools requiring more support and those leading English medium schools. For instance, more than half of headteachers who manage amber and red coded schools reported this to be a factor hindering instruction, compared to a third of teachers in green category schools.

Another key concern of headteachers in Wales is the level of absenteeism amongst their staff; a quarter of secondary pupils are taught in schools where the headteacher believes that this is hindering pupils' learning. This is above the OECD average and the average across the 10 countries with the highest average PISA science scores.

Within Wales, staff absenteeism, teachers not being prepared for class and teachers not meeting pupils' needs are key concerns amongst headteachers who manage schools in the amber and red support categories.

Pupils' aspirations and future plans

Most pupils in Wales believe that the content of their school science lessons is helping to prepare them for the future; 74 per cent agree that it will help them to get a job and 78 per cent that it will improve their career prospects. This is similar to the average across the 10 high-performing countries, and holds true irrespective of pupils' gender, socio-economic status and level of academic achievement.

Around a quarter of pupils (26 per cent) in Wales hope to be working in a science related career by age 30. This is above the average across industrialised countries (24 per cent) and the average across high-performing countries (22 per cent). Boys are more likely to want to become a scientist, engineer or ICT professional than girls, who are more likely to aspire to work in a heath related field. There is no evidence that countries with higher average PISA science scores have a greater proportion of 15-year-olds who expect to be working in a science career at age 30.

Around a third of 15-year-olds in Wales expect to obtain an undergraduate degree. Girls (40 per cent) are more likely to expect to complete university than boys (30 per cent), while over half of Welsh pupils from the most advantaged socio-economic backgrounds expect to complete university, compared to a fifth of pupils from low socio-economic households. Course content, employment prospects and entry requirements are the most important factors influencing 15-year-olds thoughts about which university to apply to, while distance from home, fitting-in and social life are the least important. Among the subset of 15-year-olds who plan to apply to university, over a third intend to leave Wales and study in another part of the UK. Over half of those who plan to apply to university listed a Russell Group institution as their first choice.

Pupils' experiences of their time in science classes at school

Secondary school pupils in Wales report having almost five hours of timetabled science lessons per week, which is more than the OECD average (3.5 hours) and the average across the high-performing countries (four hours). However, there is no evidence that countries with more hours of instruction in science have higher average PISA scores. In only two out of the 10 high-performing countries are additional study hours (i.e. hours outside of pupils' regular timetable) reported to be much higher than the 18 hours in Wales. These are Singapore (22 hours) and China (27 hours).

There is more frequent low-level disruption in science classrooms in Wales than in the average high-performing country. For instance, 41 per cent of 15-year-olds in Wales reported that pupils regularly do not listen to what their science teacher says, while 45 per cent of pupils say that there is frequent noise and disorder. This compares to an average across the 10 high-performing countries of around 20 per cent. There is a particularly stark contrast between science classrooms in Wales and science classrooms in the high-performing East Asian nations in this respect.

Less than half of pupils in Wales report that their science teacher provides them with regular feedback, such as how they are performing on their course (33 per cent), their areas of strength (36 per cent) and areas for improvement (40 per cent). However, Wales is not unusual in this respect, with an even smaller proportion of pupils saying that they receive regular feedback from their science teachers in the high-performing countries.

PISA across the UK

The average PISA science score in England (512) is significantly higher than in Northern Ireland (500) and Scotland (497). Pupils in each of these three countries achieve significantly higher science scores than pupils in Wales (485). In reading and mathematics, average scores are similar across England, Northern Ireland and Scotland, with Wales again significantly behind the rest of the UK.

Whereas average PISA scores have remained stable in England and Northern Ireland since 2006, there has been a sustained 20 point decline in science scores in Wales. Similarly, there has been a 15 point decline in PISA mathematics scores in Scotland between 2006 and 2015.

One-in-three (32 per cent) pupils in Wales lacks basic skills in at least one of the three PISA domains, compared to 29 per cent in England and Scotland, and 25 per cent in Northern Ireland. Across the UK, around 10 percent of pupils lack basic skills in all three PISA subject areas. In England, 18 per cent of pupils are classified as a high-achiever in at least one of the PISA subjects, compared to 13 per cent in Scotland, 11 per cent in Northern Ireland and eight per cent in Wales.

In Scotland, Northern Ireland and Wales, the science skills of the top 10 per cent of pupils have declined by more than 20 PISA test points between 2006 and 2015. The same is not true in England, where the PISA scores of the top 10 per cent of pupils has remained broadly stable over the last decade.

Socio-economic differences in 15-year-olds PISA scores are smaller in Wales than in the rest of the UK. This is due to the comparatively weak academic performance of pupils from the most advantaged socio-economic backgrounds in Wales, relative to their equally advantaged socio-economic peers in England, Scotland and Northern Ireland.

A lack of teaching staff and teachers not meeting individual pupils' needs stand out as a particular concern amongst headteachers in England and Scotland; more so than for headteachers in Northern Ireland and Wales.

Chapter 1. Introduction

1. The Programme for International Student Assessment (PISA) is a global benchmarking study of pupil performance by the Organisation for Economic Co-operation and Development (OECD)[1]. It provides a comparison of what 15-year-olds within participating countries know and can do in the core subjects of science, reading and mathematics. Additionally, contextual information collected from pupils and their school enables associations between performance and other factors, such as pupil engagement or teaching resources, to be compared between and within participating countries.

2. The inaugural PISA study took place in 2000, and has since been conducted on a three-year cycle. All OECD members participate in PISA, with Table 1.1 providing a list of countries and 'economies' (geographic regions within countries) that took part in 2015[2]. Members of the OECD are highlighted in bold[3].

3. Although 75 countries participated in PISA 2015, four countries have been excluded from the international report due to issues with the sampling frame, failure to meet the OECD response rate criteria, or issues with the marking. These four countries (Argentina[4], Malaysia, Kazakhstan and Cyprus) are excluded from this report, bringing the total number of countries down to 71[5].

4. In Wales, PISA was conducted between November and December 2015. A total of 140 schools and 3,451 pupils took part. The study was carried out on behalf of the Welsh Government by a consortium of RM Education, UCL Institute of Education and World Class Arena Limited. Throughout this report, we refer to this consortium as the National Centre.

---

[1] The OECD is an international organisation of industrialised countries. Its mission is to *'promote policies that will improve the economic and social well-being of people around the world*'.
[2] Four provinces within China participated in PISA 2015: Beijing, Guangdong, Jiangsu and Shanghai. For convenience, we refer to the results for these four provinces combined as 'China'. However, when interpreting the results, it is important to remember that the PISA sample for 'China' is based upon only these four regions.
[3] See NCES Website for a list of countries that have participated in each round of PISA.
[4] Although the OECD have deemed the data for Argentina to be unrepresentative, the region of Buenos Aires did satisfy the sampling criteria. This region of Argentina has therefore been included in the OECD tables. However, the whole of Argentina (including Buenos Aires) is excluded from this report.
[5] Additionally, in Albania, pupils' responses to the background questionnaire cannot be linked to the PISA test score data. Following the OECD, we will include Albania in all international comparisons of PISA test scores. However, Albania will be excluded from any analysis linking PISA scores to background information, such as gender and socio-economic status.

Table 1.1 Countries participating in PISA 2015

Albania	**Hungary**	Peru
Algeria	**Iceland**	**Poland**
Argentina+	Indonesia	**Portugal**
Australia	**Ireland**	Qatar
Austria	**Israel**	Romania
Belgium	**Italy**	Russia
Brazil	**Japan**	**Scotland**
Bulgaria	Jordan	Singapore
Canada	Kazakhstan+	**Slovakia**
Chile	**South Korea**	**Slovenia**
"China"*	Kosovo	**Spain**
Colombia	**Latvia**	**Sweden**
Costa Rica	Lebanon	**Switzerland**
Croatia	Lithuania	Taiwan
Cyprus+	**Luxembourg**	Thailand
Czech Republic	Macao	Trinidad and Tobago
Denmark	Macedonia	Tunisia
Dominican Republic	Malaysia+	**Turkey**
England	Malta	United Arab Emirates
Estonia	**Mexico**	**United States**
Finland	Moldova	Uruguay
France	Montenegro	Vietnam
Georgia	**Netherlands**	**Wales**
Germany	**New Zealand**	
Greece	**Northern Ireland**	
Hong Kong-China	**Norway**	

Notes: Table includes all countries/economies participating in PISA 2015. Members of the OECD are highlighted in **bold**. + indicates limitations with the data meaning exclusion from the OECD report. * *China* refers to the four Chinese provinces that participated (Beijing, Guangdong, Jiangsu and Shanghai).

5. There are a number of differences between PISA 2015 and previous cycles. First, PISA 2015 was a computer-based assessment (CBA). This is in contrast to the five PISA cycles that took place between 2000 and 2012, which were all paper-based tests. Second, science was the focus of the PISA 2015 study, having last been the focus in 2006[6]. Finally, in 2015 a new 'collaborative problem solving' domain was added to the PISA assessment[7].

[6] Reading was the focus of PISA 2009, and mathematics in 2012.
[7] The results for collaborative problem solving will be released by the OECD in 2017, and are therefore not covered in this report.

6. This chapter introduces PISA 2015 and our analyses of the data for Wales. It does so by addressing the following questions:

- *What is the policy background to this report?*
- *What data were collected as part of PISA 2015, and how?*
- *Have there been any methodological changes since the last PISA cycle?*
- *What can PISA tell us? (And what can it not tell us?)*
- *How will the rest of the report be structured?*

7. All analyses presented within this report are correct as of the data received by the 4th November 2016. Updates to this national report may follow, subject to any data revisions or further analyses conducted by the authors or the OECD.

1.1 What is the policy background to this report?

8. Wales has 22 local authorities and following inspection by Estyn during 2010 and 2013, significant concerns were raised about the school improvement capacity of a significant minority of these. In response, the National Model for Regional Working was initially developed and agreed in autumn 2013 in order to accelerate changes already underway. It signalled a deeper commitment to regional working and emphasised a model of school improvement, based on mutual support that was largely new across most of Wales. Following the publication of the review of the Welsh education system by the OECD in 2014, the Welsh Government has developed and embedded a more rigorous approach to accountability in the schools system through the National School Categorisation system for primary and secondary schools. It enables direct investment in supporting and challenging schools through a more strategic, sophisticated and targeted way. The commitment to regional consortia as the central pillars of the school improvement system within Wales remains. It is intended that the consortia will play a key role in taking forward the significant developments in the Welsh education system over the coming years in driving up standards for all learners.

9. Prior to the 2013/14 OECD review, the Welsh Government had begun to place a strong focus on literacy and numeracy. As part of this focus, the Welsh Government introduced the National Literacy and Numeracy Framework in September 2013 as a curriculum requirement and as an assessment requirement in September 2014. Annual Reading and Numeracy Tests were also introduced between 2013 and 2014. The numerical reasoning tests were particularly innovative, challenging learners' ability to make judgements on the most efficient ways to resolve numerical problems. We know that schools are increasingly using the test

diagnostic tools and using the data they produce to identify learners' strengths and areas for development. The online adaptive tests that the Welsh Government will introduce from 2018 will provide more sophisticated diagnostic information, allowing schools to pinpoint specific gaps in learning and to intervene where necessary.

10. The Welsh Government commissioned an independent review of its curriculum and the recommendations were published in January 2013. The Welsh Government has begun developing the new curriculum in collaboration with the sector; it is making major changes to Initial Teacher Training and working to support teaching and learning through developing a new approach to continuous professional development.

1.2 What data have been collected as part of PISA 2015?

11. The main component of PISA is a two hour test, where participating school pupils across the world are assessed in their ability to address 'real life' challenges involving reading, mathematics and science. PISA is therefore a measure of young people's 'functional competence' in these academic domains. This differentiates PISA from other international pupil assessments, such as the Trends in International Mathematics and Science Study (TIMSS), which aims to measure knowledge of particular curriculum content areas. (The most recent TIMSS study also took place in 2015, with the results published in November 2016[8]). It is also one of the differences between PISA and the General Certificate of Secondary Education (GCSE) exams – see Box 1.1 for further information.

12. The aim of this report is to provide a first insight into how young people in Wales performed on the PISA science, reading and mathematics assessment in 2015. This includes comparing scores achieved by pupils in Wales to their peers in other countries, and investigating differences between groups of pupils and schools within Wales by a set of key characteristics.

13. In addition to the PISA test, 15-year-olds in all participating countries completed the PISA 'pupil questionnaire'. This asked young people to provide detailed information about their economic and social background, attitude towards school, out-of-school activities and life satisfaction. By using data from these questionnaires, this report will also provide an analysis of 15-year-olds' perceptions of teaching practice in their schools, and their aspirations and expectations for the future.

[8] Wales did not participate in TIMSS 2015.

Box 1.1 Differences between PISA and GCSEs

PISA tests young people's skills in reading, mathematics and science; subjects that are also assessed in General Certificate of Secondary Education (GCSE) exams. Although there is a strong correlation between young people's PISA scores and GCSE grades[9], there are also important differences in terms of patterns of pupil performance[10]. In this box, we describe some of the key differences between PISA and GCSEs:

Type of skill assessed: Whereas GCSEs examine pupils' knowledge of national curricula, PISA attempts to measure young people's 'functional skills' – their ability to apply knowledge to solve problems in real world situations.

Timing: In Wales, the PISA tests were sat in November/December 2015. This is six months before GCSE exams, which were taken in May/June 2016.

Test administration mode: Whereas the PISA 2015 tests were all completed on computer, GCSEs continue to be paper-based examinations.

Question style: Previous analysis of the PISA test questions found that they typically require a greater amount of reading than GCSE exams (NFER 2006), particularly in science.

Stakes: PISA is a 'low stakes' test for pupils; they do not receive any feedback about their performance and have little riding upon the results. In contrast, GCSEs are 'high stakes' exams, with all pupils receiving a grade that potentially has an impact upon their future educational options and career.

Language versions: In GCSE examinations, pupils have the opportunity to see both Welsh and English versions of the test paper. This is not the case in PISA, where pupils were only presented the test in either English or Welsh, as chosen by the pupil.

[9] Micklewright and Schnepf (2006).
[10] Jerrim and Wyness (2016).

14. In all countries, headteachers of participating schools were also asked to complete a background questionnaire. This included questions regarding school resources, quality assurance processes, perceived barriers to learning and the impact of school inspections. Analysis of these data will also be presented within this report (see chapter 8).

15. The data for the PISA 2015 study in Wales has been augmented in two ways. First, each country is allowed to add up to five questions to the pupil background questionnaire. The National Centre took up this option, adding a set of questions asking young people about their plans regarding higher education. This included the likelihood of applying to university, names of universities to which they may apply, the factors that will be important to them when selecting a university, and with whom they have discussed their plans regarding higher education. The resulting data are analysed as part of chapter 9.

16. Second, the PISA 2015 data for Wales has been linked to national administrative records. At the school level this includes information on type of school (e.g. Welsh versus English medium), the percentage of pupils who are eligible for Free School Meals (FSM) and National School Categorisation band. At the pupil level, young people's PISA scores and survey responses have been linked to information from the Welsh pupil annual school census. This includes data on pupils' English/Welsh medium status and eligibility for FSM. The inclusion of this information allows for a richer analysis of the PISA data for Wales than would otherwise be possible.

1.3 How was the PISA 2015 sample recruited in Wales? And how representative is it of the population?

17. PISA 2015 collected information from 140 schools and 3,451 pupils in Wales. These numbers reflect official response rates in Wales of 92 per cent for schools and 88 per cent for pupils, exceeding the strict minimum response rates required by the OECD[11].

[11] The OECD requirements stipulate that the school-level response rate is at least 85 per cent, and that at least 80 per cent of selected pupils participate in the study within selected schools. School level response rate reported after replacement schools included. See Appendix B for further details.

18. PISA was conducted in Wales during November and December 2015. These dates were chosen in order to avoid a clash with national GCSE assessments and to reduce the burden on participating schools. Rather than an assessment of all pupils aged 15 in each country, a two stage survey design is used to select schools and pupils to take part in the study.

Table 1.2 The sample of schools participating in PISA 2015 in Wales

	Initial sampled schools	Final participating schools
% of FSM eligible pupils (mean)	18%	18%
% attendance during year (mean)	94%	94%
% Pupils achieving the level 2 threshold including English/Welsh and Maths (mean)	59%	60%
Key Stage 4 (capped) average point score (mean)	353	353
Medium of instruction		
English	78%	80%
Welsh	17%	17%
Unknown / Not applicable	5%	3%
School support category		
Green	19%	21%
Yellow	39%	39%
Amber	29%	29%
Red	9%	9%
Unknown / Not applicable	5%	3%
Total number of schools	**152**	**140**

Source: PISA 2015 database.

Notes: Figures based upon unweighted data, and reported only for those schools where the relevant piece of information is available.

19. Schools in Wales were randomly selected to be representative of the national distributions of school type and location. Table 1.2 provides further information on the schools included in the PISA sample. Specifically, it compares school-level characteristics of the 152 schools initially selected to participate in the PISA study to the 140 who eventually took part. Summary statistics are provided for the percentage of pupils in each school who are eligible for FSM, who achieved the level 2 threshold in GCSEs (including English/Welsh and mathematics) and the school average Key Stage 4 capped points score. The distribution of English/Welsh medium schools and school National Support Category is also shown. Overall, the achieved PISA 2015 sample is very similar to the initially selected sample at the school level. However, as there are only 24 Welsh medium schools in the PISA 2015 sample, estimates for this

particular type of school will be accompanied by relatively large margins of error. The same caveat applies to schools within the ‘red’ school support category (12 schools).

20. Within each participating school, a simple random sample of 30 pupils, who met the PISA age definition, were selected to participate[12]. In Wales, this meant an initially selected sample of 4,179 pupils from within the participating schools. A total of 3,451 of these pupils completed the PISA assessment, with 473 pupils absent on the day of the test, 275 pupils excluded from the sample (primarily due to Special Educational Needs[13]) while 40 pupils were ineligible as they did not meet the PISA population definition.

21. Table 1.5 compares the background characteristics of three nested groups of pupils:

- Column 1 = The 3,451 pupils who completed the PISA assessment
- Column 2 = The 3,924 pupils who either completed the PISA assessment or were absent on the day of the test
- Column 3 = All 4,239 initially selected pupils (including those who were eventually excluded or deemed ineligible)

22. Overall, there is relatively little difference in the distribution of pupil characteristics across the three groups. For instance, 17 per cent of pupils who completed the PISA test were taught Welsh as a first language in school. This figure falls only slightly, to 16 per cent, once those pupils who were absent on the test day are also included. Similar findings hold for Free School Meal eligibility (13 per cent versus 15 per cent), Special Educational Needs (17 per cent versus 18 per cent) and gender (51 per cent male across all groups). Table 1.3 therefore indicates that the 3,451 pupils who completed the PISA test are similar to the initially selected sample in terms of observable characteristics.

23. For many of the demographic groups presented in Table 1.3, sample sizes are relatively small. For instance, only 575 of the pupils who completed the PISA test were taught Welsh as a first language in school. Similarly, a total of 445 pupils who took part in PISA were eligible for Free School Meals (FSM). There will consequently

[12] Further details on this process can be found in Appendix B.
[13] In PISA, all countries attempt to maximise the coverage of 15-year-olds enrolled in education in their national samples. The sampling standards permit countries to exclude up to five per cent of the relevant population, for reasons such as Special Educational Needs. Of the 275 pupils excluded from the PISA sample in Wales, 71 per cent had a Special Educational Need.

be quite a large degree of sampling error in the results reported for these particular sub-groups.

Table 1.3 The sample participating in PISA 2015 in Wales

	(1)	(2)	(3)
	Assessed	**Assessed + absent**	**Assessed + absent + ineligible + excluded**
FSM eligible			
No	85%	83%	82%
Yes	13%	15%	16%
Missing data	3%	2%	2%
Study in Welsh language			
No	80%	81%	80%
Yes	17%	16%	16%
Missing data	3%	3%	4%
Gender			
Female	49%	49%	49%
Male	51%	51%	51%
SEN			
No	81%	79%	76%
Yes	17%	18%	22%
Missing data	3%	2%	2%
School traffic light			
Green	22%	21%	21%
Yellow	39%	38%	38%
Amber	29%	29%	29%
Red	8%	9%	9%
Missing data	3%	2%	2%
Total number of pupils	**3,451**	**3,924**	**4,239**

Source: PISA 2015 matched database.

Notes: Figures based upon unweighted data. Figures may not sum to 100 per cent due to rounding.

24. Although the PISA 2015 data for Wales is representative of the target population, the fact that it is based upon a sample (rather than a census) means there will be a degree of uncertainty in all results. It is therefore important that this uncertainty is reflected within our statistical analysis. This is done in two ways. First, 95 per cent confidence intervals will be presented within many of the graphs (represented using a thin black line). These refer to an upper and lower bound of the impact sampling error is likely to have upon the estimate[14]. Alternatively, we will state

[14] If one were to repeat the PISA sampling process 100 times, one would expect any given estimate for Wales to fall between the upper and lower confidence band on 95 occasions.

whether a difference is 'statistically significant' or not at the five per cent level. This simply means that the difference found (e.g. in average PISA scores between two countries) is unlikely to be due to PISA being based upon a sample from the target population, rather than a census. Note that 'statistical significance' does *not* mean a difference is big, or necessarily of substantive importance. Indeed, in large samples such as PISA, even quite small differences can reach statistical significance. Rather, such terms are used throughout this report to describe the likely impact of sampling error alone.

25. The complex survey and test design of PISA makes accurate estimation of standard errors, confidence intervals, and statistical significance tests non-trivial. Throughout this report we use the 'repest' package developed by analysts from the OECD (Avvisati and Keslair 2014) and implemented within the statistics package Stata.

1.4 Have there been any important changes since the last PISA wave?

26. A number of changes have been made to PISA in 2015. For instance, the main study used computer-based assessment (CBA), instead of the more traditional paper-based assessment (PBA), for the first time. Moreover, as PISA 2015 focussed upon science performance, a greater number of assessment items tested 15-year-olds' competence in science than in reading or mathematics. New, interactive science questions have also been introduced, while there have also been some changes to how test questions have been scored and converted into the PISA proficiency scales. Finally, pupils' collaborative problem solving skills were tested for the first time within the PISA assessment.

27. There are three main implications of science being the focus of PISA 2015. First, the assessment included a greater number of science test questions than in the previous two cycles (when mathematics and reading were the focus of the study). School pupils' science skills are therefore measured with greater precision in PISA 2015 than in previous cycles as a result. Second, a more detailed analysis of 15-year-olds' science competency is possible. This includes a breakdown of science performance by 'cognitive' (how well pupils have mastered science skills) and 'content' (knowledge of particular scientific phenomena) domains. Finally, as the background questionnaires also focused upon science, a more detailed analysis of young people's attitudes, expectations and beliefs about science is possible than in either 2009 or 2012.

28. The change to CBA offers a number of administrative advantages, including efficiencies in marking, the introduction of new interactive questions, and the provision of additional information on the techniques young people use to answer test items. The change also, however, introduces a challenge in comparing performance measured by CBA with performance measured by paper-based assessment. This includes comparisons of PISA test scores across cycles, and between countries who conducted the PISA 2015 assessment on computer to those that conducted the 2015 assessment on paper. (A total of 15 countries participating in PISA 2015 continued to use paper-based assessment)[15]. The performance measure may, for example, be impacted by changes to the administration of the test, or the ways in which pupils interact with the assessment items.

29. To adjust for the change in test administration mode, ensuring PISA 2015 scores are comparable with the scale established for the paper-based assessment, the OECD have used test questions that are not subject to large mode differences as the basis of linking PISA 2015 scores to those from previous cycles. Further details on this methodology are available from the OECD in the annex of their international PISA 2015 report (see https://www.oecd.org/pisa/keyfindings/).

30. A number of other technical aspects of the PISA study have changed in 2015 from previous rounds. These include an increase in the number of 'trend' items included in the test, alterations to the statistical model used to scale the PISA scores and changes to how test questions not reached by pupils are treated. These factors could also potentially lead to changes in the pattern of results from previous cycles. Further details regarding these changes have been provided by the OECD in the annex of the international PISA 2015 report.

31. Finally, in May 2015 an error was identified in the layout of the PISA 2012 pupil questionnaire administered in the Welsh language. The error was not large enough to have a detectable impact on the UK's PISA 2012 results. However, it does have a small impact on estimates of overall scores and gender differences for Wales, Northern Ireland and England. As the impact is only small, this report uses the original PISA 2012 results. Appendix F provides a more detailed description of the error and the revised estimates as published by the OECD in May 2015.

[15] These countries are Albania, Algeria, Argentina, Georgia, Indonesia, Jordan, Kazakhstan, Kosovo, Lebanon, Macedonia, Malta, Moldova, Romania, Trinidad and Tobago, and Vietnam.

1.5 What can PISA tell us? (And what can it not tell us?)

32. PISA provides comparative evidence on the ‘functional ability’ of 15-year-olds in key academic areas. It allows one to describe the distribution of 15-year-olds’ competence in the particular subjects that PISA tests, how this compares to young people in other countries, and how such skills vary by demographic group. For instance, PISA can be used to address questions such as ‘how big is the achievement gap between Wales and the highest performing countries’ and ‘is the relationship between socio-economic status and achievement stronger in Wales than in other members of the OECD’?

33. PISA can also be used to establish the *correlation* between academic achievement and a range of potential explanatory factors. This includes young people’s attitudes, expectations and beliefs, school-level factors (e.g. school resources and management strategies) and system-level characteristics (e.g. amount of school autonomy). It is therefore a useful benchmarking tool that can help teachers, schools and policymakers understand the relative strengths and weaknesses of young people at a particular point in their development.

34. Increasingly, PISA is also providing important contextual information about other aspects of young people’s lives. For instance, in addition to testing pupils’ skills, PISA 2015 also includes data on their ambitions, anxieties, social interactions, and life satisfaction. It can therefore assist our understanding of young people’s well-being in other important dimensions beyond school. Together, this can direct government and educators towards the areas and groups in the most need of assistance.

35. Despite these strengths, PISA also has limitations. It is therefore important to clearly state what these data, and the analysis presented in this report, can and cannot reveal.

36. First, PISA is a cross-sectional survey, providing a snapshot of pupils’ skills at one point in time. It therefore does not provide any information about the *progress* young people make during their time at school. In other words, PISA does not measure the value-added of schools (or school-systems). Consequently, it is not possible to establish whether secondary schools in any particular country (e.g. Wales) facilitate more academic progress than others (e.g. Canada, Switzerland, the Netherlands).

37. Second, PISA scores are the culmination of all the factors influencing 15-year-old pupils' skills throughout their early life. This will include schools (both primary and secondary) and government education policy. Yet it will also encompass the time and monetary investments made by parents, young people's attitudes and motivation, early lifetime conditions e.g. attending pre-school, macroeconomic forces (e.g. economic prosperity, inequality) and a host of other factors. Consequently, it is not appropriate to treat PISA as a direct indicator of the 'quality' of schools in Wales. Moreover, due to the host of factors influencing pupils' test scores, some of which cannot be observed within the data, PISA can typically only identify correlations between variables, rather than establishing cause and effect. However, what PISA can provide is a descriptive account of how the distribution of 15-year-old pupils' skills vary by school-level characteristics (e.g. by school type). It also provides contextual information on issues such as school organisation and administration.

38. Finally, PISA scores can increase or decrease for many substantive reasons. It is therefore not possible to attribute change in a country's performance as direct evidence for or against any particular national policy (or set of policies). Changes in PISA 2015 results for Wales from previous cycles should therefore not be taken as providing evidence as to the impact of any previous or ongoing educational reform.

1.6 How will the rest of the report be structured?

39. The remainder of this report will be structured as follows. Chapters 2 to 5 will focus upon comparisons of Welsh pupils' performance in the PISA science, mathematics and reading assessment. As science was the focus of PISA 2015, a detailed comparison of performance across content and cognitive domains will be presented for this particular subject in chapter 3. Each chapter includes information on the distribution of pupils' PISA test scores, an overview of how average performance in Wales has changed over time[16], and how this compares to a selection of other countries.

40. Chapter 6 then moves to the association between PISA test scores and key demographic characteristics. We start by providing separate PISA score estimates in Wales for boys and girls, between pupils from advantaged and disadvantaged backgrounds, and examining the size of the gender and disadvantage gaps in Wales compared to other countries. The latter half of the chapter focuses specifically upon

[16] Although the PISA study began in 2000, Wales did not participate as a separate benchmarking country until 2006. Moreover, the UK did not meet the strict data requirements of the OECD in the first two PISA waves (2000 and 2003). Comparisons of PISA scores for Wales can therefore not be made before 2006.

variation between groups of pupils within Wales, including differences between young people who completed the test in English versus Welsh.

41. In chapter 7, we turn to differences in performance within Wales at the school level. Following the structure of previous chapters, it focuses upon average PISA test scores, and how this varies according to a set of school characteristics. This includes Welsh/English medium status and National Categorisation System band.

42. Chapter 8 focuses upon the views of headteachers in Wales, as captured by their responses to the PISA school questionnaire. This includes an analysis of headteachers' management styles, the factors that they believe to be hindering instruction within their school, and if they feel that their school is adequately resourced. The views of headteachers in Wales are first compared to the views of headteachers in other countries, in order to provide an international comparative context for the results. We then explore variation in headteachers' responses within Wales, focusing upon differences between those leading schools in different National Categorisation System bands, and between Welsh/English medium schools. In doing so, chapter 8 will highlight what headteachers in Wales believe to be the most significant barriers to learning within their schools.

43. A host of previous research has illustrated the important role young people's aspirations play in shaping their future[17]. Chapter 9 therefore investigates the aspirations and expectations of 15-year-olds in Wales, and how this compares to the aspirations of young people in other parts of the world. As science is the focus of PISA 2015, particular attention is paid to the proportion of young people in Wales who aspire to a Science, Technology, Engineering and Mathematics (STEM) career, and the extent to which they believe that their school science lessons are relevant for their educational and occupational future. We also investigate 15-year-olds' plans regarding higher education, including the proportion who believe they will obtain at least an undergraduate degree, and the institution they hope to attend. For each of these topics, the situation in Wales is first placed into an international comparative context, before further investigation of within-country differences between certain demographic groups (including gender and socio-economic status).

44. Further investigation of pupils' responses to the PISA background questionnaire follows in chapter 10, though now with an emphasis upon how they view science teaching within their school. Wales is first compared internationally in

[17] Morgan (2005).

terms of the frequency different learning activities occur within science lessons, and the amount of feedback young people receive about their performance. Attention then turns to how much time 15-year-olds in Wales spend learning science each week compared to other subject areas, both inside and outside of school.

45. The final chapter focuses upon differences in PISA outcomes between the four constituent countries of the United Kingdom. This includes how PISA test scores vary across the UK, and whether gender and socio-economic gaps are bigger in certain parts of the UK than others. It concludes by exploring differences between England, Northern Ireland, Scotland and Wales in pupils' and headteachers' responses to the PISA background questionnaires. This includes whether there are differences in headteachers' views on the factors hindering instruction within their school, and in the amount of time 15-year-olds spend studying science compared to other subject areas.

Chapter 2. Achievement in science

- The average PISA science score in Wales is 485. This is significantly lower than the average last time science was the focus of PISA in 2006 (505).
- The average science score is more than 20 points higher than in Wales in 18 countries. Average PISA scores are between 10 and 20 points higher than in Wales in a further 11 countries.
- Wales has a similar proportion of low achieving pupils in science (22 per cent) as the average across members of the OECD (21 per cent). However, the proportion of 15-year-olds reaching the top two PISA levels is lower (five per cent in Wales versus eight per cent across the OECD).
- In science, the gap between the highest achieving pupils in Wales and the highest achieving pupils from other countries is particularly pronounced.
- The science skills of the highest achieving pupils in Wales have declined over the last decade.
- The gap between the highest and lowest achieving pupils in science is 235 test points (almost eight years of schooling). Although sizeable, this difference is smaller than in most other countries.

Box 2.1 Methods for interpreting differences between countries

1. **Country rankings**. This is where countries are ordered by the statistic of interest (e.g. average PISA scores). The position of one country in this ranking is then compared to another. Although easy to communicate, this approach is problematic for at least three reasons. First, as PISA is based upon a sample rather than a census, we cannot be certain about the exact position of any given country. Consequently, two identical countries could end up with quite different rank positions (e.g. 20th versus 30th) simply due to sampling error. Second, rank order provides no information about the size of the achievement gap between countries. Finally, the position of a country may change over time simply due to a change in the number (or selection) of countries taking part.
2. **'Statistically significant' differences**. One way to account for the fact PISA is based upon a sample is to report whether differences between countries are 'statistically significant'. A 'significant' difference between countries is then reported when we are almost certain that this is not the result of sampling error. This overcomes one limitation with the use of country rankings. However, it still reveals little about the magnitude of the difference between countries. Indeed, in large sample studies such as PISA, even relatively modest differences between countries can be reported as 'statistically significant'.
3. **Effect size differences.** Differences between countries can also be interpreted in terms of an effect size. This refers to differences between countries in terms of absolute magnitude. An advantage of this approach is that it retains some information about differences in achievement between Wales and any given country of interest. Moreover, in large samples such as PISA, effect size differences of important magnitude will also typically be statistically significant.

Throughout this report, a combination of the second and third methods listed above will be used. When reporting average PISA scores, countries will be divided into four groups, based upon the number of test points they are ahead or behind Wales. This will also be expressed in terms of 'months of schooling' differences, following the approximate rules of thumb presented in OECD (2010:110):

Group 1: Mean score at least 20 points (eight months of schooling) ahead of Wales.

Group 2: Mean score between 10 and 20 points (between four and eight months of schooling) ahead of Wales.

Group 3: Mean score within 10 points (four months of schooling) of Wales.

Group 4: Mean score at least 10 points (four months of schooling) below Wales.

A star (*) will then also be placed by any country with a mean score significantly higher or lower than Wales at the five per cent level.

2.1 What is the mean PISA science score in Wales, and how does this compare to other countries?

1. Scientific literacy matters as the world faces major challenges in providing sufficient water and food, controlling diseases, generating sufficient energy and adapting to climate change[18]. As the OECD states '*societies will therefore require a cadre of well-educated scientists to undertake the research and the scientific technological innovation that will be essential to meet the economic, social and environmental challenges which the world will face*'[19]. Ensuring sufficient scientific literacy amongst young people is also vital for Wales' economic prosperity, material well-being and growth[20]. Consequently, it is important to consider how the science proficiency of 15-year-olds in Wales compares to 15-year-olds elsewhere in the world. Table 2.1 therefore places average PISA science scores for Wales into an international context, with countries separated into one of four groups.

2. The mean PISA science score in Wales is 485. Panel (a) refers to those countries where average PISA science scores are at least 20 points higher. A metric occasionally used by the OECD (2010:110) equates differences of this magnitude to at least eight months (two terms) of additional schooling. A total of 18 countries belong to this group; including eight East Asian economies, seven European countries and three English-speaking members of the OECD (Australia, Canada and New Zealand).

3. Panel (b) of Table 2.1 turns to countries with average PISA science scores between 10 and 20 test points higher than Wales. According to the OECD (2010:110), this is broadly equivalent to a difference of between four and eight months (one and two school terms) of additional schooling. There are 11 countries within this group. The vast majority (10 out of the 11) are within Europe, and they include other parts of the UK (Northern Ireland and Scotland). The only non-European nation within this group is the United States, where the average PISA science score is 496.

4. Panel (c) includes all countries within 10 points of the mean science score in Wales. Differences of this magnitude are equivalent to less than four months (one term) of additional schooling, and generally not outside the range one would expect given sampling error[21]. A total of 10 countries are within this group (excluding

[18] UNEP (2012).
[19] OECD (2013d).
[20] World Bank (2003).
[21] Note that statistical significance, where one can largely rule out a difference between countries occurring due to sampling error, is indicated in Table 2.1 via a star next to the mean score.

Wales). This includes several nations within Eastern Europe, such as Croatia, Lithuania, Hungary, Latvia and the Czech Republic. Another notable inclusion within this group is Sweden, where the mean score is 493.

Table 2.1 Mean PISA 2015 science scores

(a) Countries more than 20 points ahead of Wales

Country	Mean	Country	Mean	Country	Mean
Singapore	**556***	**Canada**	**528***	**Slovenia**	**513***
Japan	**538***	**Vietnam**	**525***	**England**	**512***
Estonia	**534***	**Hong Kong**	**523***	**Australia**	**510***
Taiwan	**532***	**China**	**518***	**Germany**	**509***
Finland	**531***	**South Korea**	**516***	**Netherlands**	**509***
Macao	**529***	**New Zealand**	**513***	**Switzerland**	**506***

(b) Countries between 10 and 20 points ahead of Wales

Country	Mean	Country	Mean	Country	Mean
Ireland	**503***	**Portugal**	**501***	**United States**	**496***
Belgium	**502***	**Northern Ireland**	**500***	**Austria**	**495***
Denmark	**502***	**Norway**	**498***	**France**	**495***
Poland	**501***	**Scotland**	**497***		

(c) Countries within 10 points of Wales

Country	Mean	Country	Mean	Country	Mean
Sweden	**493***	Russia	487	**Hungary**	**477***
Czech Republic	**493***	Wales	485	**Lithuania**	**475***
Spain	493	Luxembourg	483	**Croatia**	**475***
Latvia	490	Italy	481		

(d) Countries between 10 and 20 points behind Wales

Country	Mean	Country	Mean
Iceland	**473***	**Malta**	**465***
Israel	**467***		

Source: PISA 2015 database.

Note: Bold font with * indicates mean score significantly different from Wales at the five per cent level. Table does not include countries with average science scores more than 20 points lower than in Wales.

5. The last panel of Table 2.1 (panel d) contains countries with average PISA science scores between 10 and 20 points below Wales. Hence average science skills of 15-year-olds within these nations are four to eight months (one to two terms)

of schooling behind young people in Wales. Three countries fall within this group, Iceland (473), Israel (467) and Malta (465).

6. It is important to note that Table 2.1 does not include any country with an average PISA science score more than 20 points below the score for Wales. Results have therefore not been presented for 28 countries, including some members of the OECD, such as Greece (455) and Slovakia (461). A full set of average PISA science scores, including all participating countries, is provided in the online data tables.

Key point

The average PISA science score in Wales is 485. There are 29 countries where the average science score is at least 10 test points higher than in Wales, and 31 countries where the average science score is at least 10 test points lower.

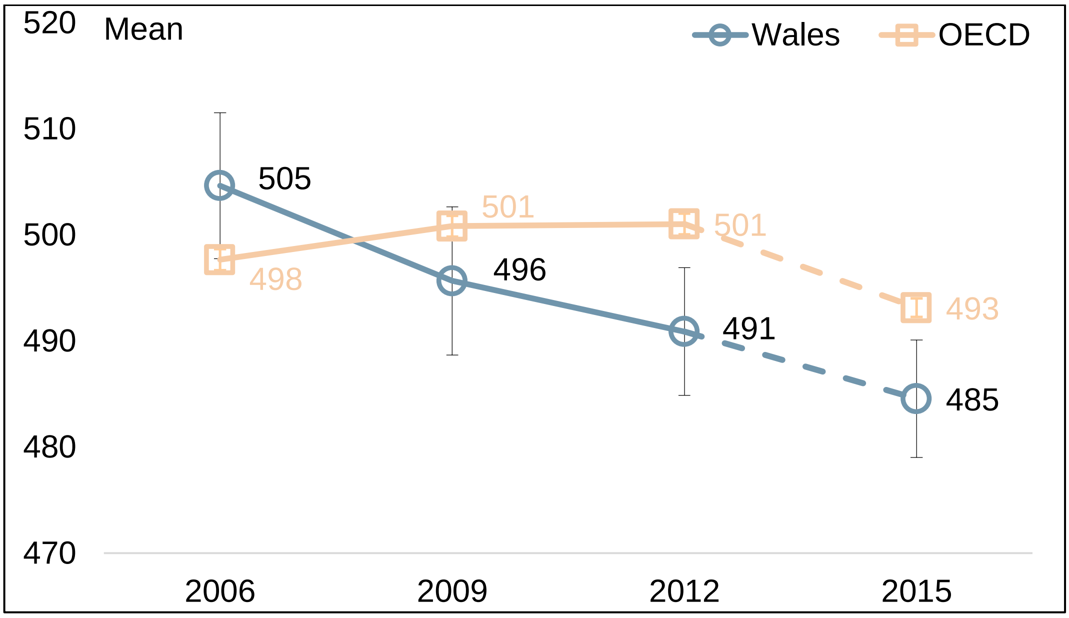
2.2 How have average PISA science scores in Wales changed over time? How does this compare to other countries?

7. The OECD has suggested that countries that manage to increase their average PISA test scores will see significant long-run improvements in their economic growth[22]. Moreover, as the previous sub-section illustrated, average science proficiency in Wales remains significantly behind some of the top-performing countries, indicating that there is room for improvement. This sub-section therefore turns to how the mean PISA score has changed since science was last the focus of PISA in 2006, and with respect to the last PISA wave conducted in 2012.

8. Figure 2.1 illustrates how the mean PISA science score in Wales has steadily declined over the past decade. Specifically, the mean has fallen from 505 in 2006, to 496 in 2009, 491 in 2012 and 485 in 2015. The difference between 2006 and 2015 is therefore 20 test points (approximately eight months of schooling) and is statistically significant at the five per cent threshold.

[22] OECD (2010:23).

Figure 2.1 Mean PISA science scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the 'AV09' results presented in the OECD international results Table I.2.4a. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

9. Table 2.2 compares the change for Wales to the five 'fastest improving' (red cells) and the five 'fastest declining' (blue cells) countries. In order to facilitate relevant comparisons, any country where the average PISA 2015 science score is below 450 points has been excluded from this table. Results are presented for both the change between 2006 and 2015 (panel a), and between 2012 and 2015 (panel b).

10. Starting with panel (a), Portugal has experienced the greatest improvement in mean science scores between 2006 to 2015, gaining approximately 27 PISA test points (moving from 474 to 501 on the PISA science scale). In contrast, Finland (-33 points, falling from 563 to 531) and Slovakia (-28 points, falling from 488 to 461) have suffered the most pronounced declines. It is notable how very few other countries have managed to substantially increase their average PISA science score over this period; Macao and Norway are the only other countries with a greater than 10 point (four months of schooling) improvement that is also statistically significant. In contrast, several other countries have seen a more than 20 test point (eight months of schooling) decline, such as Hungary and the Czech Republic. Indeed, countries with a mean PISA 2015 science score above 450 experienced, on average, a six point decrease in their average science score relative to 2006.

Table 2.2 The five fastest improving and the five fastest declining countries in science

(a) PISA 2006 to 2015

Country	From	To	Change
Portugal	474	501	**+27***
Macao	511	529	**+18***
Israel	454	467	+13
Norway	487	498	**+12***
United States	489	496	7
Czech Republic	513	493	**-20***
Wales	505	485	**-20***
Hungary	504	477	**-27***
Slovakia	488	461	**-28***
Finland	563	531	**-33***

(b) PISA 2012 to 2015

Country	From	To	Change
Portugal	489	501	**+12***
Taiwan	523	532	+9
Sweden	485	493	+9
Macao	521	529	+8
Singapore	551	556	+4
Ireland	522	503	**-19***
Lithuania	496	475	**-20***
South Korea	538	516	**-22***
Poland	526	501	**-24***
Hong Kong	555	523	**-32***

Source: PISA 2015 database.

Note: Figures refer to change between cycles in the mean PISA science score. Table restricted to only those countries with a mean score above 450 in the PISA 2015 science test. Bold font with a * indicates change statistically significant at the five per cent level. The difference between the ‘from’ and ‘to’ columns may not equal the ‘change’ column due to rounding. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

11. Panel (b) of Table 2.2 provides the analogous comparison between PISA 2012 and PISA 2015. A similar pattern emerges. There are very few countries where there is evidence of a substantial increase in mean science scores. On the other hand, the mean score has fallen by more than 20 test points (eight months of schooling) in several countries, including Hong Kong (-32 points from 555 to 523), Poland (-24 points from 526 to 501) and the Republic of Ireland (-19 points from 522

to 503). Indeed, the average country with a mean PISA 2015 science score above 450 points experienced a decline of around eight test points between 2012 and 2015.

Key point

There has been a statistically significant decline in the average PISA science score in Wales since 2006.

2.3 What proportion of pupils in Wales reach each science achievement level?

12. Although two countries may have similar average PISA science scores, there could be marked differences in terms of the distribution of pupils' performance. There may, for instance, be important differences between these countries in their share of 'top-performing' pupils and the proportion of 'low-achievers'. This matters from a policy perspective as a country's share of high-level skills is '*critical for the creation of new knowledge, technologies and innovation and therefore an important determinant of economic growth and social development*'[23]. Similarly, if a country has a large proportion of low achieving pupils, it suggests that the education system may not be equipping some young people with the basic science skills they need to function adequately in later life. This sub-section therefore focuses upon the proportion of 15-year-olds in Wales who reach each of the PISA science levels, with a particular focus upon the proportion of 'low-achievers' and 'top-performers'.

13. In order to describe the distribution of pupils' attainment, the OECD has divided the PISA science scale into different achievement levels. These range from Level 1b (very low levels of achievement) through to Level 6 (very high levels of achievement). Appendix D provides a description of these achievement levels, along with an explanation of the types of tasks to which they correspond. Throughout this report, 'low-achievers' refers to pupils scoring below PISA Level 2, while 'top-performers' score at PISA Level 5 or above.

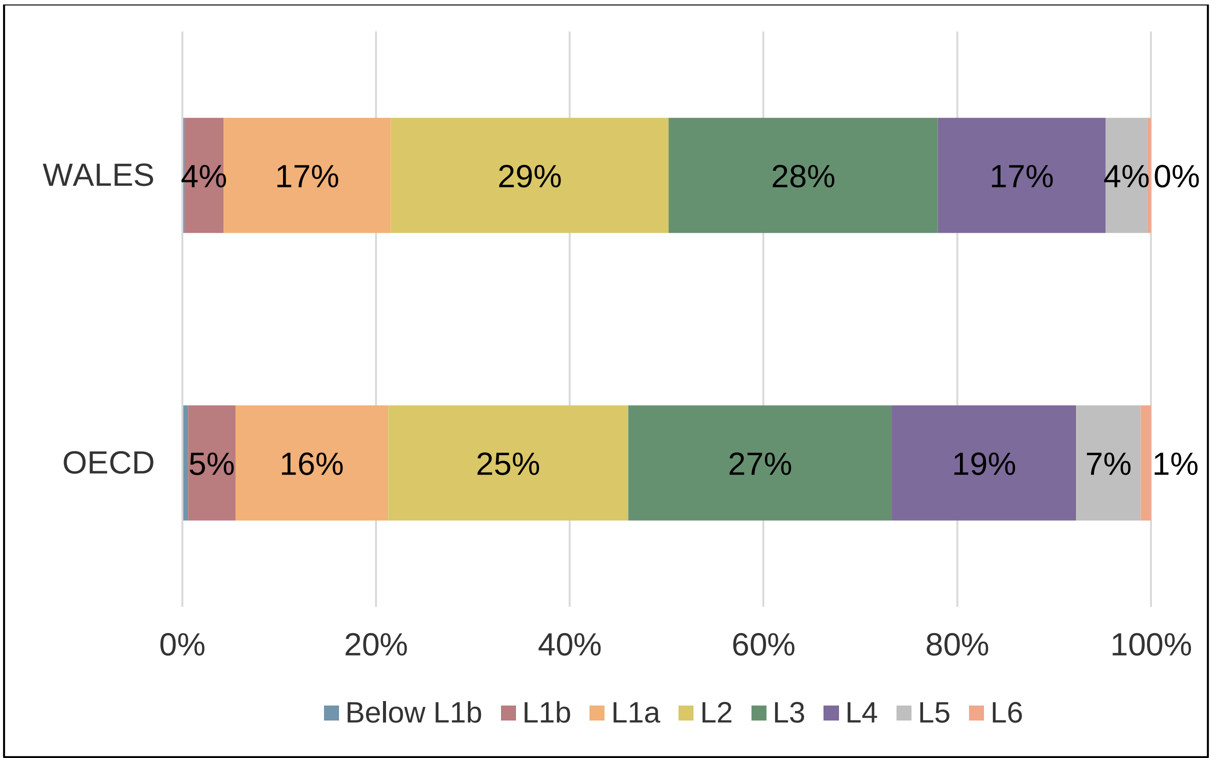
14. Figure 2.2 illustrates the proportion of pupils in Wales reaching each PISA science level, and compares this to the average across members of the OECD. In Wales, less than one per cent of 15-year-olds are below PISA science Level 1b, four per cent reach Level 1b and 17 per cent reach Level 1a. Analogous figures for the

[23] OECD (2009).

average across OECD members are one per cent (below Level 1b), five per cent (Level 1b) and 16 per cent (Level 1a). Therefore, the proportion of 'low-achievers' in Wales (22 per cent) is approximately the same as the average across members of the OECD (21 per cent).

15. However, at the other end of the distribution, Wales seems to have fewer high science achievers than the average member of the OECD. For instance, around one-in-twenty (five per cent) pupils in Wales reach one of the top two PISA science levels, compared to an OECD average of one-in-twelve (eight per cent). Consequently, it seems that the relatively low mean PISA science score in Wales is being driven by the fact that this country has relatively few high achieving pupils in this subject.

Figure 2.2 The percentage of pupils in Wales reaching each PISA science level

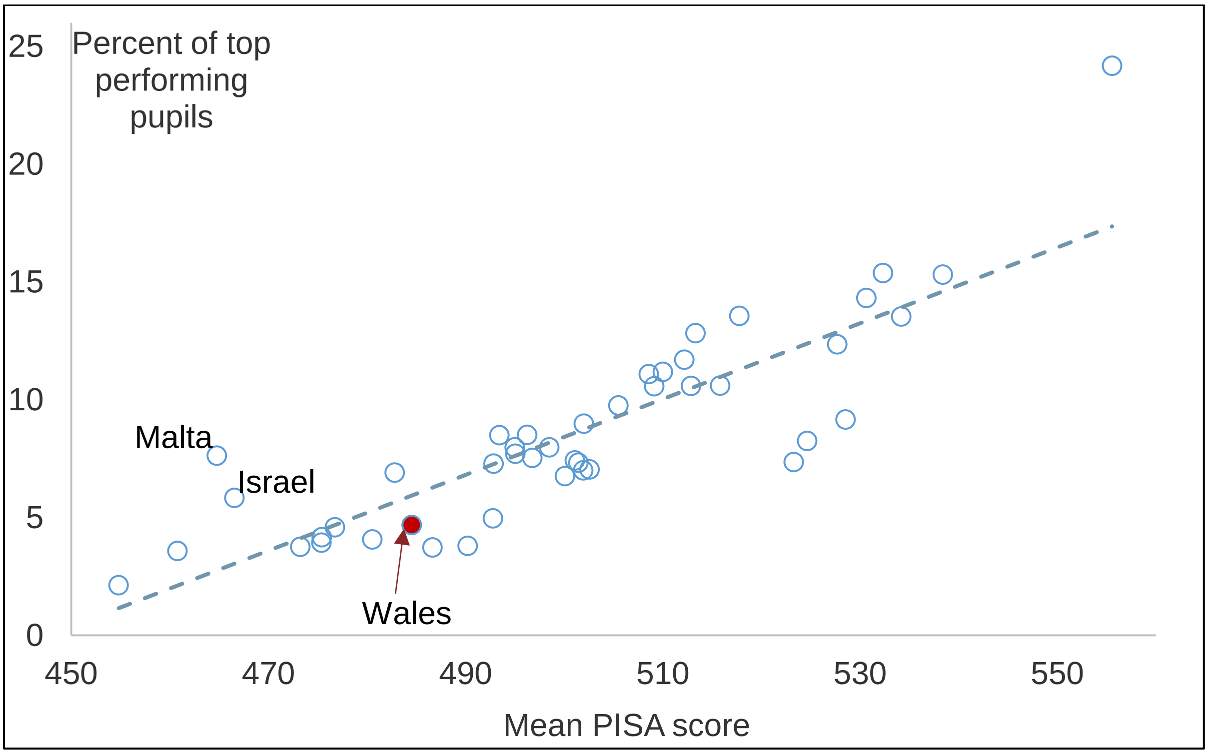


Source: PISA 2015 database.

16. Figure 2.3 provides further insight into how Wales compares to other countries in terms of the proportion of high-performing pupils. The horizontal axis plots the average PISA science score, while the vertical axis presents the proportion of pupils in each country achieving PISA Level 5 or Level 6. The dashed regression line then illustrates the cross-country relationship between these variables. In this figure, the sample of countries has been restricted to those with a mean science score above 450 points.

17. Wales sits below the dashed regression line; it is a country with a smaller proportion of high science achievers (five per cent) than one would expect given its mean score of 485. Indeed, it is particularly interesting to compare Wales in this respect to Malta and Israel. As Figure 2.3 illustrates, these countries have a significantly lower mean science score than Wales (465 and 467 respectively). Yet the proportion of pupils who reach PISA Level 5 or Level 6 in these countries is higher; six per cent in Israel and eight per cent in Malta. This again illustrates how Wales has a comparatively small proportion of 15-year-olds with high-level science skills, even in comparison to some other countries with lower average PISA science scores.

Figure 2.3 The percentage of top-performing science pupils compared to mean PISA science scores: a cross-country analysis



Source: PISA 2015 database.

Notes: The sample of countries included in this figure has been restricted to those with a mean science score above 450 points.

Key point

Wales has fewer high achieving pupils in science (five per cent) than the average across members of the OECD (eight per cent).

2.4 How do the science scores of the ***highest*** achieving pupils in Wales compare to other countries?

18. The previous sub-section highlighted how Wales has a smaller share of high-performing pupils in science than the average across members of the OECD. We now provide further insight into this issue by comparing the PISA test scores of the highest achieving Welsh pupils internationally, and considering how the performance of the highest achievers in science has changed over the last decade. Table 2.3 therefore presents the value of the 90th percentile of the science achievement distribution for Wales. As per section 2.1, countries have been divided into different groups depending upon how far ahead or behind Wales they are, but now in terms of the 90th percentile.

19. In PISA 2015, the 90th percentile of the science proficiency distribution in Wales was 602. This means that the top-performing 10 per cent of 15-year-olds in this country achieved a score of 602 test points or more. There are 21 countries where the 90th percentile is more than 20 points above the value for Wales, and a further 12 countries where the 90th percentile is between 10 and 20 points higher. In other words, the science skills of the top 10 per cent of 15-year-olds in Wales are significantly below those of the highest performing pupils in many other countries across the world. Nurturing high-level science skills therefore seems to be an area of weakness for Wales.

20. How have the science skills of the highest achieving Welsh pupils changed over time? Figure 2.4 provides the answer by plotting the 90th percentile of the PISA science distribution from 2006 to 2015, accompanied by the estimated 95 per cent confidence interval. There is evidence of a decline in this statistic since 2006 (the first time point to which one can compare). The 90th percentile of the science distribution stood at 638 in 2006. This has then steadily declined to 619 in 2009, 609 in 2012 and 602 in 2015. A difference between 2006 and 2015 of 36 test points and a sustained downward trend. Hence a key factor driving the declining mean science score in Wales over the past decade is the diminishing performance of the highest-achieving pupils.

Table 2.3 The 90th percentile of PISA 2015 science scores

(a) Countries more than 20 points ahead of Wales

Country	P90	Country	P90	Country	P90
Singapore	**683***	**Canada**	**644***	**Switzerland**	**632***
Taiwan	**655***	**England**	**642***	**Macao**	**630***
Japan	**655***	**Australia**	**639***	**Belgium**	**629***
Finland	**651***	**Netherlands**	**638***	**United States**	**626***
China	**649***	**Slovenia**	**636***	**Sweden**	**625***
Estonia	**648***	**Germany**	**636***	**Vietnam**	**624***
New Zealand	**647***	**South Korea**	**636***	**France**	**623***

(b) Countries between 10 and 20 points ahead of Wales

Country	P90	Country	P90	Country	P90
Hong Kong	**622***	**Scotland**	**619***	**Malta**	**618***
Norway	**622***	**Poland**	**619***	**Ireland**	**618***
Austria	**621***	**Northern Ireland**	**618***	**Denmark**	**617***
Portugal	**620***	**Czech Republic**	**618***	**Luxembourg**	**615***

(c) Countries within 10 points of Wales

Country	P90	Country	P90	Country	P90
Israel	606	Italy	599	Croatia	593
Spain	605	Lithuania	597	Iceland	593
Wales	602	Latvia	596		
Hungary	601	Russia	595		

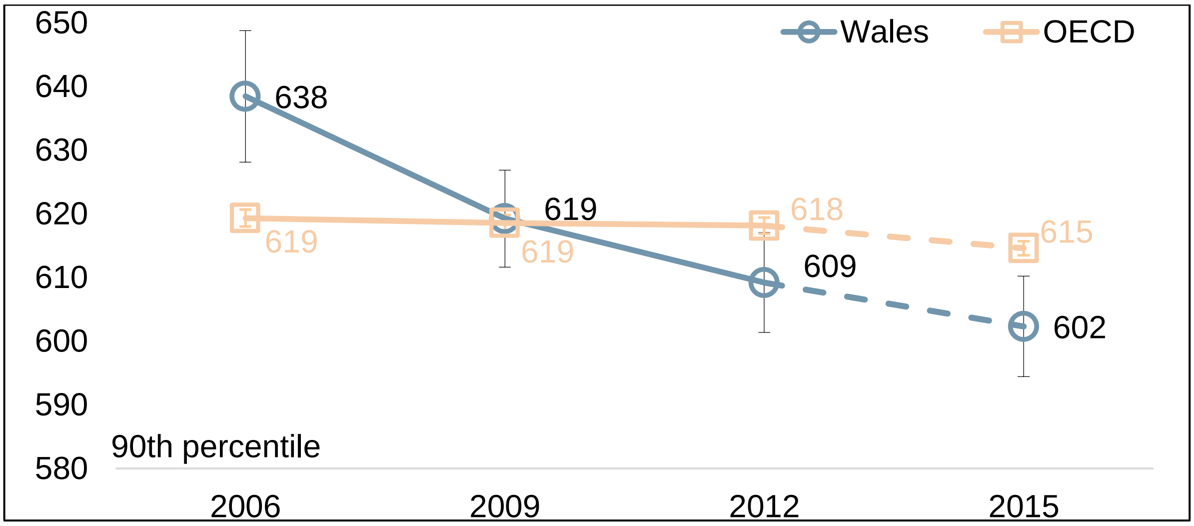
(d) Countries between 10 and 20 points behind Wales

Country	P90
Slovakia	**588***

Source: PISA 2015 database.

Note: Bold font with a * indicates significantly different from Wales at the five per cent level. Table does not include countries where the 90th percentile of the science proficiency distribution is more than 20 points below Wales.

Figure 2.4 The 90th percentile of PISA science scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. Confidence intervals do not include link error for comparing changes over time. OECD average based upon the 'AV09' results presented in the OECD international results Table I.2.4b. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

Key point

Nurturing 15-year-olds with high-level science skills is a particular challenge facing Wales. A key factor driving the decline in mean science scores in Wales over the past decade is a fall in the performance of the highest-achieving pupils.

2.5 How do the science scores of the *lowest* achieving pupils in Wales compare to other countries?

21. Although Wales may have a weakness in terms of the highest performing pupils in science, what do we know about the skills of the lowest achievers? Do their PISA scores also compare unfavourably relative to the least skilled 15-year-olds in other countries? Table 2.4 provides evidence on this matter. It does so by comparing the 10th percentile of the science proficiency distribution in Wales to other countries.

Table 2.4 The 10th percentile of PISA 2015 science scores

(a) Countries more than 20 points ahead of Wales

Country	P10	Country	P10	Country	P10
Vietnam	**428***	**Japan**	**412***	**Taiwan**	**395***
Macao	**420***	**Singapore**	**412***	**South Korea**	**388***
Estonia	**416***	**Canada**	**404***		
Hong Kong	**413***	**Finland**	**402***		

(b) Countries between 10 and 20 points ahead of Wales

Country	P10	Country	P10	Country	P10
Ireland	**387***	**Denmark**	**383***	**Russia**	**379***
Slovenia	**386***	**Latvia**	**382***	**Portugal**	**379***
Poland	**384***	**Northern Ireland**	**379***		

(c) Countries within 10 points of Wales

Country	P10	Country	P10	Country	P10
England	378	Scotland	372	Czech Republic	367
China	377	Netherlands	372	Austria	365
Germany	376	Australia	372	Belgium	364
New Zealand	374	Norway	370	Croatia	360
Spain	374	Wales	368	Italy	359
Switzerland	373	United States	368		

(d) Countries between 10 and 20 points behind Wales

Country	P10	Country	P10	Country	P10
Lithuania	**357***	**France**	**355***	**Luxembourg**	**351***
Sweden	357	**Iceland**	**354***		

Source: PISA 2015 database.

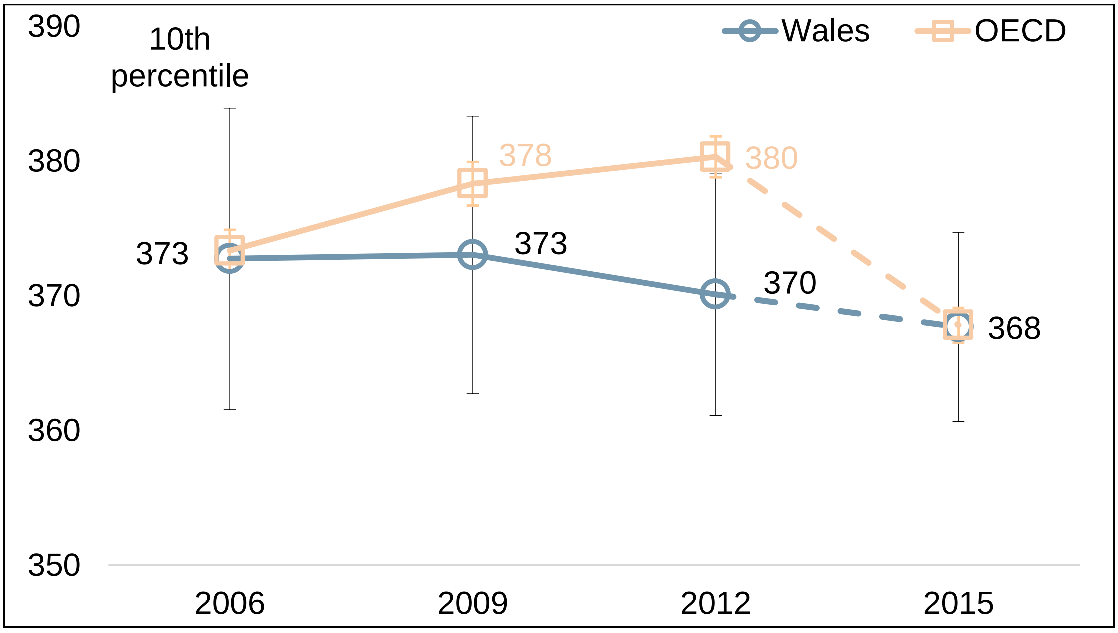
Note: * indicates significantly different from Wales at the five per cent level. Table does not include countries where the 10th percentile of the science proficiency distribution is more than 20 points below the value in Wales.

22. The value of the 10th percentile of the science proficiency distribution in Wales is 368. There are 10 countries where the 10th percentile is more than 20 points above the value for Wales, with seven of these within East Asia and just two from Europe (Finland and Estonia). In a further eight countries, the 10th percentile is between 10 and 20 points above Wales, with most of these being within Europe

(including Northern Ireland). Hence it seems that the situation towards the bottom end of the science achievement distribution for Wales is somewhat more favourable than the situation for the highest achievers (at least from an international comparative perspective).

23. Figure 2.5 proceeds by considering how the 10th percentile of PISA science scores in Wales has changed since 2006. The point estimate of the 10th percentile was very similar in 2006 (373), 2009 (373), 2012 (370) and 2015 (368). Moreover, there is little evidence of a clear trend over the past decade. Overall, Figure 2.5 therefore suggests that the science skills of the lowest-achieving pupils in Wales have remained broadly stable over time.

Figure 2.5 The 10th percentile of PISA science scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the 'AV09' results presented in the OECD international results Table I.2.4b. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

Key point

The skills of the lowest-achieving Welsh pupils in science have remained broadly unchanged between 2006 and 2015.

2.6 How big is the gap between the pupils with the strongest and weakest science skills? How does Wales compare to other countries in this respect?

24. Does Wales have an education system, society and culture that leads to large disparities in 15-year-olds science achievement? Or is this a country where there is a comparatively narrow gap between the highest and lowest performing pupils? The answer to this question matters because inequalities in education help to produce later lifetime disparities in a range of dimensions, including heath, well-being, occupational status and income[24]. This chapter therefore concludes by investigating whether the distance between the highest and lowest achieving pupils in Wales is greater than in other parts of the world.

25. To measure the gap between the highest and lowest performing pupils, we take the difference between the 10th and 90th percentiles of the PISA science achievement distribution within each country. This type of metric is commonly used to measure inequality in educational outcomes[25]. The magnitude of this gap is presented in Table 2.5. For brevity, the sample is restricted to only those countries with a mean PISA science score above 450 points. The 10 countries with the highest mean PISA science score have been highlighted.

26. The 90th percentile of the PISA science test score distribution in Wales is 602, while the 10th percentile stands at 368. Table 2.5 demonstrates that the gap is therefore 235 test score points, equivalent to almost eight years of schooling. Although this is a sizeable difference, it is smaller than in most other countries (the average across members of the OECD is 247). Indeed, in only five of the countries included in Table 2.5 is the difference between the 90th and 10th percentile significantly smaller than in Wales (three East Asian economies along with Russia and Latvia). Conversely, there are 24 countries where inequality in science achievement is significantly greater at the five per cent level. Consequently, by this metric, Wales has greater equality in 15-year-olds' science achievement than most other countries. However, as noted in the previous sub-sections of this chapter, this is being at least partly driven by the comparatively low science skills of Wales' highest achieving pupils.

[24] Micklewright and Schnepf (2006).
[25] Bruckauf and Chzhen (2016).

Table 2.5 Difference in PISA test points between the highest and lowest achievers in science 2015

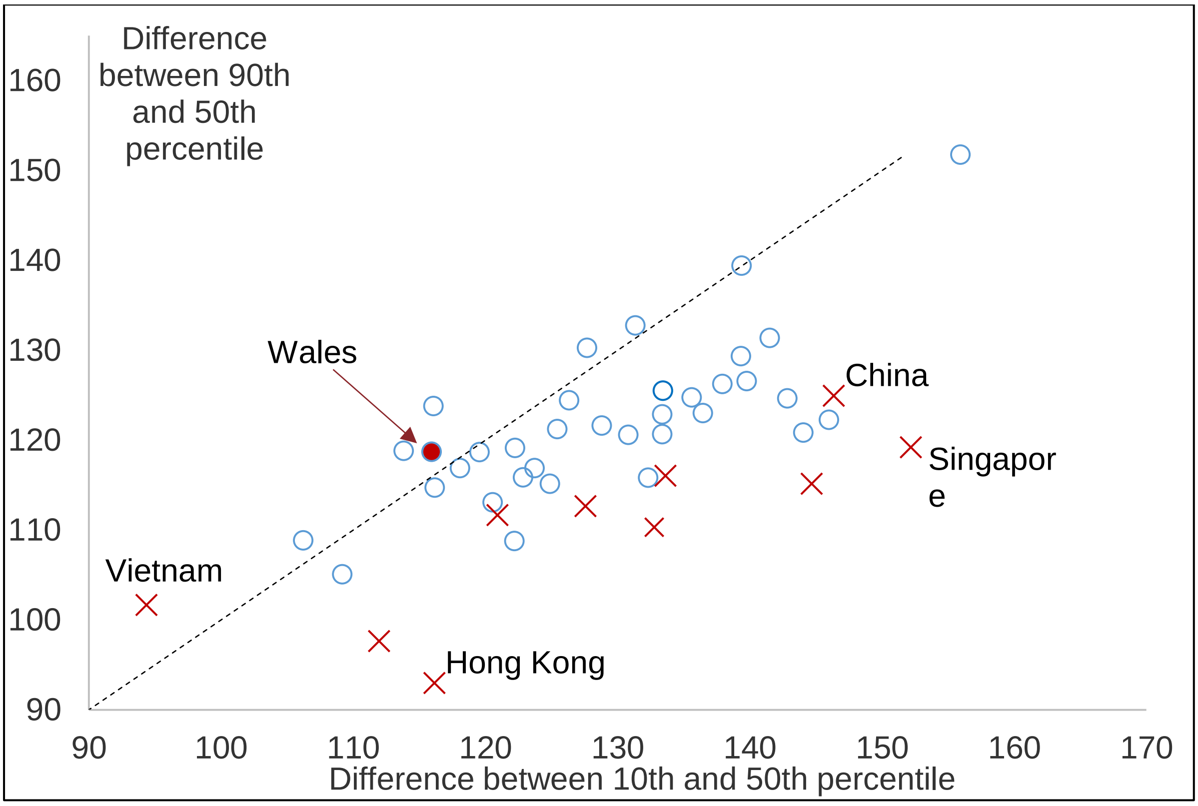
Country	Difference between the 90th and 10th percentile	Difference in years of schooling
Malta	**308***	**10.3 years**
Israel	**279***	**9.3 years**
New Zealand	**273***	**9.1 years**
Singapore	**271***	**9.0 years**
China	**271***	**9.0 years**
Sweden	**269***	**9.0 years**
France	**268***	**8.9 years**
Australia	**267***	**8.9 years**
Netherlands	**266***	**8.9 years**
Belgium	**265***	**8.8 years**
England	**264***	**8.8 years**
Luxembourg	**264***	**8.8 years**
Germany	**260***	**8.7 years**
Taiwan	**260***	**8.7 years**
Switzerland	**259***	**8.6 years**
Slovakia	**259***	**8.6 years**
United States	**258***	**8.6 years**
Austria	**256***	**8.5 years**
Hungary	**254***	**8.5 years**
Norway	**251***	**8.4 years**
Czech Republic	**251***	**8.4 years**
Slovenia	**250***	**8.3 years**
Finland	**250***	**8.3 years**
South Korea	**248***	**8.3 years**
Scotland	247	8.2 years
Japan	243	8.1 years
Greece	241	8.0 years
Portugal	241	8.0 years
Canada	240	8.0 years
Italy	240	8.0 years
Lithuania	240	8.0 years
Northern Ireland	239	8.0 years
Iceland	238	7.9 years
Poland	235	7.8 years
Wales	235	7.8 years
Denmark	234	7.8 years
Croatia	233	7.8 years
Estonia	233	7.8 years
Spain	231	7.7 years
Ireland	231	7.7 years
Russia	**215***	**7.2 years**
Latvia	**214***	**7.1 years**
Macao	**210***	**7.0 years**
Hong Kong	**209***	**7.0 years**
Vietnam	**196***	**6.5 years**

Source: PISA 2015 database.

Note: Bold font and * indicates statistically significant differences compared to Wales at the five per cent level. Table only includes countries where the mean PISA science score is above 450.

27. Figure 2.6 further explores the source of this educational inequality. The horizontal axis plots the difference between the median and the 10th percentile of the science test score distribution; the gap between the lowest achieving 10 per cent of pupils in each country and the average pupil. On the other hand, the vertical axis illustrates the difference between the median and the 90th percentile; the gap between the average pupil and the highest achieving 10 per cent within each country. This comparison therefore demonstrates whether inequality in pupils' skills is more pronounced in the bottom half of the science achievement distribution or the top half. Results have again been presented for only those countries with a mean science score above 450 points. The red crosses refer to the 10 countries with the highest mean PISA science scores ('H10').

Figure 2.6 A comparison of the 90th to 50th percentile and 50th to 10th percentile science achievement gap across countries



Source: PISA 2015 database.

Notes: Dashed diagonal line refers to where the difference between the 90th and 50th percentile is equal to the difference between the 10th and 50th percentile. Figure only includes countries and economies where the mean PISA science score is above 450. Red crosses refer to the 10 countries with the highest average PISA science score.

28. There are three important features of Figure 2.6. First, Wales is towards the left-hand side of this plot. This indicates that the difference between the 10th and 50th percentile is smaller in Wales than in most other countries. In other words, Wales stands out as a country with a comparatively small gap between the median pupil and the lowest achieving pupils. Second, the majority of countries sit below the 45 degree line. This illustrates that, in most countries, the gap between the lowest achieving pupils and the median pupil is bigger than the gap between the median pupil and the highest achievers. Wales is, however, an interesting exception to this rule, where the distance between the 50th and 10th percentile (116 test points) is roughly the same as the distance between the 90th and 50th percentile (119 test points). Finally, it is notable how patterns of educational inequality differ markedly between the 10 countries with the highest average PISA science scores. For instance, countries like Vietnam and Hong Kong sit in the bottom-left hand corner of Figure 2.6, with comparatively small differences between low, average and high achieving pupils. Conversely, there are countries like Singapore and China where inequality in achievement (particularly between low achieving and average pupils) is much greater. This illustrates how countries with the highest average PISA science scores differ markedly in terms of the distribution of performance.

Key point

The gap between the highest and lowest achieving pupils in science is smaller in Wales than in most other countries. The distance between low achieving and average achieving pupils in science is particularly small in Wales.

Chapter 3. Achievement in different aspects of scientific literacy

- PISA draws a distinction between different topics in science. These are the ‘physical system’ (which measures knowledge about matter, motion and forces), the ‘living system’ (which pertains to cells, organisms, humans), and the ‘earth and space science system’ (looking at earth’s history, the earth in space, and the universe).

- Pupils in Wales achieve equally as well across the ‘living’, ‘physical’ and ‘earth and space’ science systems in 2015. It is relatively common for a country to have equal scores across the three scientific systems – including in many of the high-achieving countries.

- The PISA 2015 test also examines skills in three core scientific competencies: ‘interpreting data and evidence scientifically’, ‘evaluating and designing scientific enquiry’ and ‘explaining phenomena scientifically’.

- Pupils in Wales are slightly stronger at explaining phenomena scientifically than they are at evaluating and designing scientific enquiry. This pattern of results is reasonably uncommon, and is not found in many of the highest performing countries. However, it should be noted that the magnitude of this difference in Wales is quite modest (around five PISA test points).

- The PISA test also attempts to measure separate types of scientific knowledge: ‘content knowledge’ and ‘procedural and epistemic knowledge’.

- Pupils in Wales are equally able in content knowledge and procedural and epistemic knowledge, which is not unusual compared to other countries. It is of note that in some of top-performing countries (e.g. Taiwan, Finland), the gap between content knowledge and procedural/epistemic knowledge is more pronounced.

1. In the previous chapter, our focus was pupils' overall achievement in the PISA science domain. However, proficiency in science is formed of several interlinking components, with the potential for 15-year-olds to have stronger skills in certain areas of this subject and weaker skills in others. For instance, do pupils in Wales have a particularly good understanding of one aspect of science (e.g. physics) but comparatively poor understanding of another (e.g. biology)? This chapter examines such issues by considering pupils' proficiency across the eight PISA science sub-domains.

2. The eight PISA science sub-domains have been divided into three broad groups:

Scientific systems (physical, living and earth and space sciences)

Scientific competencies (explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically)

Scientific knowledge (content knowledge, and procedural and epistemic knowledge)

3. The PISA 2015 test has been designed to allow comparisons to be made *within* these three broad groups. For example, average scores can be compared across physical and living science systems, or between content knowledge and procedural/epistemic knowledge. However, comparisons should not be made between sub-domains that fall within different groups; it is not possible to directly compare the mean score for the 'living system' to the mean score for the 'explaining phenomena scientifically' competency, for example. In order to provide a more detailed insight into the content of the PISA test, the latter half of the chapter turns to analysis of two exemplar science questions. This includes one of the new interactive test items that have been introduced into PISA as part of the move to computer-based assessment. We also provide some descriptive evidence on how pupils in Wales performed on these two tasks, relative to 15-year-olds in other parts of the world.

4. In summary, this chapter will address the following questions:

- *Do pupils have the same proficiency across the PISA 'physical', 'living' and 'earth and space' science systems? How does Wales compare to other countries in this respect?*
- *How do average PISA scores vary in Wales across three core scientific competencies: 'explaining phenomena scientifically', 'evaluating and designing scientific enquiry' and 'interpreting data and evidence scientifically'?*

- *How does pupils' knowledge of scientific content compare to their knowledge of scientific processes and procedures? Is this similar to the situation in other countries?*
- *What types of questions were pupils asked as part of the PISA science test? What proportion of pupils in Wales answered these exemplar items correctly?*

3.1 Do pupils have the same proficiency across the PISA physical, living and earth and space science systems?

5. Science is a broad term used to encapsulate many different topics. For instance, in the Welsh educational system, a clear distinction is made between specific areas such as physics, chemistry and biology, with pupils being able to complete separate GCSEs and A-Levels in these particular fields. PISA also draws a distinction between different topics in science, based upon the OECD definition of different scientific systems. These are the 'physical system', the 'living system', and the 'earth and space science system'. Details on the types of topics each of these covers can be found in Table 3.1, with further information available within the PISA 2015 science framework[26].

Table 3.1 Content of the PISA science 'systems'

Physical systems	Living systems	Earth and Space systems
Structure and properties of matter	Cells	Structures of the Earth
Chemical changes of matter	Organisms	Energy in the Earth
Motion and forces	Humans	Change in the Earth
Energy and its transformation	Populations	Earth's history
Interactions between energy and matter	Ecosystems	Earth in space
	Biosphere	The Universe

Source: OECD (2016:26)

[26] See OECD (2016).

Table 3.2 Average scores across the PISA 'scientific systems' sub-domains

Country	Physical	Living	Earth and Space
Singapore	**555***	**558***	**554***
Japan	**538***	**538***	**541***
Estonia	**535***	**532***	**539***
Taiwan	**531***	**532***	**534***
Finland	**534***	**527***	**534***
Macao	**533***	**524***	**533***
Canada	**527***	**528***	**529***
Vietnam	-	-	-
Hong Kong	**523***	**523***	**523***
China	**520***	**517***	**516***
South Korea	**517***	**511***	**521***
New Zealand	**515***	**512***	**513***
Slovenia	**514***	**512***	**514***
England	**512***	**512***	**513***
Australia	**511***	**510***	**509***
Germany	**505***	**509***	**512***
Netherlands	**511***	**503***	**513***
Switzerland	**503***	**506***	**508***
Ireland	**507***	**500***	**502***
Belgium	**499***	**503***	**503***
Denmark	**508***	**496***	**505***
Poland	**503***	**501***	**501***
Portugal	**499***	**503***	**500***
Northern Ireland	**501***	**498***	**498***
Norway	**503***	**494***	**499***
Scotland	**499***	**497***	**494***
United States	494	**498***	**496***
Austria	**497***	**492***	**497***
France	492	**496***	**496***
Sweden	**500***	488	**495***
Czech Republic	492	**493***	**493***
Spain	487	**493***	**496***
Latvia	490	489	**493***
Russia	488	483	489
Wales	486	482	485
Luxembourg	**478***	485	483
Italy	479	479	485
Hungary	481	**473***	**477***
Lithuania	478	476	**471***
Croatia	**472***	476	477
Iceland	**472***	476	**469***
Israel	**469***	**469***	**457***
Malta	-	-	-
Slovakia	**466***	**458***	**458***
Greece	**452***	**456***	**453***

Notes: Table only includes countries with an average score above 450 points on the overall PISA science scale. Green/red cells indicate where the mean score for the country is at least five points higher/lower than for the mean score for the 'living' system. Information on sub-domain scores is not available for Malta and Vietnam. Bold font and * indicates statistically significant difference from Wales.

6. In all three scientific systems, there are around 25 to 30 countries with statistically significant higher scores than Wales. This includes the high-performing East Asian nations (e.g. Singapore, Japan, Macao, Hong Kong), but also several European countries as well (e.g. England, Finland, Estonia and, in some domains, Germany and the Netherlands). Several other English-speaking nations also have higher average scores across the board than Wales, including Australia, New Zealand, Canada and all other parts of the UK. Further details are provided in Table 3.2. The mean score for the living system (482) in Wales is also very similar to the mean score for either the physical (486) or earth and space science (485) systems.

Key point

Pupils in Wales achieve similar scores across the three PISA scientific systems.

3.2 How do average scores vary in Wales across the three core scientific 'competencies' measured by PISA?

7. For pupils to be able to understand and engage in critical discussions about science, they need to be able to demonstrate proficiency in three separate areas. First, they need to be able to explain and understand key scientific phenomena; for example, how a microwave oven works or why it is possible to compress gasses but not liquids. Second, pupils must understand the key principles of scientific investigation, such as what things should be measured, or what variables should be controlled, so that accurate and precise data can be collected. Finally, pupils need to be able to interpret data and evidence scientifically, in order to reach appropriate conclusions. For instance, they should recognise that an article within a peer-reviewed academic journal is a more trustworthy source of scientific information than a newspaper report.

8. The PISA 2015 test examined pupils' skills in these three core scientific competencies. They can be summarised under the following headings:

Explaining phenomena scientifically. Pupils' ability to recall knowledge of a particular aspect of science and then use that knowledge to explain some phenomena (e.g. why antibiotics do not kill viruses). This includes the use of such knowledge to make predictions of what is likely to occur in a particular real-world situation.

Evaluate and design scientific enquiry. This captures pupils' ability to identify questions that could be explored in a scientific study, to propose ways of explaining a question using a rigorous scientific method and to evaluate the quality of scientific investigations that have been conducted. This could also include an evaluation of how scientists ensure reliability of data and the generalisability of their findings.

<u>*Interpret data and evidence scientifically*</u>. Pupils' ability to understand the strengths and limitations of a scientific investigation, and how the reliability of the evidence may vary depending upon the source. This captures young people's understanding of uncertainty in science, the quality assurance processes needed to ensure reliability and objectivity, and to distinguish arguments based upon evidence from other considerations.

A summary of the skills each of these core competencies encapsulates can be found in Table 3.3.

Table 3.3 The scientific competencies examined in the PISA 2015 assessment

Explain phenomena scientifically	Evaluate and design scientific enquiry	Interpret data and evidence scientifically
Recall and apply scientific knowledge	Identify questions explored in a scientific study	Transform data into different representations
Identify, use and generate explanatory models	Distinguish questions that could be explored scientifically	Analyse and interpret data to reach appropriate conclusions
Make and justify predictions	Propose and evaluate ways of exploring a question scientifically	Identify assumptions, evidence and reasoning in texts
Explain implications of scientific knowledge for society	Evaluate how scientists ensure reliability, objectivity and generalisability of data and explanations	Distinguish arguments based upon theory and evidence from other considerations
Offer explanatory hypotheses		Evaluate evidence from different sources (e.g. journals, newspapers)

Source: OECD (2016:24-26)

9. Pupils in Wales are slightly stronger at explaining phenomena scientifically (486) than at evaluating and designing scientific enquiry (481). This pattern of results is reasonably uncommon, and is not found in many other countries – see Table 3.4. This includes several of the highest-performers, such as Japan, Estonia, Canada and Hong Kong, where average scores across the three scientific competencies are broadly equal. Table 3.4 illustrates there are only a few exceptions to this pattern amongst the high-performers, such as Singapore (where pupils have a particular strength in evaluating and designing scientific enquiry), Taiwan and Macao (where pupils are weaker at evaluating and designing scientific enquiry). Within the UK, pupils in England and Northern Ireland have similar average scores in the three PISA competencies, while 15-year-olds in Scotland have weaker skills in interpreting data and evidence scientifically than the other two areas. Nevertheless, the overall message of Table 3.4 is that, in most countries, differences across the three scientific competencies are relatively modest (including Wales).

Table 3.4 Average scores for the scientific 'competencies' tested in PISA

Country	Explain phenomena scientifically	Evaluate and design scientific enquiry	Interpret data and evidence scientifically
Singapore	**553***	**560***	**556***
Japan	**539***	**536***	**541***
Estonia	**533***	**535***	**537***
Taiwan	**536***	**525***	**533***
Finland	**534***	**529***	**529***
Macao	**528***	**525***	**532***
Canada	**530***	**530***	**525***
Vietnam	-	-	-
Hong Kong	**524***	**524***	**521***
China	**520***	**517***	**516***
South Korea	**510***	**515***	**523***
New Zealand	**511***	**517***	**512***
Slovenia	**515***	**511***	**512***
England	**512***	**510***	**512***
Australia	**510***	**512***	**508***
Germany	**511***	**506***	**509***
Netherlands	**509***	**511***	**506***
Switzerland	**505***	**507***	**506***
Ireland	**505***	**500***	**500***
Belgium	**499***	**507***	**503***
Denmark	**502***	**504***	**500***
Poland	**501***	**502***	**501***
Portugal	**498***	**502***	**503***
Northern Ireland	**500***	**497***	**501***
Norway	**502***	**493***	**498***
Scotland	**498***	**498***	**493***
United States	492	**503***	**497***
Austria	**499***	488	**493***
France	488	**498***	**501***
Sweden	**498***	491	490
Czech Republic	**496***	486	**493***
Spain	**494***	489	**493***
Latvia	488	489	**494***
Russia	486	484	489
Wales	486	481	483
Luxembourg	482	479	486
Italy	481	477	482
Hungary	**478***	474	476
Lithuania	**478***	478	**471***
Croatia	**476***	473	476
Iceland	**468***	476	478
Israel	**463***	471	**467***
Malta	**-**	-	-
Slovakia	**464***	**457***	**459***
Greece	**454***	**453***	**454***

Notes: Table only includes countries with an average score above 450 points on the overall PISA science scale. Green/red cells indicate where the mean score for the country is at least five points higher/lower than the mean score for 'evaluating and designing scientific enquiry'. Information on sub-domain scores is not available for Malta and Vietnam. Bold font and * indicates statistically significant difference from Wales.

Key point

Pupils in Wales are slightly stronger at explaining phenomena scientifically than they are at evaluating and designing scientific enquiry. This pattern of results is reasonably uncommon, and is not found in many of the highest performing countries. However, it should be noted that the magnitude of this difference in Wales is quite modest (around five PISA test points).

3.3 How does pupils' knowledge of scientific content compare to their knowledge of scientific processes and procedures?

10. The PISA test attempts to measure three separate types of scientific knowledge, which together demonstrates pupils' understanding of the natural world. This not only includes knowledge of the science systems (as listed in Table 3.1), but also of the rigorous processes and procedures that must be applied in order to generate high quality evidence. It also encompasses how knowledge in science is built.

11. In PISA 2015, these three complementary forms of knowledge are reported on two separate sub-scales:

Content knowledge. Pupils' knowledge and understanding of the content of the physical, living and earth and space science systems.

Procedural and epistemic knowledge. Pupils' understanding of key concepts and procedures underpinning scientific methods, which are used to produce reliable and valid data. Those with such knowledge can explain, with examples, the difference between an observation and an established scientific fact.

Table 3.5 provides further details on the definition of procedural and epistemic knowledge within the PISA science framework.

Table 3.5 The key components of procedural and epistemic knowledge in the PISA 2015 science framework

Procedural knowledge	Epistemic knowledge
Concept of variables	How claims are supported by data and reasoning
Concepts of measurement	The function of different forms of scientific enquiry
Ways of assessing and minimising uncertainty	How measurement error affects confidence in scientific knowledge
Mechanisms to ensure replicability and accuracy of data	The use and limitations of physical, system and abstract models
Methods of representing and using data	The role of collaboration and critique in establishing scientific claims
The use of control-of-variables and randomised controlled trials to identify possible causal mechanisms	The role of scientific knowledge in identifying societal and technological issues
The nature of an appropriate design for a given scientific question	

Source: OECD (2016:26-27)

12. Pupils in Wales are equally able in content knowledge (486) and procedural and epistemic knowledge (484). A similar pattern occurs in several of the top-performing countries, and the rest of the UK. Notable exceptions include Taiwan and Finland, where pupils have stronger content knowledge than procedural and epistemic knowledge – see Table 3.6. In Singapore, South Korea, France and the United States the opposite holds true, with pupils having stronger skills in procedural and epistemic knowledge.

Key point

In Wales, pupils' knowledge of science content is approximately equal to their knowledge of scientific practices and procedures. Wales is not unusual in this respect, with a similar pattern occurring in many other countries, including some of the top-performers in science.

Table 3.6 Average scores across the PISA 'scientific knowledge' sub-domains

Country	Content knowledge	Procedural and epistemic knowledge
Singapore	**553***	**558***
Japan	**539***	**538***
Estonia	**534***	**535***
Taiwan	**538***	**528***
Finland	**534***	**528***
Macao	**527***	**531***
Canada	**528***	**528***
Vietnam	-	-
Hong Kong	**526***	**521***
China	**520***	**516***
South Korea	**513***	**519***
New Zealand	**512***	**514***
Slovenia	**515***	**512***
England	**511***	**513***
Australia	**508***	**511***
Germany	**512***	**507***
Netherlands	**507***	**509***
Switzerland	**506***	**505***
Ireland	**504***	**501***
Belgium	**498***	**506***
Denmark	**502***	**502***
Poland	**502***	**501***
Portugal	**500***	**502***
Northern Ireland	**499***	**501***
Norway	**502***	**496***
Scotland	**496***	**496***
United States	490	**501***
Austria	**501***	490
France	489	**499***
Sweden	**498***	491
Czech Republic	**499***	488
Spain	**494***	**492***
Latvia	489	**492***
Russia	488	485
Wales	486	484
Luxembourg	483	482
Italy	483	479
Hungary	480	**474***
Lithuania	**478***	**474***
Croatia	**476***	**475***
Iceland	**468***	477
Israel	**462***	**470***
Malta	-	-
Slovakia	**463***	**458***
Greece	**455***	**454***

Notes: Table only includes countries with an average score above 450 points on the overall PISA science scale. Green/red cells indicate where the mean score for the country is at least five points higher/lower than for the mean score on the content knowledge scale. Information on sub-domain scores is not available for Malta and Vietnam. Bold font and * indicates statistically significant difference from Wales.

3.4 Example question 1. Slope face investigation.

13. To further illustrate the content of the PISA science test, we conclude this chapter by providing an analysis of two of the released PISA test questions. The first is the slope face investigation task[27]. To begin, pupils were shown an introductory information screen, as depicted in the top half of Figure 3.1. This includes a visual stimulus of two hills in a valley, one with plentiful green vegetation and one without. The screen then informs pupils how an investigation is taking place to determine which of three environmental factors (solar radiation, soil moisture and rainfall) is likely to be causing the difference in vegetation.

14. In the following screen, pupils are then told how the individuals who are conducting this investigation have placed two sets of instruments upon each hill slope. This is accompanied by the visual stimulus shown in the lower half of Figure 3.1. They are then asked the following question, with responses to be provided in an open text field:

'In investigating the difference in vegetation from one slope to the other, why did the students place two of each instrument on each slope?'

Pupils who succeeded at this question recognised the potential for measurement error to occur in this scientific study. Moreover, they recognised that collecting data from more than one instrument may help to identify and resolve this problem.

[27] Although this question is formed of several independently scored parts, our description and analysis focuses upon the first task.

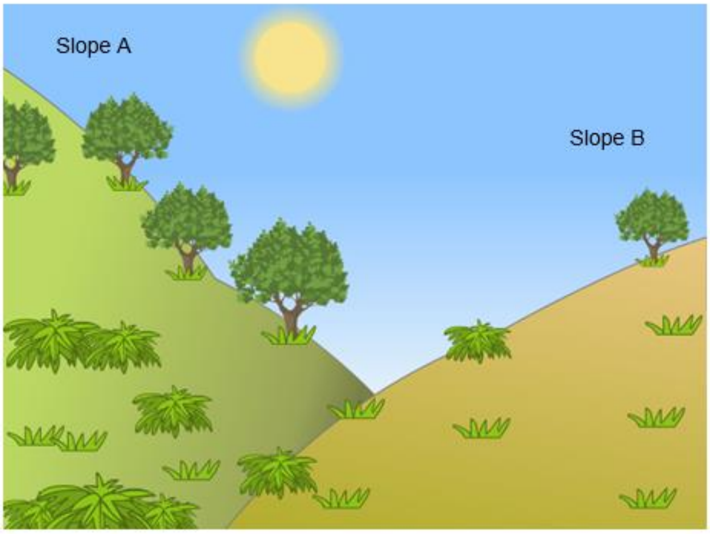
Figure 3.1 The 'slope face investigation' item

SLOPE FACE INVESTIGATION

A group of students notices a dramatic difference in the vegetation on the two slopes of a valley: the vegetation is much greener and more abundant on slope A than on slope B. This difference is shown in the illustration on the right.

The students investigate why the vegetation on the slopes is so different from one slope to the other. As part of this investigation, the students measure three environmental factors over a given period of time:

- **Solar radiation:** how much sunlight falls on a given location
- **Soil moisture:** how wet the soil is in a given location
- **Rainfall:** how much rain falls on a given location



The students place two of each of the following three instruments on each slope, as shown below.

Solar radiation sensor: measures the amount of sunlight, in megajoules per square metre (MJ/m^2)

Soil moisture sensor: measures the amount of water as a percentage of a volume of soil

Rain gauge: measures the amount of rainfall, in millimetres (mm)

Source: PISA 2015 science test.

15. Table 3.7 describes the key properties of this question. It is testing pupils' epistemic knowledge in the context of the earth and space science system. In terms of scientific competencies, it captures pupils' ability to evaluate and design scientific enquiry (and, in particular, the methods scientists use to ensure the reliability of their results). The difficulty of the question is around 517 points on the PISA science scale; pupils achieving at PISA Level 3 have around a 50/50 chance of answering this question correctly. In Wales, 57 per cent of pupils who took this question provided the correct response, with girls (60 per cent) performing slightly better than boys (55 per cent). Finally, as the PISA 2015 test was taken on computer, we know the median response time of pupils in Wales who answered this question correctly was around 70 seconds (66 seconds for boys and 74 seconds for girls). This compares to approximately 55 seconds for individuals who provided an incorrect response.

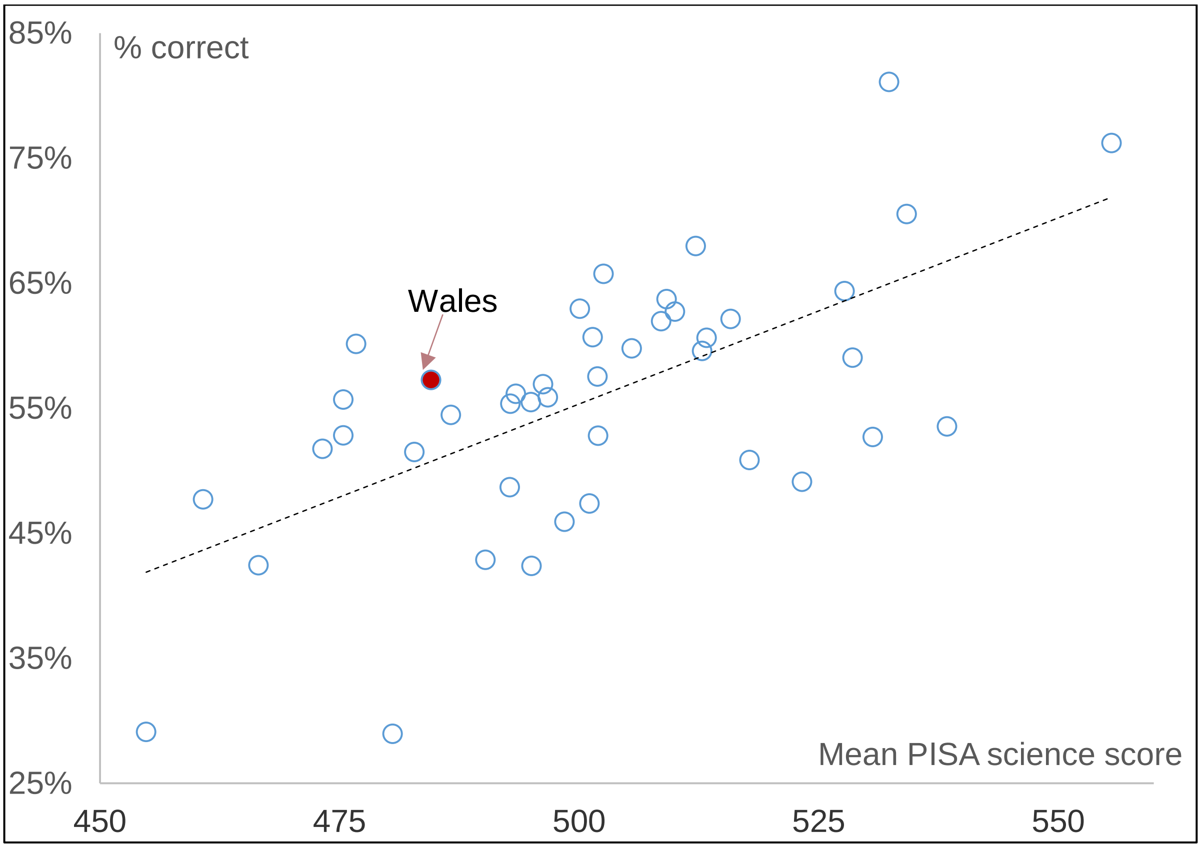
Table 3.7 Properties of the exemplar PISA science questions

	Slope face investigation	Bird migration
Item code	CS637Q01	CS656Q01
Science content system	Earth and space	Living
Scientific competency	Evaluate and design scientific enquiry	Explain phenomena scientifically
Knowledge category	Epistemic	Content
Difficulty	517 science points	501 science points
PISA level	Level 3	Level 3
% correct Wales	57%	58%
% correct girls in Wales	60%	56%
% correct boys in Wales	55%	59%
Median response time (girls correct)	74 seconds	66 seconds
Median response time (boys correct)	66 seconds	63 seconds
Median response time (girls incorrect)	53 seconds	72 seconds
Median response time (boys incorrect)	55 seconds	72 seconds

Source: PISA 2015 database and OECD (2016).

16. Figure 3.2 places Welsh pupils' performance on this question into an international context. Average PISA science scores are plotted along the horizontal axis, with the percentage of pupils providing the correct response on the vertical axis. Wales sits well above the dashed regression line; this is a question where Welsh pupils perform better than one would anticipate, given its average PISA science score. Specifically, 57 per cent of pupils in Wales answered this question correctly, compared to the 51 per cent one would expect based upon the fitted regression line. It is also interesting to note that the percentage correct for this question is similar in Scotland (56 per cent), but higher in Northern Ireland (63 per cent) and England (68 per cent).

Figure 3.2 The percentage of pupils who answer the slope face investigation question correctly across countries



Source: PISA 2015 database

3.5 Example question 2. Bird migration.

17. The second example question is from the 'bird migration' module. To begin, pupils were provided with the following information on their computer screen, along with a visual stimulus of a tagged bird.

> *'Bird migration is a seasonal large-scale movement of birds to and from their breeding grounds. Every year volunteers count migrating birds at specific locations. Scientists capture some of the birds and tag their legs with a combination of coloured rings and flags. The scientists use sightings of tagged birds together with volunteers' counts to determine the migratory routes of birds.'*

They were then asked the following question, and told to select one of the four multiple choice options:

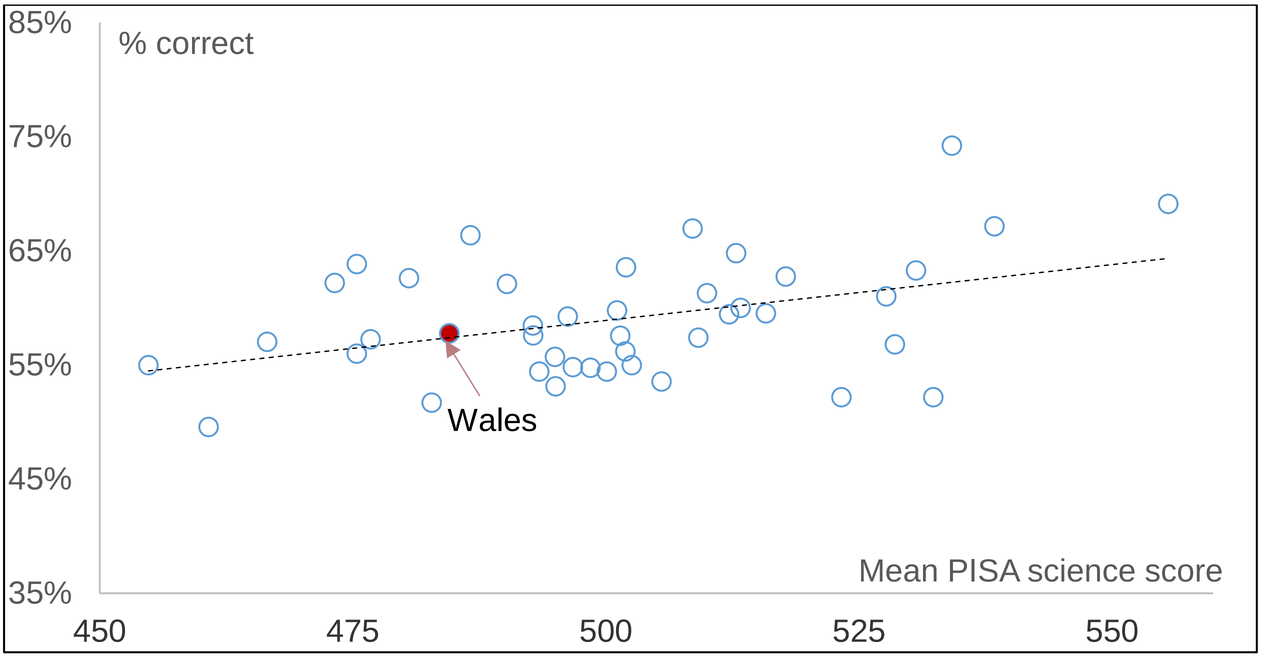
Most migratory birds gather in one area and then migrate in large groups rather than individually. This behaviour is the result of evolution. Which of the following is the best scientific explanation for the evolution of this behaviour in most migratory birds?

- *Birds that migrated individually or in small groups were less likely to survive and have offspring.*
- *Birds that migrated individually or in small groups were more likely to find adequate food.*
- *Flying in large groups allowed other bird species to join the migration.*
- *Flying in large groups allowed each bird to have a better chance of finding a nesting site*

18. Returning to Table 3.7, this question examined pupils' content knowledge of a key element within the living scientific system. In terms of scientific competencies, it captures pupils' ability to explain a particular scientific phenomenon. The difficulty of the question is around 501 points on the PISA science scale; pupils achieving at PISA Level 3 have around a 50/50 chance of answering this question correctly. In Wales, 58 per cent of pupils who took this question provided the correct response, with little difference between girls and boys. Finally the median response time of pupils in Wales who answered correctly was just over 60 seconds. This is slightly lower than the amount of time that was spent by pupils who answered incorrectly (median time of 72 seconds for both boys and girls).

19. How does Welsh pupils' performance on this question compare to pupils in other countries? The answer is provided in Figure 3.3. Wales sits below the dashed regression line; the 58 per cent who answered this question correctly is exactly what one would anticipate for a country with a mean science score of 485. Countries where pupils perform notably better on this question than in Wales include Estonia (74 per cent correct) and the Netherlands (67 per cent correct). On the other hand, the percentage correct in Hong Kong (52 per cent) and Taiwan (52 per cent) is somewhat lower than one might anticipate, given their comparatively high average PISA science scores.

Figure 3.3 Proportion of pupils answering the 'bird migration' question correctly versus average PISA science scores



Source: PISA 2015 database

Chapter 4. Achievement in mathematics

- Young people in Wales score, on average, 478 on the PISA 2015 mathematics test. This figure is not significantly different to the average score in 2006 (484).
- There are 19 countries where the average mathematics score is more than 20 points higher than in Wales. There are a further 14 countries where the average PISA mathematics score is between 10 and 20 points higher than in Wales.
- Approximately a quarter of Welsh 15-year-olds lack basic skills in mathematics. This is the same proportion as the average across members of the OECD.
- Wales has a smaller proportion of high achieving pupils in mathematics (five per cent) than the average across members of the OECD (11 per cent).
- There is a pronounced difference in mathematics achievement between the highest achieving pupils in Wales and the highest achieving pupils in other countries.
- There is no evidence that the mathematics skills of the highest achieving pupils in Wales have improved over the last decade.
- The gap between the highest and lowest achieving pupils in mathematics is 201 points (around six and a half years of schooling) in Wales. This is a significantly smaller difference than in most other countries.

4.1 What is the average PISA mathematics score in Wales, and how does this compare to other countries?

1. An understanding of mathematics is central to a young person's preparedness for life in modern society. A growing proportion of problems and situations encountered in daily life, including in professional contexts, require some level of understanding of mathematics, mathematical reasoning and mathematical tools, before they can be fully understood and addressed. Mathematics is a critical tool for young people as they confront issues and challenges in personal, occupational, societal, and scientific aspects of their lives. It is therefore important to have an understanding of the degree to which young people emerging from school are adequately prepared to apply mathematics to understanding important issues and solving meaningful problems. The results from PISA 2015 provide such insight, helping us to understand whether 15-year-olds in Wales are able to use their knowledge and skills in mathematics to solve real world problems. Table 4.1 presents the average PISA mathematics score for Wales, and how this compares in an international comparative context.

2. The mean PISA mathematics score in Wales is 478. Panel (a) refers to those countries where the average PISA mathematics score is at least 20 points higher. A total of 19 countries belong to this group; the top seven being from East Asia (Singapore, Hong Kong, Macao, Taiwan, Japan, China and South Korea). Panel A also includes 11 countries from Europe, and one from North America (Canada).

3. Panel (b) of Table 4.1 turns to countries where the average PISA mathematics score is between 10 and 20 test points higher than Wales. There are 14 countries within this group, mostly from Europe. This includes England (493), Northern Ireland (493) and Scotland (491), along with Sweden (494), France (493) and Italy (490). For each of these countries, the average PISA mathematics score ranges between 488 and 497 test points.

4. Panel (c) includes all countries within 10 points of the mean mathematics score for Wales. Differences of this magnitude are equivalent to less than four months (one term) of schooling, and generally not outside the range one would expect given sampling error. A total of nine countries are within this group (excluding Wales). It includes several Eastern European nations, such as Latvia (482), Lithuania (478) and Slovakia (475). Another notable country with a similar average PISA mathematics score to Wales is the United States (470).

Table 4.1 Mean PISA 2015 mathematics scores

(a) Countries more than 20 points ahead of Wales

Country	Mean	Country	Mean	Country	Mean
Singapore	**564***	**Switzerland**	**521***	**Belgium**	**507***
Hong Kong	**548***	**Estonia**	**520***	**Germany**	**506***
Macao	**544***	**Canada**	**516***	**Poland**	**504***
Taiwan	**542***	**Netherlands**	**512***	**Ireland**	**504***
Japan	**532***	**Denmark**	**511***	**Norway**	**502***
China	**531***	**Finland**	**511***		
South Korea	**524***	**Slovenia**	**510***		

(b) Countries between 10 and 20 points ahead of Wales

Country	Mean	Country	Mean	Country	Mean
Austria	**497***	**Australia**	**494***	**Portugal**	**492***
New Zealand	**495***	**England**	**493***	**Scotland**	**491***
Vietnam	**495***	**France**	**493***	**Italy**	**490***
Russia	**494***	**Northern Ireland**	**493***	**Iceland**	**488***
Sweden	**494***	**Czech Republic**	**492***		

(c) Countries within 10 points of Wales

Country	Mean	Country	Mean	Country	Mean
Spain	486	Lithuania	478	Israel	470
Luxembourg	486	Wales	478	United States	470
Latvia	482	Hungary	477		
Malta	479	Slovakia	475		

(d) Countries between 10 and 20 points behind Wales

Country	Mean
Croatia	**464***

Source: PISA 2015 database.

Note: Bold with a * indicates mean score significantly different from Wales at the five per cent level. Table does not include countries with average mathematics scores more than 20 points lower than Wales.

5. The final panel of Table 4.1 (panel d) contains countries where the average PISA mathematics score is between 10 and 20 points below the mean score for Wales. Just one country belongs to this group (Croatia with a mean of 464). However, it is important to note that Table 4.1 does not include any country with a mean PISA mathematics score more than 20 points below the score for Wales. Results have therefore not been presented for 27 countries, including some members of the OECD, such as Greece (454). A full set of average PISA mathematics scores, including all participating countries, is provided in the online data tables.

Key point

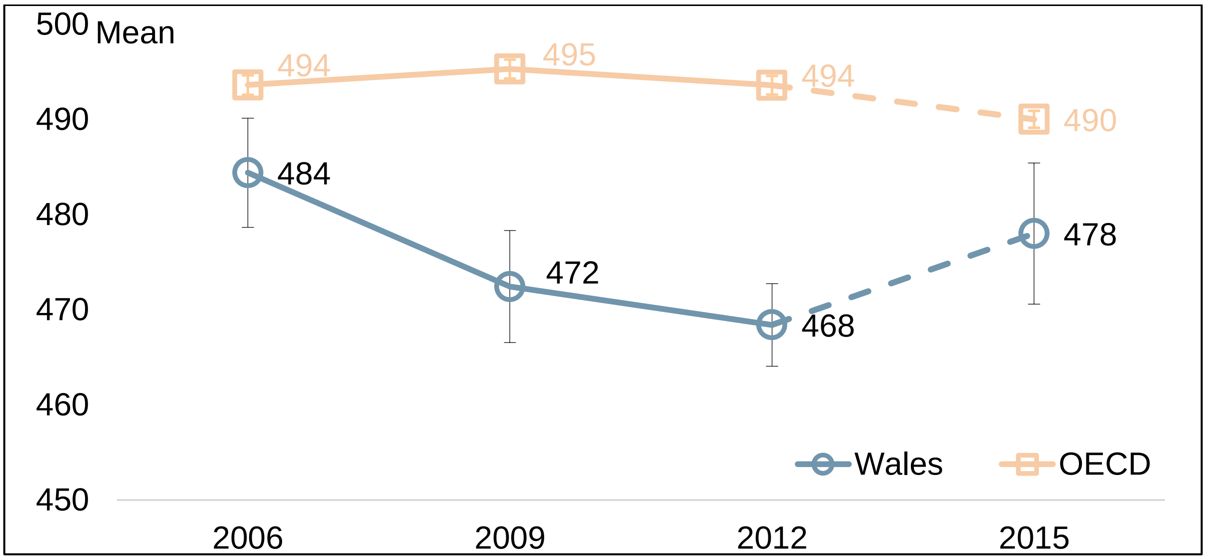
The average PISA mathematics score in Wales is 478. There are 33 countries where the average is at least 10 test points higher than in Wales, and 28 countries where the average is at least 10 test points lower.

4.2 How have average PISA mathematics scores in Wales changed over time? How does this compare to other countries?

6. Figure 4.1 illustrates that, although there has been some fluctuation over time, the mean PISA mathematics score for Wales is around the same level in 2015 as it was in 2006. Indeed, the difference between the mean score in 2015 (478) and 2006 (484) is not statistically significant at the five per cent level. There is hence no evidence of any significant increase or decrease in average PISA mathematics scores in Wales over the last decade.

7. Table 4.2 compares the change for Wales to the five ‘fastest improving’ (red cells) and the five ‘fastest declining’ (blue cells) countries. In order to facilitate relevant comparisons, any country where the average PISA 2015 mathematics score is below 450 points has been excluded from this table. Results are presented for both the change between 2006 and 2015 (panel a), and between 2012 and 2015 (panel b).

Figure 4.1 Mean PISA mathematics scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the 'AV09' results presented in the OECD international results Table I.5.4a. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

8. Starting with panel (a), Italy has experienced the greatest improvement in mean mathematics scores between 2006 to 2015, gaining approximately 28 PISA test points (moving from 462 to 490 on the PISA mathematics scale). Other countries with a more than 20 test point (eight months of schooling) increase include Israel and Portugal. In contrast, Finland (-37 points, falling from 548 to 511), New Zealand (-27 points, falling from 522 to 495) and Australia (-26 points, from 520 to 494) have suffered the most pronounced declines.

9. Panel (b) of Table 4.2 provides the analogous comparison between PISA 2012 and PISA 2015. Some countries, such as Vietnam, took part in PISA for the first time in 2012. The countries included in the comparison are now rather different. Sweden saw the biggest increase in mathematics scores between 2012 and 2015 (from 478 to 494), returning the mean for Sweden back to its level in 2009. On the other hand, a 30 point fall has occurred in South Korea, though it is too early to tell whether this is a one-off decline or part of a sustained trend[28]. Other countries with a notable improvement or decline in mean mathematics scores since 2012 include Norway (+12 points), Taiwan (-18 points) and Vietnam (-17 points).

[28] In particular, note that the mean mathematics score in South Korea was 547 in 2006, 546 in 2009 and 554 in 2012, before a sharp drop to 524 in 2015.

Table 4.2 The five fastest improving and declining countries in mathematics

(a) PISA 2006 to 2015

Country	From	To	Change
Italy	462	490	**+28***
Israel	442	470	**+28***
Portugal	466	492	**+25***
Macao	525	544	**+19***
Russia	476	494	**+18***
Netherlands	531	512	**-18***
South Korea	547	524	**-23***
Australia	520	494	**-26***
New Zealand	522	495	**-27***
Finland	548	511	**-37***

(b) PISA 2012 to 2015

Country	From	To	Change
Sweden	478	494	**+16***
Norway	489	502	**+12***
Russia	482	494	**+12***
Denmark	500	511	**+11***
Wales	468	478	+10
Poland	518	504	**-13***
Hong Kong	561	548	**-13***
Vietnam	511	495	**-17***
Taiwan	560	542	**-18***
South Korea	554	524	**-30***

Source: PISA 2015 database.

Note: Figures refer to change between cycles in the mean PISA mathematics score. Table restricted to only those countries with a mean score above 450 in the PISA 2015 mathematics test. Bold font with * indicates change between cycles statistically significant at the five per cent level. The difference between the 'from' and 'to' columns may not equal the 'change' column due to rounding. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

Key point

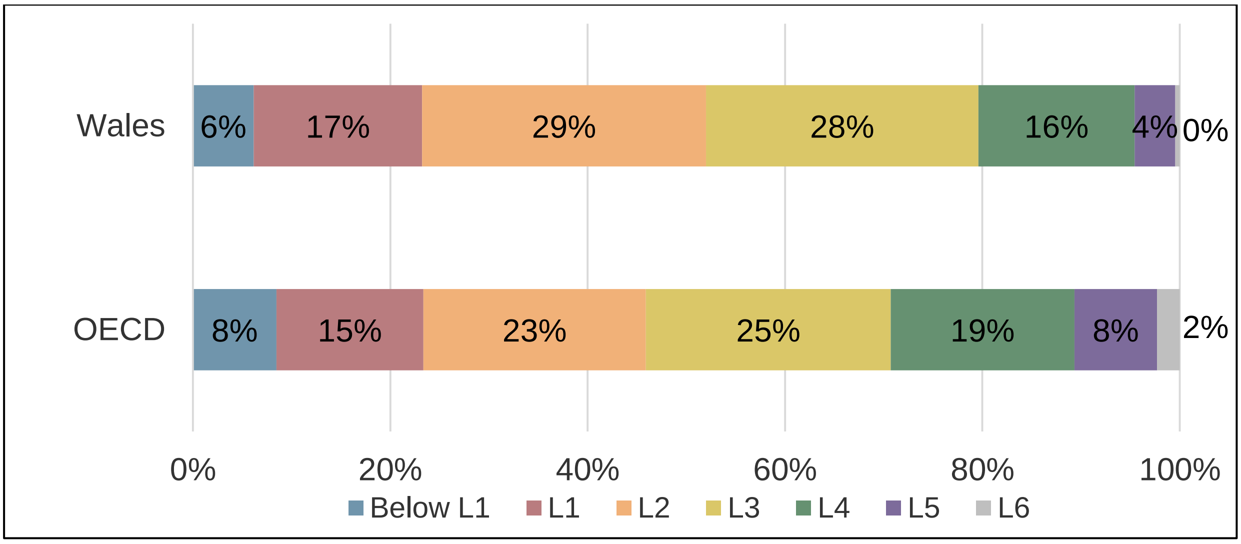
There is little evidence of a sustained change in average PISA mathematics scores over the last decade in Wales.

4.3 What proportion of pupils in Wales reach each mathematics proficiency level?

10. Figure 4.2 illustrates the proportion of pupils in Wales reaching each PISA mathematics level, and compares this to the average across members of the OECD. In Wales, six per cent of 15-year-olds are working below PISA mathematics Level 1, while 17 per cent of 15-year-olds reach Level 1. Analogous figures for the average across OECD members are eight per cent below Level 1 and 15 per cent at Level 1. Therefore, the proportion of 'low-achievers' in Wales (23 per cent) is similar to the OECD average (23 per cent).

11. However, there is a more notable difference between Wales and other industrialised countries in the proportion of pupils who reach the top two PISA levels. Specifically, around one-in-twenty (five per cent) Welsh pupils reach PISA Level 5 or Level 6 in mathematics. This is substantially lower than the average across members of the OECD (11 per cent). It therefore seems that Wales faces a particular challenge in developing young people with high-level mathematics skills.

Figure 4.2 The proportion of pupils reaching each mathematics proficiency level



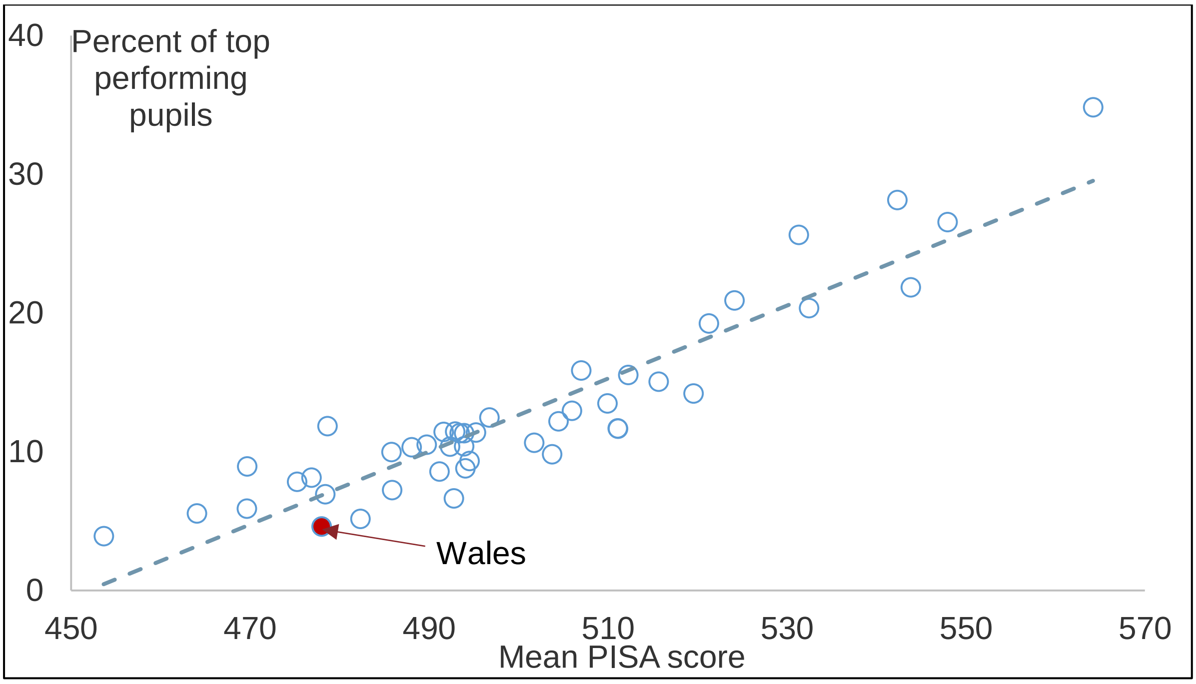
Source: PISA 2015 database.

12. Figure 4.3 provides further insight into how Wales compares to other countries in terms of the proportion of high-performing pupils in mathematics. The horizontal axis plots the average PISA mathematics score, while the vertical axis presents the proportion of pupils in each country achieving PISA Level 5 or Level 6. The dashed regression line then illustrates the cross-country relationship between

these variables. In this figure, the sample of countries has been restricted to those with a mean mathematics score above 450 points.

13. Wales sits below the dashed regression line; the proportion of high achieving pupils in mathematics is lower than one would expect given its mean score. In particular, the regression line suggests that a typical country with a mean PISA mathematics score of 478 will have around eight per cent of its 15-year-olds reaching one of the top two PISA levels. Yet, in Wales, only around five per cent of pupils achieve this benchmark. This further highlights the lack of high achieving pupils in mathematics in Wales.

Figure 4.3 The percent of top-performing pupils in mathematics compared to mean PISA mathematics scores: a cross-country analysis



Source: PISA 2015 database.

Notes: The sample of countries included in this figure has been restricted to those with a mean mathematics score above 450 points.

Key point

Wales has a similar proportion of low achieving pupils in mathematics as the average across members of the OECD. However, Wales seems to face a particular challenge in developing enough young people with high-level mathematics skills.

4.4 How do the PISA mathematics scores of the ***highest*** achieving pupils in Wales compare to other countries?

14. The previous sub-section highlighted how Wales has a smaller proportion of high-performing pupils in mathematics than the average member of the OECD. We now provide further insight into the proficiency of the highest achieving pupils by comparing the 90th percentile of the mathematics distribution for Wales to the 90th percentile in other countries. We then consider whether the PISA mathematics score of the highest achieving pupils in Wales has changed over time.

15. Table 4.3 compares the 90th percentile of the PISA mathematics distribution for Wales to a range of other countries. In 2015, the 90th percentile of the mathematics proficiency distribution in Wales was 578. This means that the top-performing 10 per cent of 15-year-olds in Wales achieved a PISA score of 578 test points or more. This figure is lower than in many other countries that participated in the PISA 2015 assessment. In particular, there are 36 countries where the 90th percentile is more than 20 points above the value for Wales, with a further four countries where the 90th percentile is between 10 and 20 points higher. Conversely, there are relatively few industrialised nations where the value of the 90th percentile is significantly lower than in Wales. (Turkey, Mexico and Chile are the only members of the OECD where the 90th percentile is lower – see the online data tables for further details). Overall, Table 4.3 illustrates how the mathematics skills of the highest achieving pupils in Wales are significantly below the skills of the highest achieving pupils in a number of other countries.

16. How have the mathematics skills of the highest achieving pupils in Wales changed over the last decade? Figure 4.4 provides the answer by plotting the 90th percentile of the PISA mathematics distribution from 2006 to 2015, accompanied by the estimated 95 per cent confidence interval. Overall, there is relatively little evidence of a change in the 90th percentile in Wales since 2006. Specifically, the 90th percentile stood at a very similar value in 2015 (578), 2012 (578) and 2009 (578). Although the point estimate was slightly higher in 2006 (592), there is no consistent evidence of a sustained upwards or downwards trend.

Table 4.3 The 90th percentile of PISA 2015 mathematics scores

(a) Countries more than 20 points ahead of Wales

Country	P90	Country	P90	Country	P90
Singapore	**682***	**Slovenia**	**622***	**Norway**	**610***
Taiwan	**670***	**Germany**	**620***	**Italy**	**610***
China	**664***	**Austria**	**618***	**Sweden**	**609***
Hong Kong	**659***	**Poland**	**617***	**Czech Republic**	**608***
South Korea	**649***	**Malta**	**616***	**Iceland**	**608***
Macao	**643***	**Finland**	**614***	**Luxembourg**	**607***
Japan	**643***	**Denmark**	**614***	**Ireland**	**606***
Switzerland	**641***	**Portugal**	**614***	**Vietnam**	**604***
Belgium	**630***	**Australia**	**613***	**Russia**	**601***
Canada	**627***	**England**	**613***	**Israel**	**601***
Netherlands	**627***	**New Zealand**	**613***	**Scotland**	**601***
Estonia	**623***	**France**	**613***		

(b) Countries between 10 and 20 points ahead of Wales

Country	P90	Country	P90
Hungary	**598***	Northern Ireland	592
Slovakia	**596***	**Lithuania**	**590***
Spain	**593***		

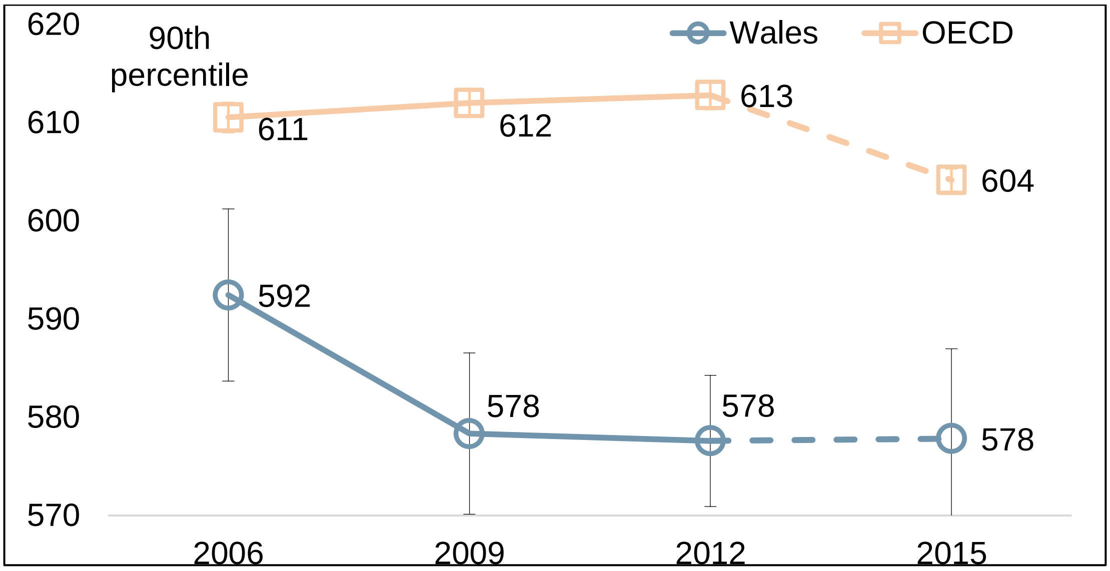
(c) Countries within 10 points of Wales

Country	P90	Country	P90	Country	P90
United States	585	Croatia	580	Greece	570
Latvia	582	Wales	578	Bulgaria	568

Source: PISA 2015 database.

Note: Bold with a * indicates significant difference from Wales at the five per cent level. Table does not include countries where the 90th percentile of the mathematics distribution is more than 20 points below Wales. There are no countries where the 90th percentile of the mathematics distribution is between 10 and 20 points lower than in Wales.

Figure 4.4 The 90th percentile of PISA mathematics scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the 'AV09' results presented in the OECD international results Table I.5.4b. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

Key point

There has been little improvement in the PISA scores of the highest achieving pupils in mathematics in Wales since 2006.

4.5 How do the mathematics scores of the *lowest* achieving pupils in Wales compare to other countries?

17. Although the mathematics skills of the highest achieving pupils in Wales may be lower than the top performing pupils in other countries, does the same hold true for the lowest achievers? Table 4.4 provides evidence on this matter by comparing the 10th percentile of the PISA mathematics distribution across countries.

Table 4.4 The 10th percentile of PISA 2015 mathematics scores

(a) Countries more than 20 points ahead of Wales

Country	P10	Country	P10	Country	P10
Macao	**439***	**Estonia**	**415***	**Canada**	**400***
Singapore	**436***	**Denmark**	**405***	**Ireland**	**400***
Hong Kong	**426***	**Finland**	**404***		
Japan	**416***	**Taiwan**	**404***		

(b) Countries between 10 and 20 points ahead of Wales

Country	P10	Country	P10	Country	P10
Switzerland	**394***	**Norway**	**391***	China	388
Slovenia	**394***	Netherlands	390	Northern Ireland	388
Poland	**391***	Germany	389	Russia	387
South Korea	391	Vietnam	388		

(c) Countries within 10 points of Wales

Country	P10	Country	P10	Country	P10
Scotland	382	New Zealand	375	Australia	371
Latvia	382	Belgium	374	Austria	370
Wales	377	Spain	374	England	369
Sweden	376	Czech Republic	373	Italy	368

(d) Countries between 10 and 20 points behind Wales

Country	P10	Country	P10	Country	P10
Iceland	367	Portugal	365	**Luxembourg**	**363***
Lithuania	365	France	364		

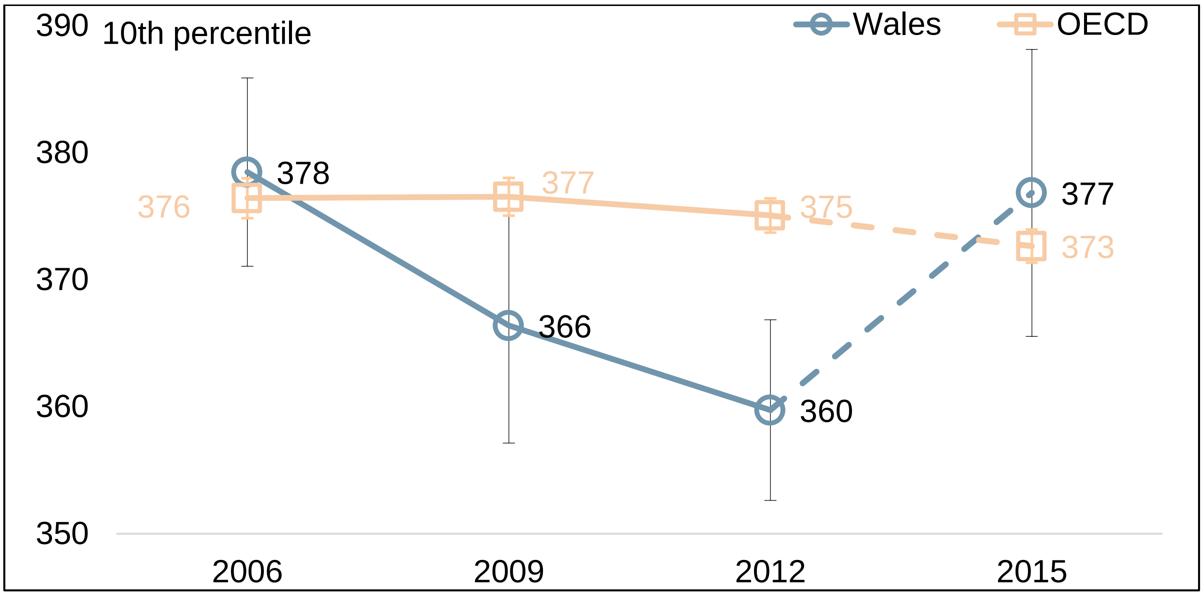
Source: PISA 2015 database.

Note: Bold font with a * indicates significant difference from Wales at the five per cent level. Table does not include countries where the 10th percentile of the mathematics distribution is more than 20 points below Wales.

18. The value of the 10th percentile of the PISA mathematics distribution in Wales is 377. There are 10 countries where the 10th percentile is more than 20 points above the value for Wales, with five of these within East Asia. In a further 11 countries, the 10th percentile is between 10 and 20 points above Wales. However, Table 4.4 also indicates that low achieving pupils in Wales achieve similar PISA mathematics scores to 15-year-olds in a number of other OECD countries, including

England (10th percentile = 369), Australia (371), New Zealand (375) and Sweden (376). Likewise, the 10th percentile of the mathematics distribution in Wales is higher than in a selection of other industrialised nations, including France (364) and the United States (355). Overall, the position of Wales in this international comparison of low-achievers is somewhat more favourable than the results previously presented for the highest achievers in Table 4.3.

Figure 4.5 The 10th percentile of PISA mathematics scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the 'AV09' results presented in the OECD international results Table I.5.4b. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

19. Figure 4.5 proceeds by considering how the 10th percentile of PISA mathematics scores in Wales has changed since 2006. The point estimate of the 10th percentile was 378 in 2006, 366 in 2009, 360 in 2012 and 377 in 2015. Although there has been some fluctuation over this period, the difference between the 2006 and 2015 value is not statistically significant at the five per cent level. There is hence no evidence of either a sustained decline or increase in the mathematics skills of the lowest achieving pupils over this period.

Key point

There is no evidence that the mathematics skills of the lowest performing pupils in Wales have improved or declined over the last decade.

4.6 How big is the gap between the pupils with the strongest and weakest mathematics skills? How does Wales compare to other countries in this respect?

20. To conclude this chapter, we consider inequality in 15-year-olds' mathematics skills, as measured by the difference between the 90th percentile and the 10th percentile. The magnitude of this gap is presented in Table 4.5. For brevity, the sample is restricted to only those countries with a mean PISA mathematics score above 450 points. The 10 countries with the highest mean PISA mathematics scores have been highlighted in red.

21. The 90th percentile of the PISA mathematics distribution in Wales is 578, while the 10th percentile stands at 377. Table 4.5 demonstrates that the gap is therefore 201 test points, equivalent to almost seven years of schooling. This is smaller than in almost every other country included in the comparison (OECD average = 232). Indeed, there is no country in Table 4.5 where the difference between the 90th and 10th percentile is significantly smaller than in Wales at the five per cent level. Conversely, there are 35 countries where inequality in mathematics achievement is significantly greater, and often by more than 30 PISA test points. Consequently, by this metric, Wales has a particularly equal distribution of 15-year-olds' mathematics achievement. However, as illustrated in the previous sub-sections, this finding is being at least partially driven by the lack of Welsh pupils with high-level mathematics skills.

Key point

The gap between the highest and lowest achieving pupils in mathematics is smaller in Wales than in most other countries.

Table 4.5 Difference between the highest and lowest achievers in mathematics

Country	Difference between the 90th and 10th percentile	Difference in years of schooling
Malta	**285***	**9.5 years**
China	**276***	**9.2 years**
Israel	**269***	**9.0 years**
Taiwan	**266***	**8.9 years**
South Korea	**258***	**8.6 years**
Belgium	**255***	**8.5 years**
France	**249***	**8.3 years**
Portugal	**249***	**8.3 years**
Switzerland	**247***	**8.2 years**
Slovakia	**247***	**8.2 years**
Austria	**247***	**8.2 years**
Singapore	**247***	**8.2 years**
Hungary	**246***	**8.2 years**
England	**245***	**8.2 years**
Luxembourg	**244***	**8.1 years**
Australia	**242***	**8.1 years**
Iceland	**241***	**8.0 years**
Italy	**241***	**8.0 years**
New Zealand	**238***	**7.9 years**
Netherlands	**237***	**7.9 years**
Czech Republic	**235***	**7.8 years**
Greece	**234***	**7.8 years**
Sweden	**233***	**7.8 years**
Hong Kong	**232***	**7.7 years**
Germany	**230***	**7.7 years**
United States	**230***	**7.7 years**
Croatia	**229***	**7.6 years**
Slovenia	**228***	**7.6 years**
Canada	**227***	**7.6 years**
Japan	**227***	**7.6 years**
Poland	**226***	**7.5 years**
Lithuania	**225***	**7.5 years**
Spain	**220***	**7.3 years**
Scotland	**219***	**7.3 years**
Norway	**219***	**7.3 years**
Vietnam	215	7.2 years
Russia	214	7.1 years
Finland	210	7.0 years
Denmark	209	7.0 years
Estonia	209	7.0 years
Ireland	206	6.9 years
Macao	204	6.8 years
Northern Ireland	204	6.8 years
Wales	201	6.7 years
Latvia	200	6.7 years

Source: PISA 2015 database.

Note: Bold with a * indicates statistically significant difference compared to Wales at the five per cent level. Table only includes countries where the mean PISA mathematics score is above 450.

Chapter 5. Achievement in reading

- The average PISA reading score in Wales is 477. This figure is not significantly different to the average score in 2006 (481).

- There are 22 countries where the average reading score is more than 20 points higher than in Wales. There are a further nine countries where the average PISA reading score is between 10 and 20 points higher.

- Around one-in-five (21 per cent) 15-year-olds in Wales lack basic reading skills. This is similar to the average across members of the OECD (20 per cent).

- Wales has fewer high achieving pupils in reading (four per cent) than the average across OECD countries (eight per cent).

- The reading skills of the highest achieving Welsh pupils remain more than 20 PISA test points (eight months of schooling) behind the highest achieving pupils in 32 other countries.

- The difference in reading skills between the highest and lowest achieving pupils in Wales is 219 test points (seven and a quarter years of schooling). This is a significantly smaller difference than in most other countries, suggesting that inequality in 15-year-olds' reading skills is lower in Wales than in most other parts of the industrialised world.

5.1 What is the average PISA reading score in Wales, and how does this compare to other countries?

1. Achievement in reading literacy is not only a foundation for achievement in other subject areas, but also a prerequisite for successful participation in most areas of adult life. Indeed, although greater levels of reading literacy are associated with higher economic returns[29], the impact of reading literacy upon personal well-being and social cohesion is likely to be just as important[30]. This foundational nature of reading literacy has been summed up by the European Commission[31], which noted such skills to be '*key to all areas of education and beyond, facilitating participation in the wider context of lifelong learning and contributing to individuals' social integration and personal development.*' Throughout this chapter we therefore consider the reading proficiency of 15-year-olds in Wales, and how this compares to the reading skills of young people living in other countries. This particular sub-section focuses upon average PISA reading scores.

2. The mean PISA reading score in Wales is 477. Panel (a) of Table 5.1 lists the countries where the average PISA reading score is at least 20 points higher than in Wales. A total of 22 countries belong to this group; five from East Asia, 14 from Europe (including England) along with Australia, Canada and New Zealand. The average PISA reading score in all of these countries is at least 498 test points.

3. Panel (b) of Table 5.1 turns to countries where the average PISA reading score is between 10 and 20 test points higher than Wales. There are nine countries within this group, the majority from Europe (Latvia, Spain, Switzerland, Scotland and Northern Ireland), as well as Russia (495), China (494), Taiwan (497) and the United States (497).

4. Panel (c) includes all countries within 10 points of the mean reading score in Wales. Differences of this magnitude are equivalent to less than four months (one term) of schooling, and generally not outside the range one would expect given sampling error. A total of 10 countries are within this group (excluding Wales). These are mostly European nations, including several from Eastern Europe, such as Croatia (487), Lithuania (472) and Hungary (470). Other non-European countries with a similar average PISA reading score to Wales include Vietnam (487) and Israel (479).

[29] Machin and McNally (2008).
[30] Friedman (2005) and OECD (2001).
[31] European Commission (2001).

Table 5.1 Mean PISA 2015 reading scores

(a) Countries more than 20 points ahead of Wales

Country	Mean	Country	Mean	Country	Mean
Singapore	**535***	**Norway**	**513***	**Australia**	**503***
Hong Kong	**527***	**New Zealand**	**509***	**Sweden**	**500***
Canada	**527***	**Germany**	**509***	**Denmark**	**500***
Finland	**526***	**Macao**	**509***	**England**	**500***
Ireland	**521***	**Poland**	**506***	**France**	**499***
Estonia	**519***	**Slovenia**	**505***	**Belgium**	**499***
South Korea	**517***	**Netherlands**	**503***	**Portugal**	**498***
Japan	**516***				

(b) Countries between 10 and 20 points ahead of Wales

Country	Mean	Country	Mean
Taiwan	**497***	**China**	**494***
Northern Ireland	**497***	**Scotland**	**493***
United States	**497***	**Switzerland**	**492***
Spain	**496***	**Latvia**	**488***
Russia	**495***		

(c) Countries within 10 points of Wales

Country	Mean	Country	Mean	Country	Mean
Czech Republic	**487***	Italy	485	Wales	477
Croatia	**487***	Iceland	482	Lithuania	472
Vietnam	487	Luxembourg	481	Hungary	470
Austria	485	Israel	479		

(d) Countries between 10 and 20 points behind Wales

Country	Mean	Country	Mean
Greece	467	**Chile**	**459***

Source: PISA 2015 database.

Note: Bold with a * indicates mean score significantly different from Wales at the five per cent level. Table does not include countries with average reading scores more than 20 points lower than in Wales.

5. The final panel of Table 5.1 (panel d) contains countries where the average PISA reading score is between 10 and 20 points below Wales. Just two nations belong to this group; Greece (467) and Chile (459). However, it is important to note that Table 5.1 does not include any country with a mean PISA reading score more than 20 points below the score for Wales. Results have therefore not been presented for 27 countries, including some members of the OECD, such as Slovakia (453), Turkey (428) and Mexico (423). A full set of average PISA reading scores, including all participating countries, is provided in the online data tables.

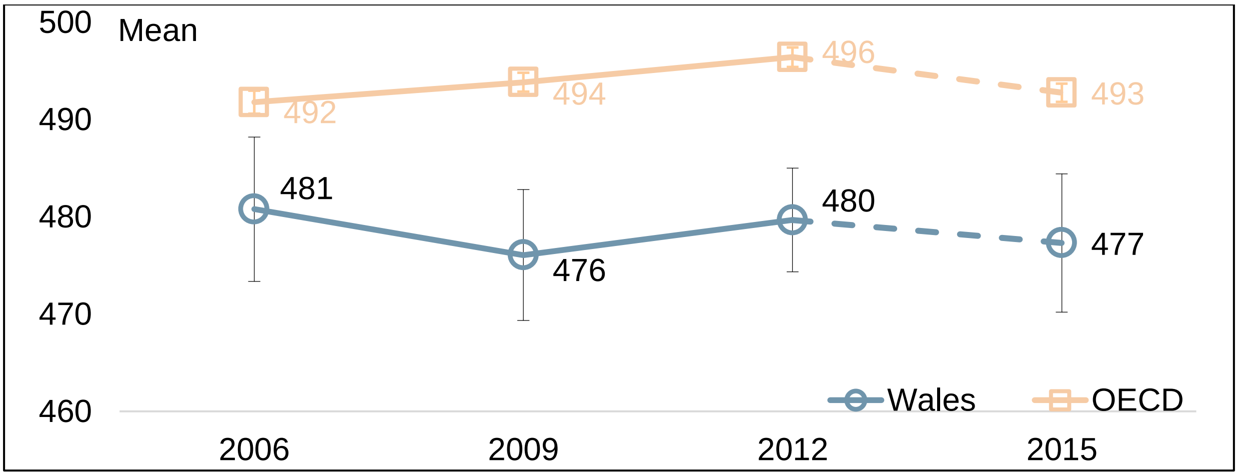
Key point

The average PISA reading score in Wales is 477. There are 31 countries where the average is at least 10 test points higher than in Wales, and 29 countries where the average is at least 10 test points lower.

5.2 How have average PISA reading scores in Wales changed over time? How does this compare to other countries?

6. Figure 5.1 illustrates that the mean PISA reading score for Wales has remained stable over time. Specifically, the average PISA reading score in 2015 for Wales (477) is not significantly different from the mean score in 2012 (480), 2009 (476) or 2006 (481). There is hence no evidence of any significant increase or decrease in average PISA reading scores in Wales over the last decade.

Figure 5.1 Mean PISA reading scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the ‘AV09’ results presented in the OECD international results Table I.4.4a. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

7. Table 5.2 compares the change for Wales to the five 'fastest improving' (red cells) and the five 'fastest declining' (blue cells) countries. In order to facilitate relevant comparisons, any country where the average PISA 2015 reading score is below 450 points has been excluded from this table. Results are presented for both the change between 2006 and 2015 (panel a), and between 2012 and 2015 (panel b).

Table 5.2 The five fastest improving and the five fastest declining countries in reading

(a) PISA 2006 to 2015

Country	From	To	Change
Russia	440	495	**+55***
Israel	439	479	**+40***
Spain	461	496	**+35***
Norway	484	513	**+29***
Portugal	472	498	**+26***
New Zealand	521	509	-12
Hungary	482	470	-13
Slovakia	466	453	-14
Finland	547	526	**-20***
South Korea	556	517	**-39***

(b) PISA 2012 to 2015

Country	From	To	Change
Slovenia	481	505	**+24***
Russia	475	495	**+19***
Chile	441	459	**+17***
Sweden	483	500	**+17***
Portugal	488	498	+10
South Korea	536	517	**-18***
Hungary	488	470	**-19***
Vietnam	508	487	**-21***
Japan	538	516	**-22***
Taiwan	523	497	**-26***

Source: PISA 2015 database.

Note: Figures refer to change between cycles in the mean PISA reading score. Table restricted to only those countries with a mean score above 450 in the PISA 2015 reading test. Bold font with a * indicates where change between cycles statistically significant at the five per cent level. The difference between the 'from' and 'to' columns may not equal the 'change' column due to rounding.

8. Starting with panel (a), Russia has experienced the greatest improvement in mean reading scores during the 2006 to 2015 period, gaining approximately 55 test points (moving from 440 to 495 on the PISA reading scale). Other countries with a greater than 20 test point (eight months of schooling) increase include Israel (+40, from 439 to 479), Spain (+35, from 461 to 496), Norway (+29, from 484 to 513) and Portugal (+26, from 472 to 498). In contrast, South Korea (-39 points, falling from 556 to 517) and Finland (-20 points, from 547 to 526) have suffered the most pronounced declines.

9. Panel (b) of Table 5.2 provides the analogous comparison between PISA 2012 and PISA 2015. Perhaps the most notable feature of this table is that four of the five countries with the biggest decline since 2012 are East Asian. This includes South Korea (-18 points, from 536 to 517), Japan (-22 points, from 538 to 516), Vietnam (-21 points, from 508 to 487) and Taiwan (-26 points, from 523 to 497). However, for many of these countries, it is too early to tell whether this is due to a one-off fall or part of a sustained trend. On the other hand, Slovenia (+24 points), Russia (+19 points), Sweden (+17 points) and Chile (+17 points) have demonstrated the greatest improvement in average PISA reading scores since PISA 2012.

Key point

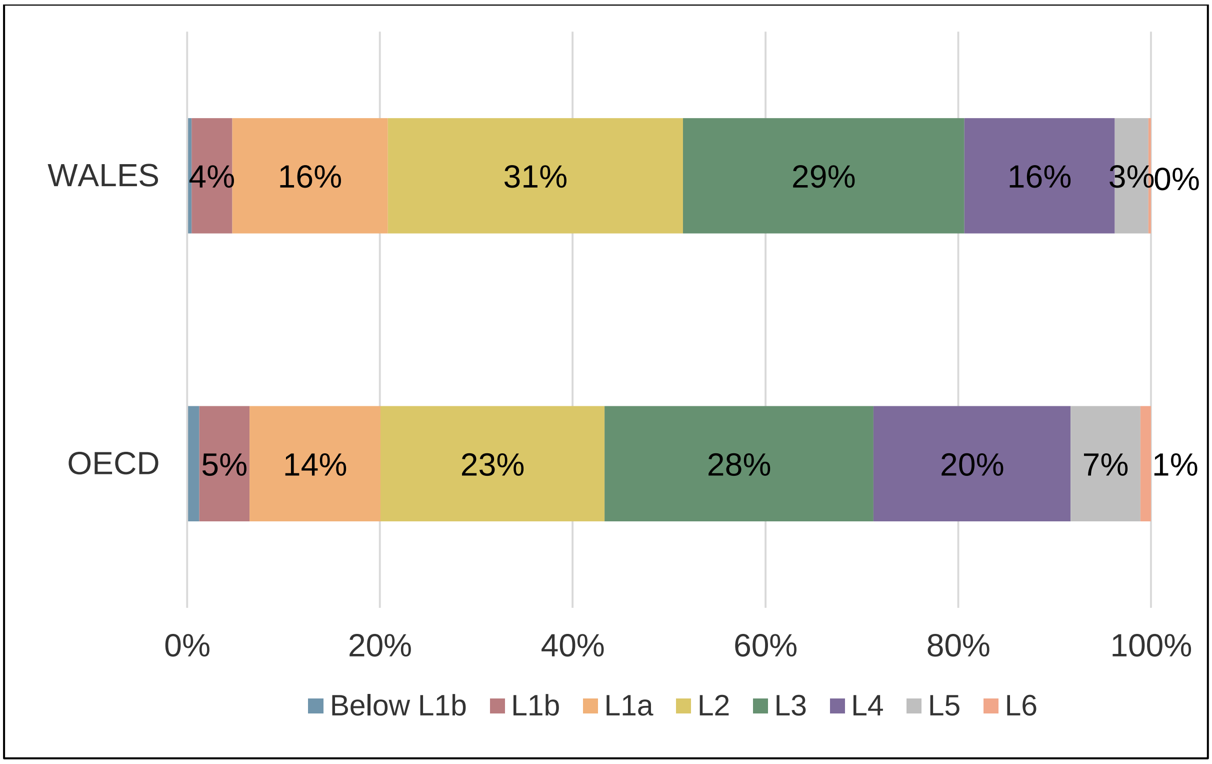
There has been no statistically significant change in the average PISA reading score for Wales since 2006.

5.3 What proportion of pupils in Wales reach each reading proficiency level?

10. Figure 5.2 illustrates the proportion of pupils in Wales reaching each PISA reading level, and compares this to the average across members of the OECD. In Wales, one per cent of 15-year-olds are working below PISA reading Level 1b, four per cent reach Level 1b, while 16 per cent reach Level 1a. Analogous figures for the average across OECD members are one per cent below Level 1b, five per cent at Level 1b and 14 per cent at Level 1a. The proportion of ‘low-achievers’ in Wales (21 per cent) is therefore similar to the average across members of the OECD (20 per cent).

11. On the other hand, the proportion of high achieving pupils in reading in Wales is somewhat below the OECD average. Specifically, around four per cent of pupils in Wales reach one of the top two PISA achievement levels in reading. This compares to an average across OECD members of approximately eight per cent.

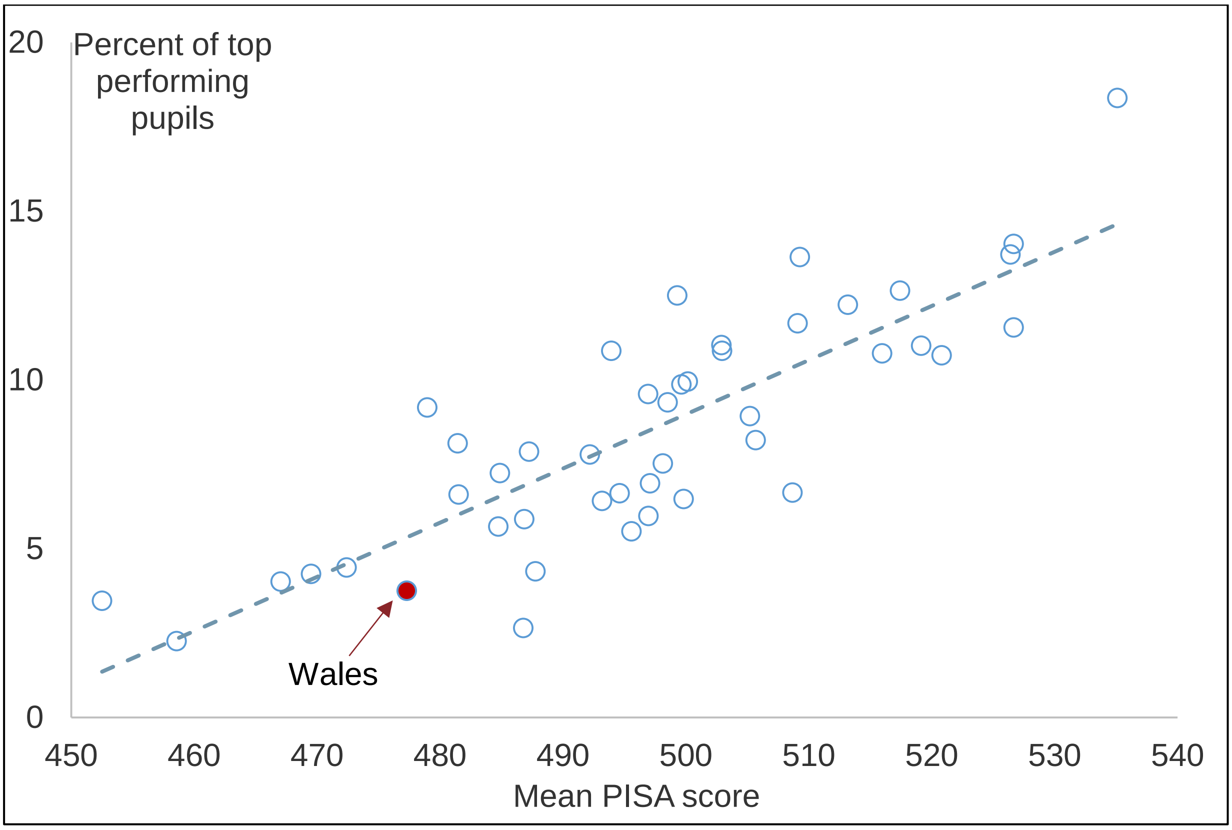
Figure 5.2 The proportion of pupils in Wales reaching each PISA reading level



Source: PISA 2015 database.

12. Figure 5.3 provides further insight into how Wales compares to other countries in terms of the proportion of high-performing pupils in reading. The horizontal axis plots the average PISA reading score, while the vertical axis presents the proportion of pupils in each country achieving PISA Level 5 or Level 6. The dashed regression line then illustrates the cross-country relationship between these variables. In this figure, the sample of countries has been restricted to those with a mean reading score above 450 points. Wales sits below the dashed regression line; this means that Wales has fewer high achieving pupils than one would anticipate given its mean reading score. Specifically, the fitted regression line suggests that around six per cent of pupils will reach PISA Level 5 or 6 in the typical country with a mean PISA reading score of 477; yet only four per cent of 15-year-olds reach this benchmark in Wales. This illustrates how Wales has a low proportion of 15-year-olds with high-level reading skills, even compared to other countries with similar mean scores.

Figure 5.3 The percentage of top-performing pupils in reading compared to mean PISA reading scores: a cross-country analysis



Source: PISA 2015 database.

Notes: The sample of countries included in this figure has been restricted to those with a mean reading score above 450 points.

Key point

21 per cent of 15-year-olds in Wales lack basic reading skills; this is similar to the average across members of the OECD. On the other hand, Wales has a comparatively small proportion of pupils with high-level reading skills.

5.4 How do the PISA reading scores of the ***highest*** achieving pupils in Wales compare to other countries?

13. The previous sub-section highlighted how Wales has a smaller proportion of its pupils reaching the top two PISA achievement levels in reading than the average member of the OECD. We now provide further insight into the proficiency of the highest achieving pupils by comparing the 90th percentile of the reading distribution for Wales to the 90th percentile in other countries. We then consider whether the PISA reading scores of the highest achievers in Wales have changed over the last decade.

Table 5.3 The 90th percentile of PISA 2015 reading scores

(a) Countries more than 20 points ahead of Wales

Country	P90	Country	P90	Country	P90
Singapore	**657***	**China**	**630***	**Luxembourg**	**616***
New Zealand	**643***	**Netherlands**	**630***	**Czech Republic**	**614***
Canada	**642***	**Japan**	**629***	**Switzerland**	**614***
Finland	**640***	**Ireland**	**629***	**Portugal**	**614***
South Korea	**637***	**Sweden**	**625***	**Austria**	**611***
France	**637***	**England**	**625***	**Taiwan**	**611***
Norway	**636***	**United States**	**624***	**Macao**	**610***
Germany	**634***	**Belgium**	**623***	**Denmark**	**608***
Hong Kong	**632***	**Israel**	**621***	**Scotland**	**608***
Australia	**631***	**Slovenia**	**621***	**Russia**	**608***
Estonia	**630***	**Poland**	**617***		

(b) Countries between 10 and 20 points ahead of Wales

Country	P90	Country	P90	Country	P90
Iceland	**607***	**Croatia**	**603***	**Italy**	**602***
Northern Ireland	**605***	**Spain**	**603***		

(c) Countries within 10 points of Wales

Country	P90	Country	P90	Country	P90
Malta	595	Hungary	593	Slovakia	583
Latvia	595	Greece	590	Vietnam	580
Lithuania	593	Wales	588		

(d) Countries between 10 and 20 points behind Wales

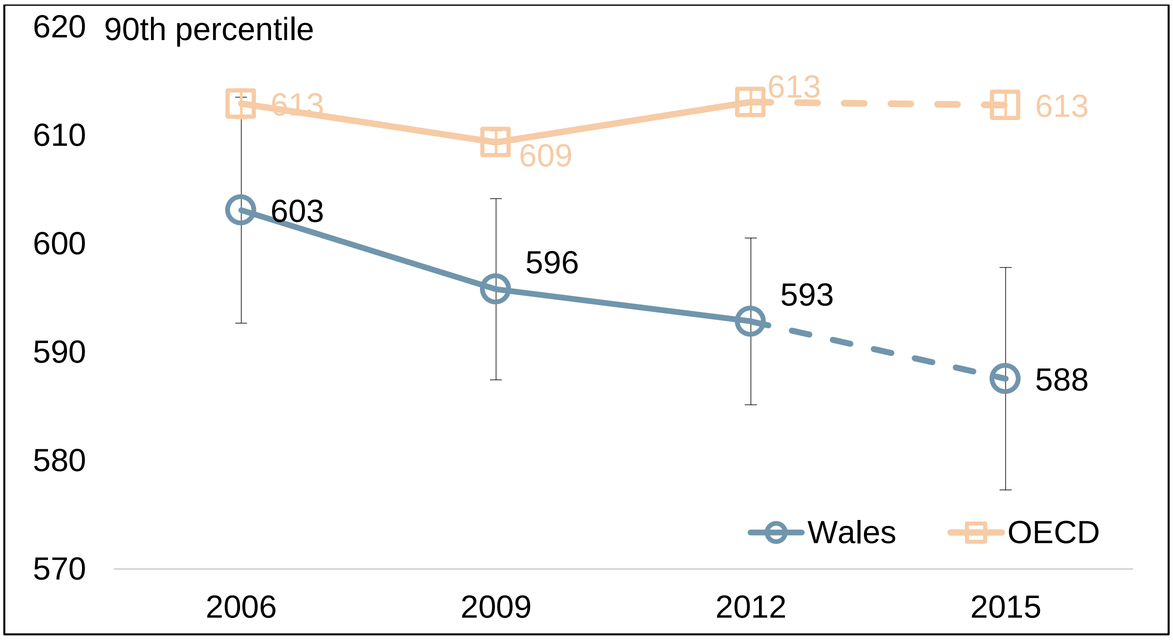
Country	P90	Country	P90	Country	P90
Bulgaria	578	**Chile**	**572***	**United Arab Emirates**	**572***

Note: Bold font with a * indicates significantly different from Wales at the five per cent level. Table does not include countries where the 90th percentile of the reading proficiency distribution is more than 20 points below Wales.

14. Table 5.3 compares the 90th percentile of the PISA reading distribution for Wales to a range of other countries. In 2015, the 90th percentile of the reading proficiency distribution in Wales was 588. This means that the top-performing 10 per cent of 15-year-olds in this country achieved a score of 588 reading test points or

more. There are 32 countries where the value of the 90th percentile is more than 20 points above the value for Wales, with a further five countries where the 90th percentile is between 10 and 20 points higher. Conversely, there are relatively few industrialised nations where the value of the 90th percentile is significantly lower than in Wales. (Turkey, Mexico and Chile are the only members of the OECD where the 90th percentile is lower – see the online data tables for further details). Overall, Table 5.3 illustrates how the reading skills of the highest achieving pupils in Wales is significantly below the skills of the highest achieving pupils in a number of other countries.

Figure 5.4 The 90th percentile of reading scores for Wales: 2006 to 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. OECD average based upon the 'AV09' results presented in the OECD international results Table I.4.4b. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

15. How has the performance of Wales' highest achieving pupils in reading changed over time? Figure 5.4 provides the answer by plotting the 90th percentile of the PISA reading distribution from 2006 to 2015, accompanied by the estimated 95 per cent confidence interval. There is some evidence of a trend, with a steady decline in the 90th percentile over the last decade. In particular, the 90th percentile stood at 603 in 2006, 596 in 2009, 593 in 2012 and 588 in 2015. The point estimate has hence fallen in the last four consecutive rounds. However, it should also be noted that the difference between 2006 and 2015 is not statistically significant at the five per cent level, meaning sampling error remains a possible explanation for this result.

Key point

There is a particularly pronounced gap in reading skills between the highest achieving pupils in Wales and the highest achieving pupils in other countries.

5.5 How do the reading scores of the ***lowest*** achieving pupils in Wales compare to other countries?

16. Although the reading skills of the highest achieving pupils in Wales may be lower than the top performing pupils in other countries, does the same hold true for the lowest achievers? Table 5.4 provides evidence on this matter by comparing the 10^{th} percentile of the PISA reading distribution across countries.

17. The value of the 10^{th} percentile of the reading proficiency distribution in Wales is 368. There are nine countries where the 10^{th} percentile is more than 20 points above the value for Wales, including five East Asian economies (Hong Kong, Singapore, Macao, Vietnam and Japan), along with Ireland (406), Estonia (404), Canada (404) and Finland (401). In a further eight countries, the 10^{th} percentile is between 10 and 20 points above Wales, including South Korea (386), Poland (386) and Northern Ireland (385). However, Table 5.4 also indicates that low achieving pupils in Wales achieve similar PISA reading scores to 15-year-olds in a number of other OECD countries, including England (371), Germany (375), Australia (365) and the United States (364). Likewise, the 10^{th} percentile of the reading distribution in Wales is more than 20 points higher than in a selection of other countries, including France (344), Austria (347) and China (346). Indeed, in more than half of the participating countries, the 10^{th} percentile is at least 10 points lower than the value in Wales. Overall, the position of Wales in this international comparison of low-achievers is somewhat more favourable than the results previously presented for the highest achievers in Table 5.3.

Table 5.4 The 10th percentile of PISA 2015 reading scores

(a) Countries more than 20 points ahead of Wales

Country	P10	Country	P10	Country	P10
Hong Kong	**412***	**Canada**	**404***	**Macao**	**399***
Ireland	**406***	**Finland**	**401***	**Vietnam**	**393***
Estonia	**404***	**Singapore**	**400***	**Japan**	**391***

(b) Countries between 10 and 20 points ahead of Wales

Country	P10	Country	P10	Country	P10
South Korea	**386***	**Denmark**	**383***	**Russia**	**381***
Poland	**386***	**Slovenia**	**382***	Spain	379
Northern Ireland	**385***	**Norway**	**381***		

(c) Countries within 10 points of Wales

Country	P10	Country	P10	Country	P10
Germany	375	New Zealand	368	Sweden	364
Portugal	374	Wales	368	Belgium	360
Latvia	374	Netherlands	368	Switzerland	360
Scotland	373	Croatia	367	Italy	359
England	371	Australia	365		
Taiwan	371	United States	364		

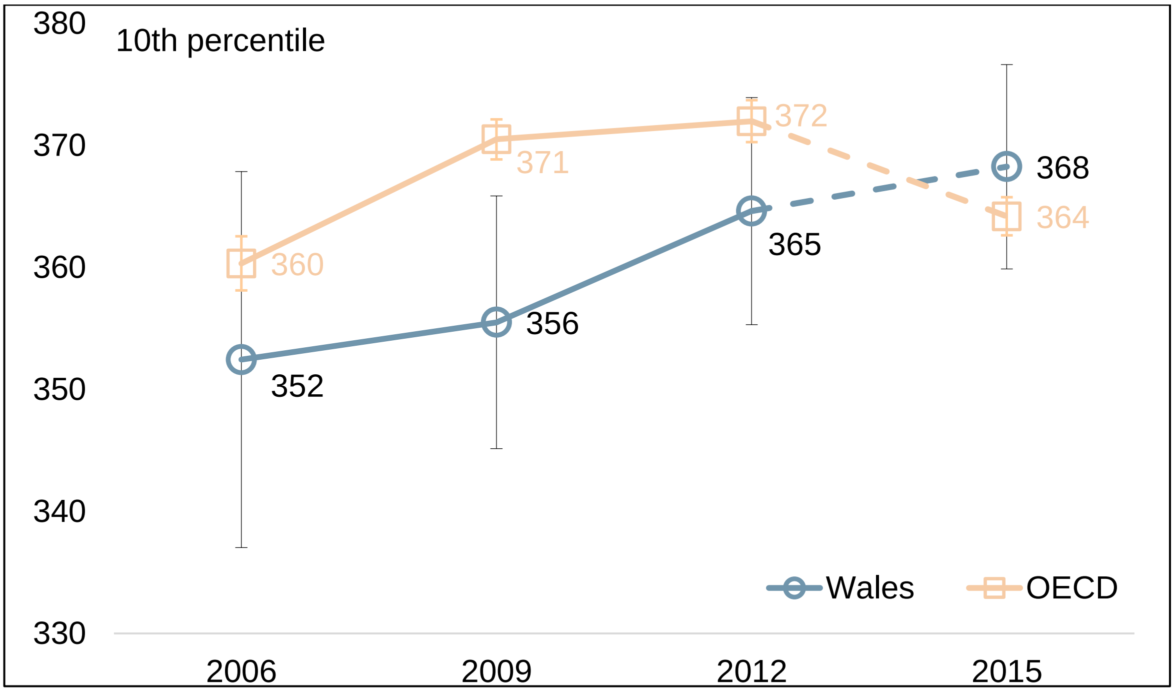
(d) Countries between 10 and 20 points behind Wales

Country	P10	Country	P10
Czech Republic	**352***	**Iceland**	**350***

Source: PISA 2015 database.

Note: * indicates significantly different from Wales at the five per cent level. Table does not include countries where the 10th percentile of the reading distribution is more than 20 points below Wales.

Figure 5.5 The 10th percentile of PISA reading scores for Wales between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: The dashed line between 2012 and 2015 refers to the introduction of computer based testing. Thin black line through each data point refers to the estimated 95 per cent confidence interval. Confidence intervals do not include link error for comparing changes over time. OECD average based upon the 'AV09' results presented in the OECD international results Table I.4.4b. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

18. Figure 5.5 proceeds by considering how the 10th percentile of the PISA reading distribution in Wales has changed since 2006. There is some evidence of a trend, with a steady increase in the 10th percentile since 2006. In particular, the 10th percentile stood at 352 in 2006, 356 in 2009, 365 in 2012 and 368 in 2015. The point estimate has therefore increased in the last four consecutive PISA rounds. However, it should also be noted that the difference between 2006 and 2015 is not statistically significant at the five per cent level, meaning sampling error remains a possible explanation for this result.

Key point

The PISA reading scores of the lowest achieving pupils in Wales are similar to the scores achieved by the lowest achieving pupils in several other industrialised countries.

5.6 How big is the gap between the pupils with the strongest and weakest reading skills? How does Wales compare to other countries in this respect?

19. To conclude this chapter, we consider inequality in 15-year-olds' reading skills, as measured by the difference between the 90th percentile and the 10th percentile. The magnitude of this gap is presented in Table 5.5. For brevity, the sample is restricted to only those countries with a mean PISA reading score above 450 points. The 10 countries with the highest mean PISA reading scores have been highlighted in red.

20. The 90th percentile of the PISA reading distribution in Wales is 588, while the 10th percentile stands at 368. Table 5.5 demonstrates that the gap is therefore 219 test points, equivalent to around seven and a quarter years of schooling. This is smaller than in most other countries included in the comparison (OECD average = 249). Indeed, there are no countries included in Table 5.5 where the difference between the 90th and 10th percentile is significantly smaller than in Wales. Conversely, there are 32 countries where inequality in reading achievement is significantly greater. Consequently, by this metric, Wales seems to be one of the most equal countries in the world in terms of 15-year-olds' reading skills.

Key point

The difference in reading skills between the highest and lowest achieving pupils is smaller in Wales than in most other countries.

Table 5.5 Difference between the highest and lowest achievers in reading

Country	Difference between the 90th and 10th percentile	Difference in years of schooling
Israel	**295***	**9.8 years**
France	**293***	**9.8 years**
China	**283***	**9.4 years**
Luxembourg	**279***	**9.3 years**
New Zealand	**274***	**9.1 years**
Slovakia	**271***	**9.0 years**
Australia	**265***	8.8 years
Austria	**265***	8.8 years
Belgium	**263***	8.8 years
Czech Republic	**262***	8.7 years
Netherlands	**262***	8.7 years
Sweden	**262***	8.7 years
United States	**259***	8.6 years
Germany	**258***	8.6 years
Singapore	**257***	8.6 years
Iceland	**256***	8.5 years
Greece	**256***	8.5 years
Norway	**255***	8.5 years
Hungary	**255***	8.5 years
Switzerland	**254***	8.5 years
England	**254***	8.5 years
South Korea	**251***	8.4 years
Lithuania	**246***	8.2 years
Italy	**244***	8.1 years
Taiwan	**240***	8.0 years
Portugal	**240***	8.0 **years**
Finland	**239***	**8.0 years**
Slovenia	**239***	**8.0 years**
Canada	**238***	**7.9 years**
Japan	**238***	**7.9 years**
Croatia	**237***	**7.9 years**
Scotland	**235***	**7.8 years**
Poland	231	7.7 years
Chile	229	7.6 years
Russia	227	7.6 years
Estonia	226	7.5 years
Denmark	225	7.5 years
Spain	224	7.5 years
Ireland	222	7.4 years
Latvia	221	7.4 years
Hong Kong	220	7.3 years
Northern Ireland	220	7.3 years
Wales	219	7.3 years
Macao	212	7.1 years
Vietnam	187	6.2 years

Source: PISA 2015 database.

Note: Bold with a * indicates statistically significant difference compared to Wales at the five per cent level. Table only includes countries where the mean PISA reading score is above 450.

Chapter 6. Variation in PISA scores by pupil characteristics

- There is no statistically significant gender difference in Wales on the overall PISA science scale. However, Welsh boys are (on average) stronger than girls in particular areas of scientific literacy, including the physical scientific system and the ability to explain phenomena scientifically.
- The gender gap in 15-year-olds' reading skills is smaller in Wales than in most other countries. However, this is due to the low reading skills of Welsh girls (relative to the reading skills of girls in other parts of the world).
- Family background has a smaller impact upon pupils' achievement in Wales than in most other countries. 15-year-olds from high socio-economic backgrounds in Wales achieve significantly lower PISA scores than socio-economically advantaged pupils in other OECD countries.
- Around one-in-four pupils in Wales overcome a disadvantaged socio-economic background to achieve a top score on the PISA science test. There is no evidence that countries with selective schooling systems have a greater proportion of resilient pupils.
- Differences in average PISA scores between pupils from immigrant backgrounds and pupils whose family were born in the UK are not statistically significant at the five per cent level.
- Pupils who took the Welsh language version of the PISA science test achieved lower scores than pupils who completed the test in English.

1. This chapter explores differences in pupils' PISA test scores according to selected demographic characteristics – gender, socioeconomic status, immigrant status and Welsh or English language. Variation in achievement by these characteristics is a key policy concern in Wales, where there is growing emphasis on reducing educational inequalities. Although we already know much about differences in achievement by these characteristics from national GCSE examination data, PISA provides an opportunity to consider the magnitude of these gaps in a comparative context. For instance, although there are socio-economic disparities in educational achievement in Wales, are these disparities bigger or smaller in Wales than elsewhere? PISA also allows us to examine differences between demographic groups using a rather different measure to GCSEs, one with a greater emphasis upon young people's 'functional skills' (see Box 1.1 for further details).

2. In summary, this chapter will address the following questions:

- *How do boys and girls in Wales perform on the PISA science, mathematics and reading test? Is this gender gap bigger or smaller in Wales than in other countries?*
- *What is the 'strength' and 'impact' of socio-economic status upon pupils' PISA test scores? How does Wales compare to other countries in this respect?*
- *What proportion of young people in Wales are classified as 'resilient' – overcoming the odds to achieve highly in science, despite a disadvantaged socio-economic background?*
- *Do immigrants in Wales achieve lower average PISA test scores than young people who were born in the UK?*
- *Do PISA scores differ between pupils who took the PISA test in English versus those who took the test in Welsh?*

3. Due to limited sample sizes for certain groups, caution is needed when interpreting some results. Only 339 pupils took the Welsh language version of the PISA 2015 test. Likewise, only 178 pupils are first generation immigrants (meaning they were born outside of the UK). These results will be subject to a degree of uncertainty, with relatively wide margins-of-error.

6.1 How big is the gender gap in PISA test scores?

4. In GCSE examinations, girls tend to achieve higher grades than boys in most subject areas. For instance, in the 2014/15 academic year, 85 per cent of girls received an A*-C grade in GCSE science, compared to 83 per cent of boys[32]. The difference between genders is bigger for GCSE English or Welsh first language (78 per cent A*-C for girls versus 62 per cent for boys), though smaller for GCSE mathematics (65 per cent A*-C for girls versus 64 per cent for boys). Yet the PISA assessment differs from GCSE examinations in a number of ways, including the precise type of knowledge and skill each is attempting to measure (see Box 1.1 for further details). This raises the question, how does the gender gap in PISA test scores in Wales compare to the gender gap in GCSE grades? Moreover, how does the gender gap in Wales, as measured by PISA, compare to other countries?

5. Evidence on this matter is presented in Table 6.1. This documents the gender gap in average PISA test scores, with positive figures indicating a higher mean for boys than girls. Estimates are presented for countries with a mean PISA science score above 450 points.

6. In Wales, the mean science score for boys (487) is above the mean for girls (482), though this five point difference is not statistically significant at the five per cent level. Nevertheless, this is slightly different to the pattern observed for science GCSEs, where girls achieve slightly higher grades than boys (85 per cent versus 83 per cent A*-C). It is also reasonably similar to the results for PISA 2006, 2009 and 2012, where the difference in mean scores between boys and girls was consistently around 10 test points (in favour of boys).

7. Table 6.1 suggests the estimated gender difference in 15-year-olds science skills in Wales is of similar magnitude to that in most other countries. Indeed, the gender difference in science scores is rarely greater than 10 points and is typically statistically insignificant at the five per cent level. There is also little evidence of a consistent pattern emerging across the 10 countries with the highest average PISA science scores. For instance, in Finland and Macao, girls achieve significantly higher average science scores than boys, while in China and Japan, the opposite holds true (scores for boys are higher for boys than for girls). Yet in others (e.g. Singapore and Taiwan) the situation is very much the same as in Wales, with little difference in science achievement by gender.

[32] StatsWales: Examination achievements of pupils aged 15 by gender 2014/15

Table 6.1 Difference in mean PISA test scores between boys and girls

Science		Maths		Reading	
Country	**Gap**	**Country**	**Gap**	**Country**	**Gap**
Austria	**19***	Austria	**27***	Wales	**-11***
Italy	**17***	Italy	**20***	Chile	**-12***
Japan	**14***	Germany	**17***	Ireland	**-12***
Belgium	**12***	Ireland	**16***	Japan	**-13***
Ireland	**11***	Spain	**16***	Northern Ireland	**-14***
Germany	**10***	Belgium	**14***	Belgium	**-16***
Portugal	**10***	Japan	**14***	Italy	**-16***
China	**9***	Croatia	**13***	China	**-16***
Czech Republic	**9***	England	**12***	Portugal	**-17***
Luxembourg	**8***	Switzerland	**12***	United States	**-20***
United States	**7***	Poland	**11***	Austria	**-20***
Spain	**7***	Luxembourg	**11***	Singapore	**-20***
Singapore	**6***	Portugal	**10***	Spain	**-20***
Poland	**6***	Wales	**10***	Germany	**-21***
Switzerland	6	Denmark	9	Scotland	**-21***
Denmark	6	Canada	9	Luxembourg	**-21***
Croatia	6	New Zealand	9	Denmark	**-22***
Wales	5	United States	9	Israel	**-23***
New Zealand	5	Israel	8	England	**-23***
Taiwan	4	Hungary	8	Netherlands	**-24***
Russia	4	Czech Republic	7	Hungary	**-25***
Israel	4	Scotland	7	Vietnam	**-25***
Netherlands	4	Northern Ireland	7	Taiwan	**-25***
Estonia	3	France	6	Switzerland	**-25***
Hungary	3	Russia	6	Czech Republic	**-26***
Norway	3	China	6	Russia	**-26***
Northern Ireland	3	Australia	6	Canada	**-26***
Australia	2	Slovakia	6	Croatia	**-26***
France	2	Taiwan	6	Estonia	**-28***
Scotland	1	Estonia	5	Hong Kong	**-28***
Canada	1	Slovenia	4	France	**-29***
England	0	Netherlands	2	Poland	**-29***
Slovakia	-1	Hong Kong	2	Macao	**-32***
Hong Kong	-1	Greece	0	Australia	**-32***
Vietnam	-3	Singapore	0	New Zealand	**-32***
Iceland	-3	Iceland	-1	Slovakia	**-36***
Sweden	-5	Lithuania	-1	Greece	**-37***
Slovenia	**-6***	Latvia	-2	Lithuania	**-39***
Lithuania	**-7***	Sweden	-2	Sweden	**-39***
Macao	**-8***	Norway	-2	Norway	**-40***
Greece	**-9***	Vietnam	-3	South Korea	**-41***
South Korea	**-10***	Malta	-4	Iceland	**-42***
Latvia	**-11***	South Korea	-7	Latvia	**-42***
Malta	**-11***	Finland	**-8***	Slovenia	**-43***
Finland	**-19***	Macao	**-8***	Finland	**-47***

Source: PISA 2015 database.

Notes: Table restricted to those countries with a mean science score greater than 450 test points. Positive figures refer to higher average score for boys than girls. Gender differences that are statistically significant indicated by bold font with a *.

8. Although there may be little evidence of gender differences on the overall PISA science scale in Wales, there could be marked differences within some particular sub-domains. For instance, might boys achieve higher average scores in one area of science (e.g. understanding physical systems) with girls achieving higher average scores in another (e.g. knowledge of living systems)? Table 6.2 provides insight into this matter by presenting average PISA scores by gender for each of the eight separate science skills that the PISA test examines.

Table 6.2 Gender differences in PISA science scores by sub-domain in Wales

		Girls mean	Boys Mean	Gender gap
System	Physical	481	491	**+9***
	Living	482	482	0
	Earth and Space science	482	488	+7
Competency	Explain phenomena scientifically	480	492	**+12***
	Evaluate and design	481	481	0
	Interpret data and evidence	483	484	1
Knowledge	Content knowledge	480	491	**+11***
	Procedural and epistemic	483	484	1

Source: PISA 2015 database.

Notes: Positive figures refer to higher average score for boys than girls. Gender differences that are statistically significant indicated by bold font with a *. Difference between boys and girls columns may not equal the gender gap column due to rounding.

9. Across the three science systems, there are signs of some important gender differences in Wales. In particular, the mean score for boys (491) on the physical scientific system is around nine points higher than the mean score for girls (481), with the difference statistically significant at the five per cent level. There is also a difference of seven points (in favour of boys) in the earth and space science domain, though this does not quite reach statistical significance at the five per cent level[33]. A similar finding emerges in terms of science competencies; in one particular area ('explaining phenomena scientifically') the mean for boys (492) is significantly above the mean for girls (480). Finally, although there is little difference between genders in pupils' epistemic and procedural knowledge, there is a statistically significant difference between boys and girls in their knowledge of science content (mean score of 491 for boys versus 480 for girls). It therefore seems that, although gender differences in science achievement overall are relatively small in Wales, there are a

[33] The online data tables provide further details by illustrating how Wales compares to other countries in terms of gender differences across these three science systems.

few specific areas (namely physical sciences, explaining phenomena scientifically and content knowledge) where boys are more proficient (on average) than girls.

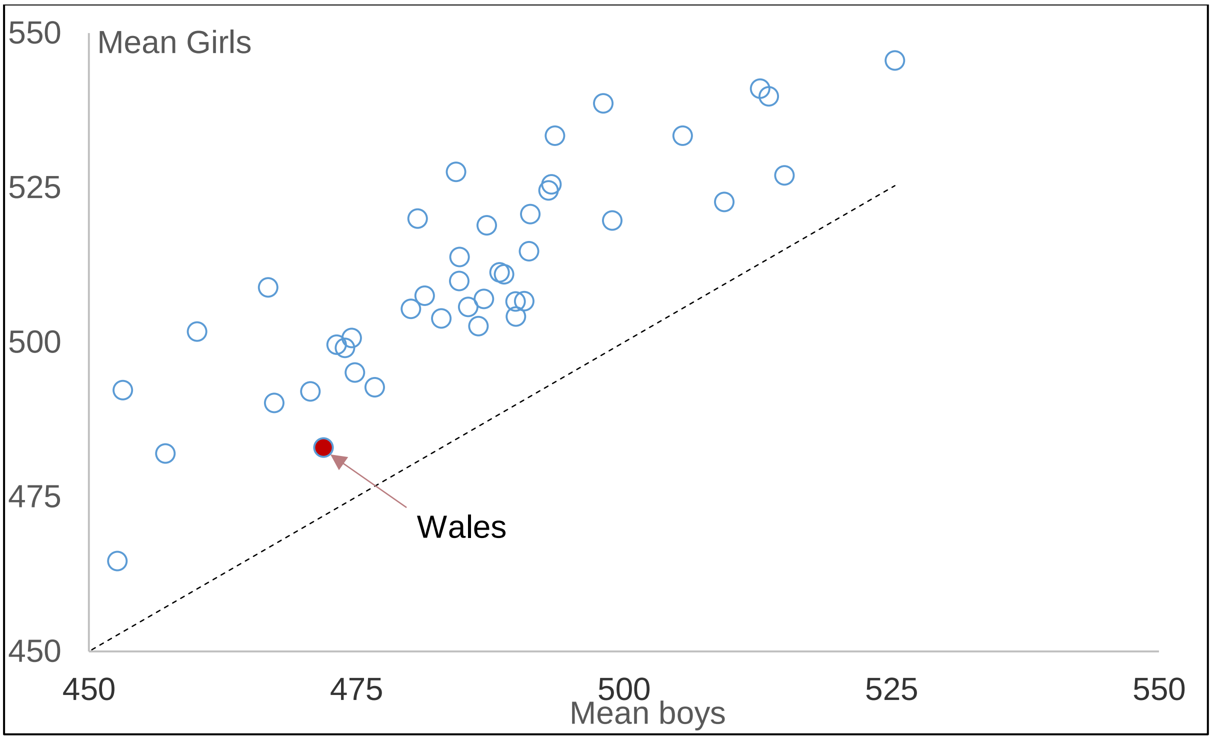
10. Returning to Table 6.1, the middle columns highlights gender differences within the PISA mathematics domain. In most countries the average PISA mathematics scores are higher for boys than for girls. The gender gap in Wales is equal to 10 PISA test points; boys achieve a mean score of 483 compared to 473 for girls. Hence the difference in mathematics performance between boys and girls is quite pronounced in this country, though of a reasonably similar magnitude to many of the other countries that have been included in this cross-national comparison (the gender gap in mathematics is, on average, seven test points in favour of boys across the countries included in Table 6.1). This is a somewhat different pattern to that observed for GCSE mathematics grades in Wales, where the proportion who achieve an A* to C grade is very similar for boys (64 per cent) and girls (65 per cent). On the other hand, the gender gap in mathematics in PISA 2012 was of a similar magnitude (a nine point difference in favour of boys).

11. The final two columns of Table 6.1 provide analogous results for gender differences in pupils' reading skills. In every country, the average PISA reading score for girls is higher than the average score for boys. The average magnitude of this difference across members of the OECD is approximately 27 test points. This pattern is also observed in Wales, with the mean PISA reading score for boys (472 points) significantly below the score for girls (483 points). It is also consistent with GCSE results, where 78 per cent of girls obtain an A*-C grade in English/Welsh language compared to only 62 per cent of boys. However, it is also notable how the gender gap in pupils' reading skills (11 points) is substantially lower in Wales than in most other countries. Indeed, no other country included in this comparison has a significantly smaller gender gap in pupils' reading skills than Wales.

12. Figure 6.1 provides further insight into this issue. It plots the average PISA reading scores for boys along the horizontal axis, with average scores for girls along the vertical axis. The dashed 45 degree line then indicates where the two scores are equal (i.e. where there is no gender gap). From a Welsh perspective, there are two points of particular note. First, although all countries are above the 45 degree line, Wales sits closer to it than most. This further highlights the comparatively small gender gap that exists in Welsh pupils' reading skills. Second, from an international comparative perspective, it illustrates the relatively low performance of Welsh girls on the PISA reading test (which, in turn, is leading to the small gender gap). In particular, out of all the countries with an overall mean reading score above 450 points, only in Chile do girls have a substantially lower average level of reading

proficiency. Thus, although Welsh boys may be worse readers than Welsh girls in an absolute sense, the comparatively low reading skills of Welsh girls stand out as a particular challenge facing Wales from an international comparative perspective.

Figure 6.1 The average PISA reading score for boys versus girls



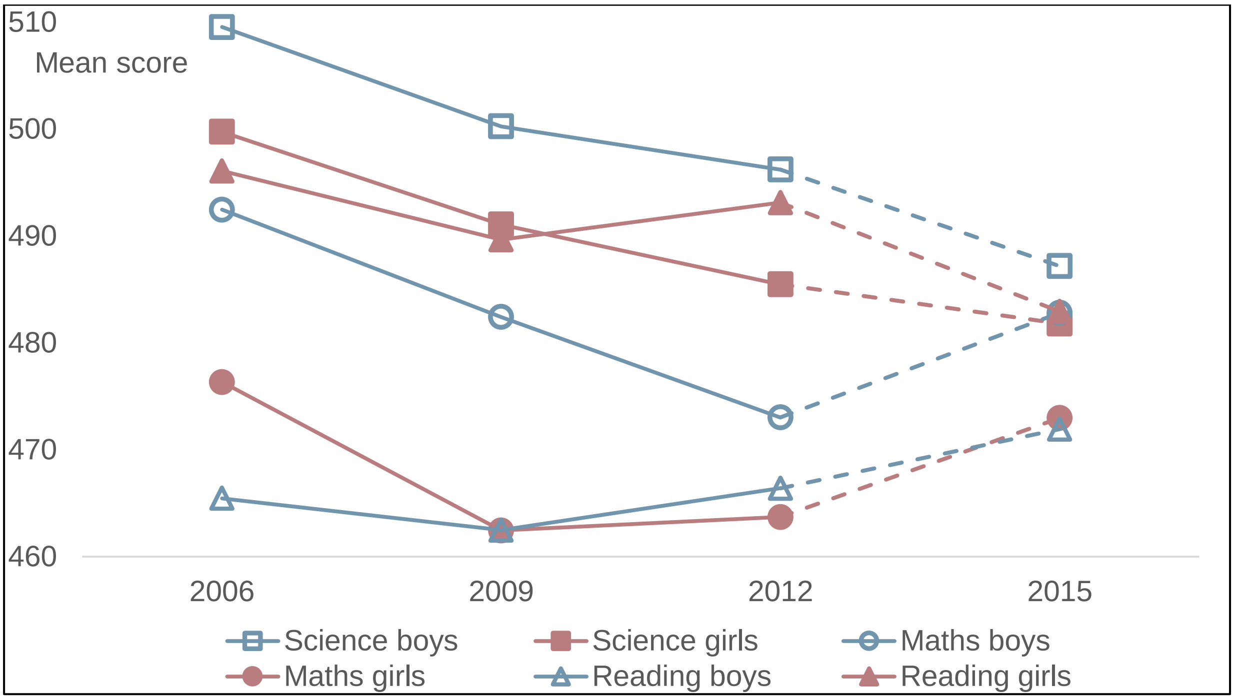
Source: PISA 2015 database.

Note: Table restricted to those countries with a mean science score greater than 450 test points.

13. To conclude this sub-section, Figure 6.2 illustrates how average PISA science, mathematics and reading test scores for boys and girls have changed since 2006. Solid red markers provide the results for girls and hollow blue markers the results for boys. The first interesting feature to highlight is how the collection of data points in Figure 6.2 were much more spread out in 2006 than they are in 2015 (i.e. they are now much closer together). This is the result of a general shrinking of differences in Welsh pupils' skills across the various PISA domains (i.e. in 2006 Welsh pupils had a clear comparative advantage in science relative to mathematics, but in 2015 this is no longer the case) accompanied by a decline in the gender gap in specific subject areas. The clearest example of the latter is in reading. During the 2006 to 2012 PISA cycles, Welsh girls held a 25 to 30 point lead over Welsh boys in this domain. Yet the difference has shrunk to 11 points in 2015, which Figure 6.2 illustrates is due to the combined effect of a six point increase in the mean reading score for boys (466 in 2012 versus 472 in 2015) and a 10 point decrease for girls

(from 493 to 483). However, this change in results in 2015 should be carefully interpreted. A number of possible explanations exist, including sampling error, the move to computer-based assessment, changes to the scoring procedures, in addition to a genuine substantive change in boys and girls reading skills (recall the discussion in chapter 1 for further details).

Figure 6.2 Average PISA scores for boys and girls in Wales since 2006



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

Key points

There is no statistically significant gender difference in Wales on the overall PISA science scale. However, Welsh boys are (on average) stronger than girls in particular areas of scientific literacy, including the physical scientific system and the ability to explain phenomena scientifically.

The gender gap in 15-year-olds' reading skills is smaller in Wales than in most other countries. However, this is due to the low reading skills of Welsh girls (relative to reading skills of girls in other parts of the world).

6.2 How pronounced is the relationship between socio-economic status and pupils’ PISA test scores?

14. The relationship between family background and young people’s academic achievement has long been recognised as a challenge facing the Welsh education system. A wealth of previous research has documented the achievement gap between young people from socio-economically advantaged and disadvantaged backgrounds, with a widespread belief that this is hindering the prospects of greater social mobility. This sub-section therefore provides evidence on the relationship between socio-economic status and the PISA test scores of 15-year-olds in Wales, and how this compares to other countries. It will therefore illustrate the challenge Wales faces in narrowing educational inequalities by family background.

15. The main measure of socio-economic status in PISA is the Economic, Social and Cultural Status (ESCS) index. This is a continuous index that has been derived by the OECD based upon pupils’ responses to the background questionnaire. It encompasses the following information:

- Maternal and paternal education
- Maternal and paternal occupation
- Household possessions

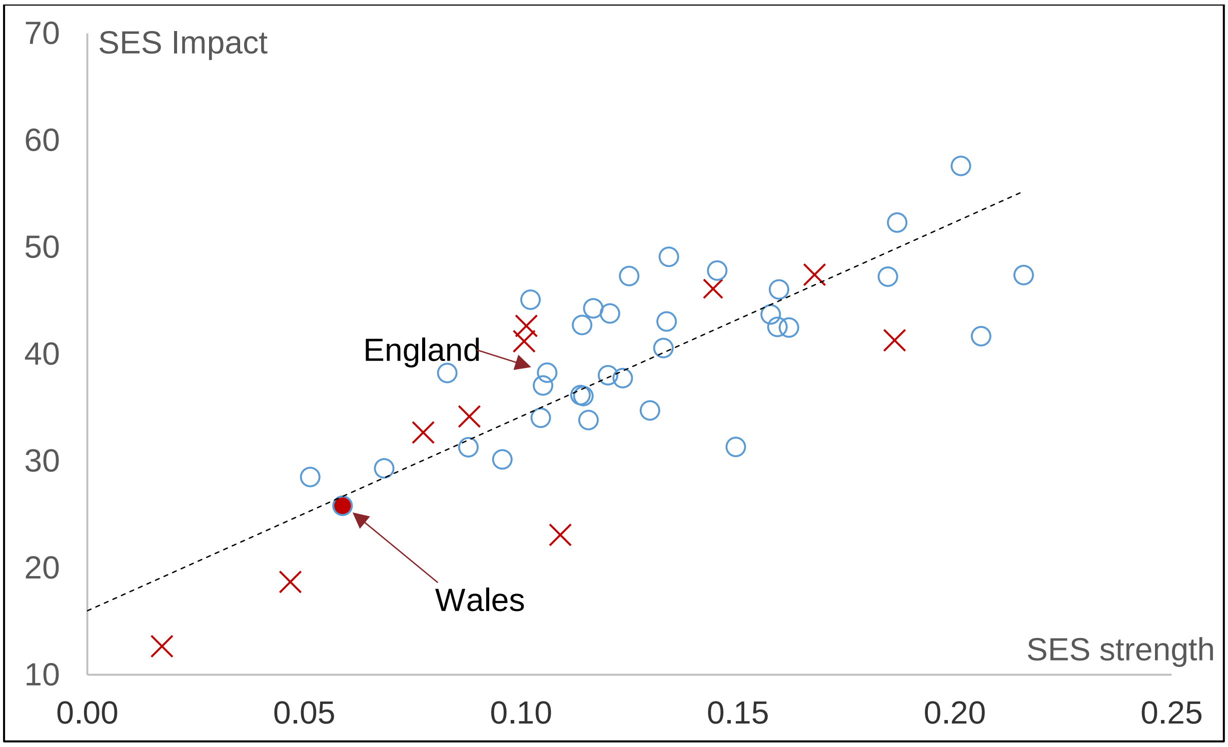
The OECD use this measure to estimate the *impact* socio-economic status has upon achievement and the *strength* of this relationship.

16. The OECD measure the *impact* of the relationship between pupil’s socio-economic backgrounds (ESCS score) and their attainment in terms of the steepness of the socio-economic gradient for each participating country. Specifically, these figures refer to the change in PISA science scores when comparing the median young person to a young person at approximately the 85th ESCS percentile[34]. Low values indicate that socio-economic background has less impact upon pupil attainment; high values indicate socio-economic background has more impact upon pupil attainment. In Wales, the impact of socio-economic status upon pupils’ science scores is estimated to be around 25 points. The average impact of socioeconomic status upon pupils’ science scores amongst OECD countries is 38 test points.

[34] In other words, these figures refer to the change in PISA science scores per each international standard deviation increase in the ESCS index. It is the parameter estimate generated by a simple Ordinary Least Squares regression of the ESCS index upon PISA test scores.

17. The OECD measure the *strength* of the relationship between pupil's socio-economic backgrounds and their attainment in terms of the percentage of variance in PISA scores explained by the pupils' backgrounds. The key difference is that whereas the 'impact' measure is influenced by the dispersion of the ESCS index relative to PISA test scores, the 'strength' measure is not. Low values indicate that pupil attainment varies widely, even for pupils with similar backgrounds, while high values indicate that pupil attainment is strongly determined by background. In Wales, approximately six per cent of the variation in pupils' science achievement can be explained by the ESCS index. The average amount of variation in pupil's science achievement explained by the ESCS index amongst OECD countries is 13 per cent.

Figure 6.3 The 'impact' and 'strength' of the relationship between socio-economic status and PISA science scores



Source: PISA 2015 database.

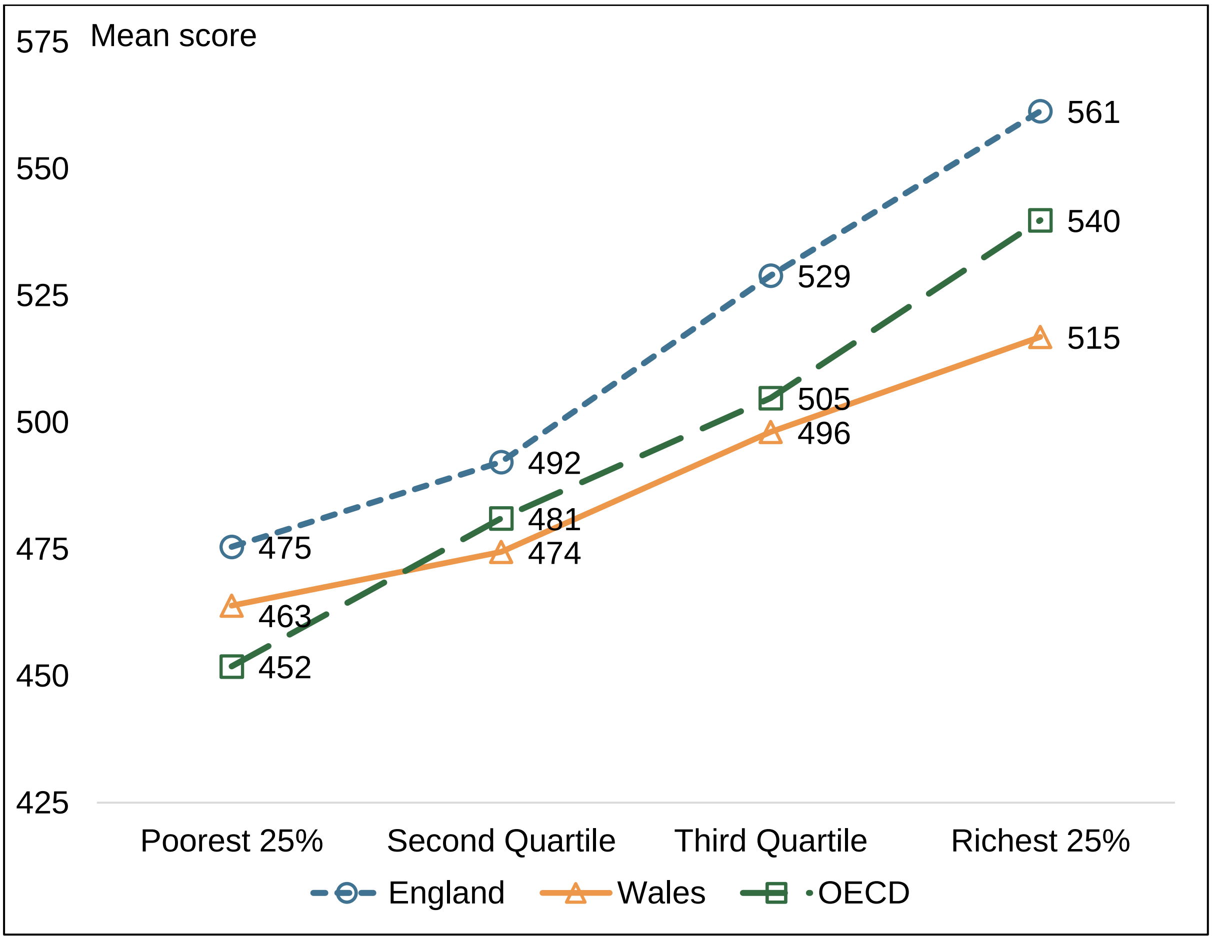
Notes: 'Impact' refers to the bivariate relationship between the ESCS index and PISA science scores, estimated in PISA test points using OLS regression. 'Strength' refers to the percent of variance in PISA science scores that is explained by the ESCS index. Sample of countries restricted to those with a mean science score above 450 points. Spain and Latvia have been excluded due to recoding of the ESCS index required at the time of writing.

18. These two measures of socio-economic inequality in pupils' science achievement are plotted against one another in Figure 6.3. Countries towards the top right are where family background matters a lot for pupils' science achievement, while family background has less of an influence in those countries towards the

bottom left. Wales is very much towards the lower left hand corner of this graph, indicating that there is *less* socio-economic inequality in Welsh pupils' PISA test scores than in most other countries. This includes the rest of the UK, where inequality in pupils' achievement is significantly higher. Moreover, this finding is not specific to science; similar findings emerge regarding the link between family background and pupils' achievement in reading and mathematics as well (see the online data tables for further details). Consequently, Figure 6.3 highlights how, in terms of PISA outcomes, Wales is one of the most equitable countries in the world.

19. Figure 6.4 provides further detail to this result. We have now divided pupils into four equal groups (quartiles) within each country according to their ESCS index score. These groups, from the least advantaged (bottom quartile) up to the most advantaged (top quartile), run along the horizontal axis. Mean PISA scores for each quartile are then plotted along the vertical axis. Results are presented for Wales, England and the OECD average.

Figure 6.4 Average PISA science scores in Wales by national quartiles of the ESCS index



Source: PISA 2015 database.

20. Figure 6.4 reveals an important finding; differences in mean scores between Wales, England and the OECD average are much more pronounced for socio-economically advantaged pupils than for socio-economically disadvantaged pupils. For instance, pupils from low socio-economic backgrounds in Wales score, on average, 463 on the PISA science test. This is lower than the equivalent socio-economic group in England (475) but is actually above the average across OECD countries (452). On the other hand, the mean score for the most advantaged pupils in Wales (515) is more than half a year of schooling below the OECD average (540) and more than a year of schooling behind the value for the equivalent socio-economic group in England (561). There are two related implications of this result. First, the key reason why socio-economic inequality is lower in Wales than in other countries is the comparatively weak performance of the top socio-economic quartile (relative to other countries). Second, the comparatively low score of socio-economically advantaged pupils in Wales is one of the major reasons why the overall mean score for Wales in Table 2.1 is lower than the mean score for England and several other members of the OECD.

21. An alternative measure of socio-economic disadvantage that is often used in Wales is eligibility for Free School Meals (FSM). Table 6.3 therefore considers how average PISA test scores vary by this characteristic. Unsurprisingly, there are statistically significant differences between FSM-eligible and FSM-ineligible pupils within each domain. In science, FSM pupils (446 points) score, on average, 43 PISA test points below their non-FSM peers (489 points). This difference is equivalent to more than an additional year of schooling. A similarly sized gap between FSM eligible and non-FSM pupils exists for reading (40 points) and mathematics (43 points).

Table 6.3 The relationship between FSM eligibility and PISA test scores

	Not eligible for FSM	Eligible for FSM
Science	489	**446***
Mathematics	483	**440***
Reading	482	**441***
Observations	**2,915**	**445**

Source: PISA 2015 – school census matched database.

Notes: Bold font with a * indicates statistically significant difference from the 'not FSM' group at the five per cent level. Estimates presented for pupils where PISA has been successfully linked to the school census.

Key points

Family background has a smaller impact upon pupils' achievement in Wales than in most other countries.

15-year-olds from high socio-economic backgrounds in Wales achieve significantly lower PISA scores than similarly socio-economically advantaged pupils in other OECD countries.

6.3 To what extent do socio-economically disadvantaged pupils succeed against the odds?

22. A number of studies have highlighted the challenges socio-economically disadvantaged young people face when trying to access professional jobs[35]. Many believe that improving the educational achievement of young people from low income backgrounds is key to breaking this glass ceiling[36] – and, in particular, increasing the proportion of disadvantaged pupils who achieve the highest grades. At the same time, there remains some debate as to whether comprehensive or selective (grammar style) schooling systems are more effective at reaching this goal. This sub-section provides some descriptive evidence on these issues. Specifically, it documents the proportion of socio-economically disadvantaged 15-year-olds in Wales who succeed in PISA against the odds (see Box 6.1), and compares this to the situation in other countries - particularly those with a more selective 'grammar' style education system.

Box 6.1 The OECD definition of 'resilience'

A pupil is classified as resilient if he or she is in the bottom quarter of the PISA index of economic, social and cultural status (ESCS) in the country of assessment and performs in the top quarter of pupils in the focus subject (science in PISA 2015) among all countries, after accounting for socio-economic status. It therefore captures the proportion of pupils who are amongst the most socio-economically disadvantaged within their country, but who are amongst the highest performing 15-year-olds in science internationally.

23. In Table 6.4, we document the proportion of resilient pupils in countries where the mean science score is above 450 points. In Wales, over a quarter (29 per cent) of pupils from low socio-economic backgrounds are classified as 'resilient'. This is similar to countries like New Zealand (30 per cent), the Netherlands (31 per cent)

[35] See Macmillan et al. (2015).
[36] Economic and Social Research Council (2012).

and Ireland (30 per cent). However, it is lower than in several East Asian nations, which tend to dominate the top of Table 6.4. Indeed, eight of the top 10 countries with the greatest proportion of resilient pupils are within East Asia (Finland and Estonia are the exceptions). Moreover, the fact that the majority of disadvantaged pupils in Vietnam (76 per cent), Macao (65 per cent) and Hong Kong (62 per cent) are classified as resilient is particularly striking. Likewise, it is notable how all of the 10 countries with the highest average PISA science scores have a comparatively large proportion of resilient pupils (these are the countries highlighted in orange).

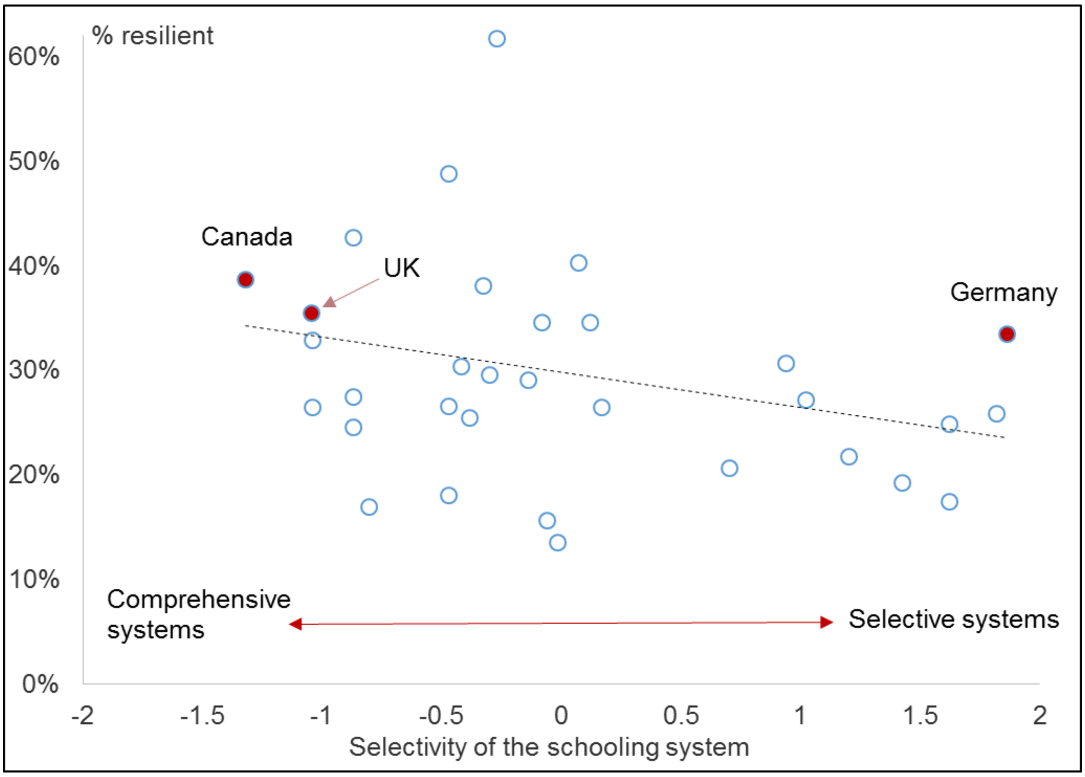
Table 6.4 The proportion of resilient pupils across countries

Country	Percentage of resilient pupils	Country	Percentage of resilient pupils
Vietnam	76%	Switzerland	29%
Macao	65%	Wales	29%
Hong Kong	62%	Denmark	28%
Singapore	49%	Scotland	27%
Japan	49%	Belgium	27%
Estonia	48%	France	27%
Taiwan	46%	Italy	27%
China	45%	Norway	26%
Finland	43%	Austria	26%
South Korea	40%	Russia	26%
Canada	39%	Czech Republic	25%
Portugal	38%	Sweden	25%
England	36%	Croatia	24%
Slovenia	35%	Lithuania	23%
Poland	35%	Malta	22%
Germany	34%	Luxembourg	21%
Australia	33%	Hungary	19%
United States	32%	Greece	18%
Netherlands	31%	Slovakia	18%
New Zealand	30%	Iceland	17%
Northern Ireland	30%	Israel	16%
Ireland	30%		

Source: PISA 2015 database.

Notes: The sample of countries has been restricted to those with an average PISA science score greater than 450 points. Countries highlighted in red are the 10 with the highest average PISA science scores. Spain and Latvia have been excluded due to recoding of the ESCS data required at the time of writing.

Figure 6.5 The proportion of 'resilient' pupils in a country compared to the academic selectivity of its secondary-schooling system



Source: PISA 2015 database and Bol et al. (2014).

Notes: Sample restricted to the countries included in Bol et al. (2014). Spain and Latvia have been excluded due to recoding of the ESCS index required at the time of writing.

24. In debates about the pros and cons of grammar schools, it is often suggested that they may help disadvantaged young people to excel academically and overcome their low socio-economic background. Evidence from PISA can help to guide this debate by illustrating how the proportion of resilient pupils varies across countries. Specifically, do countries with more selective secondary education systems have more resilient pupils? This is the focus of Figure 6.5. The vertical axis plots the proportion of 15-year-olds in each country who have been classified as 'resilient' by the OECD (following the definition in Box 6.1). The horizontal axis provides an index of the selectivity of schooling-systems across the world[37]. Higher values on this index indicate greater segregation of 15-year-olds into different types of school based upon their prior academic achievements[38]. Note that Figure 6.5 has been restricted to the 34 countries included in the study by Bol et al. (2014), and that

[37] This information has been drawn from Bol et al. (2014).
[38] Countries with a comprehensive schooling system, such as Finland and Norway, are therefore towards the left-hand side of this graph. In contrast, countries like Germany, where early academic selection is common, are towards the right.

the United Kingdom has been treated here as a single entity (rather than as separate data points for England, Scotland, Northern Ireland and Wales)[39].

25. There is little evidence of an association between the selectivity of the secondary education system and the chances of young people from disadvantaged backgrounds succeeding academically against the odds. Rather, if anything, the opposite may hold true, with the downward sloping regression line indicating a weak, negative relationship (i.e. countries with more academic selection into secondary schools have fewer resilient pupils). For instance, the proportion of resilient pupils in countries like the UK and Canada (where most pupils are within a non-selective comprehensive system) is similar to countries like Germany (where the secondary education system is highly selective). Consequently, evidence from PISA provides little support for the notion that pupils from disadvantaged backgrounds are more likely to succeed if they live in a country with an academically selective secondary education system.

Key points

Around 29 per cent of socio-economically disadvantaged pupils in Wales are classified as 'resilient'.

There is no evidence that countries with selective schooling systems have a greater proportion of resilient pupils.

6.4 Do immigrants in Wales achieve lower PISA test scores than their peers who were born in the UK?

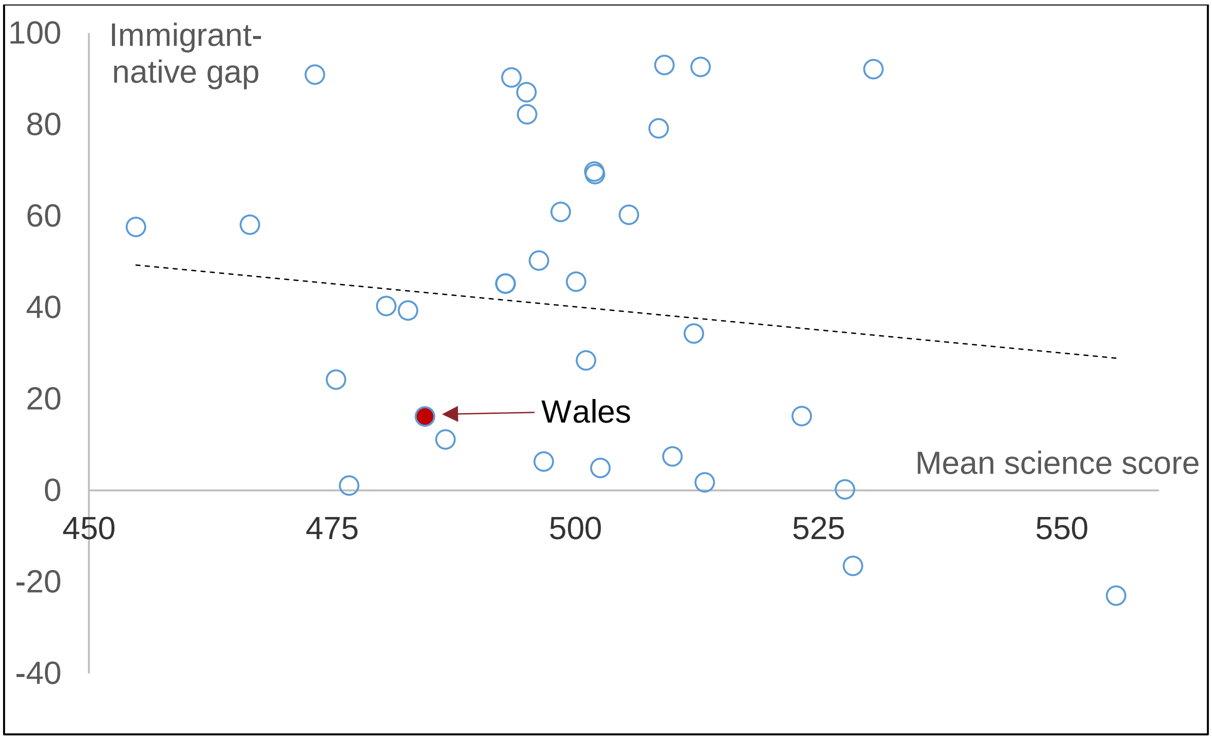
26. Since 2000, net migration into the United Kingdom has totalled approximately 250,000 individuals per year[40]. The increase in the number of Eastern Europeans now living in the UK has been well documented[41], following earlier waves of migration from India and Pakistan in the 1950s and 1960s. Consequently, almost one-in-twelve (eight per cent) 15-year-olds in Wales are now classified as either a first or second generation immigrant (meaning either they or their parent were born outside of the UK)[42].

[39] This has been done as the information on school-system selectivity in Bol et al. (2014) is only provided for the United Kingdom as a whole, and not separately for England, Northern Ireland, Scotland and Wales.
[40] ONS (2015).
[41] ONS (2015).
[42] PISA 2015 database.

27. There has been much debate about the impact such migration has upon public services, including the education system. While popular opinion has focused upon the strain that this could place upon resources[43], and the challenges that this then poses for teachers[44], others have suggested that there is no link between the number of migrant pupils in a school-system and its level of performance[45].

28. This then raises the question, how did immigrant pupils in Wales perform on the PISA test? Figure 6.6 provides evidence on this matter for science. This plots the size of the native-immigrant achievement gap in each country (vertical axis) against the average PISA science score (horizontal axis). Note that the sample of countries included in the graph has been restricted to those with (a) at least 50 pupils identified as first-generation immigrants (b) a mean PISA science score greater than 450 points.

Figure 6.6 The native-immigrant gap in PISA science scores across countries



Source: PISA 2015 database.

Notes: 'Immigrants' includes first generation immigrants only.

[43] See Reynolds (2008) for a discussion
[44] See Eleftheriou-Smith (2014).
[45] OECD (2015) and Coughlan (2015).

29. In the vast majority of countries, average PISA science scores are lower for immigrants than for country natives. Therefore, although the mean science score of first-generation immigrants in Wales (471 points) is below the mean score for individuals born inside the UK (488 points), this is consistent with the pattern observed in most other countries included in this comparison. Indeed, Wales is actually towards the bottom of Figure 6.6, with a much more pronounced difference in average PISA science scores between natives and immigrants existing elsewhere in the world (e.g. in Scandinavian countries such as Sweden, Norway and Denmark there is a difference of more than 60 test points). Moreover, there is no evidence of an association between the size of the immigrant-native test score gap and average PISA science scores at the country level (correlation = -0.11). Although Figure 6.6 refers specifically to science, similar findings emerge for reading and mathematics (see the online data tables for further details).

30. Further detail on how achievement differs in Wales by immigration status is provided in Table 6.5. This compares science, mathematics and reading mean scores for first generation immigrants, second generation immigrants and 15-year-olds who were born in the UK ('natives'). Across all three core PISA domains, UK natives are the highest-achieving group. However, partly due to the limited sample size, the average PISA science, mathematics and reading scores of this group are not significantly different from the average for first-generation or second-generation immigrants. Consequently, in Wales, the difference between immigrants and natives average PISA test scores is within the range one would expect given sampling error.

Table 6.5 Average PISA test scores by immigrant status in Wales

	Pupil and parents born in UK	First-generation immigrant	Second-generation immigrant	Missing
Science	488	471	482	**452***
Mathematics	480	471	484	**445***
Reading	480	464	481	**445***
Observations	**3,012**	**178**	**73**	**188**

Source: PISA 2015 database.

Notes: * indicates significantly different from the 'pupils and parents born in UK' ('natives') category at the five per cent level.

Key point

Differences in average PISA scores between pupils from immigrant backgrounds and pupils whose family were born in the UK are not statistically significant at the five per cent level.

6.5 How do PISA scores differ between pupils who completed the test in English versus Welsh?

31. Wales is a bilingual country, with this reflected in the teaching and learning of Welsh language within schools. Indeed, around 10 per cent of pupils sit GCSE Welsh as their first language, with many more completing GCSE Welsh as a second language option[46]. In 2014/15, there were around 5,500 GCSE Welsh First Language entries (representing around 16 per cent of the number of 15 year olds); around 10,900 GCSE Welsh Second Language (full course) entries (representing around 32 per cent of the number of 15 year olds); and around 10,400 GCSE Welsh Second Language (short course) entries (representing around 30 per cent of the number of 15 year olds)[47].

32. In PISA, all pupils had the option of completing *either* an English language or Welsh language version of the test. This differs from current practice in GCSE examinations, where for most subjects (bar English / Welsh language) pupils are presented test questions in both mediums. In this sub-section, we explore differences in reading, science and mathematics proficiency between pupils who completed the PISA test in English versus Welsh.

33. A total of 575 pupils in the Welsh PISA sample study Welsh as a first language in school. Out of these 575 pupils, 337 (59 per cent) chose to complete the PISA test in Welsh while 238 (41 per cent) took the test in English. In addition, one pupil who studies English as a first language and one pupil whose information could not be linked to the administrative data also completed the test in Welsh. Throughout this sub-section, we draw comparisons between the 339 pupils who completed the Welsh language version of the PISA test to the 3,112 pupils who completed the test in English.

34. Table 6.6 begins by providing some descriptive information about the characteristics of pupils who took the PISA test in English and Welsh. Although there is little difference in terms of gender, it does seem that young people who decided to take the test in Welsh came from slightly more advantaged socio-economic backgrounds. For instance, 48 per cent of those answering in Welsh reported that at least one of their parents holds a university degree (compared to 36 per cent of English-language pupils), while 24 per cent reported that there are more than 200 books in their home (compared to 16 per cent for the English-language group). In

[46] See Jones (2012: 78).
[47] https://statswales.gov.wales/Catalogue/Education-and-Skills/Schools-and-Teachers/Examinations-and-Assessments/Key-Stage-4

terms of other markers of Welsh-language use, the vast majority of pupils who took the PISA test in Welsh attended a Welsh medium school (94 per cent compared to nine per cent of those who completed the test in English). Similarly, only two per cent who took the English version spoke Welsh as the main language at home, compared to 38 per cent who completed the Welsh-language version of the PISA test.

Table 6.6 The characteristics of pupils who completed the English and Welsh versions of the PISA 2015 test

	Test in English	Test in Welsh
Gender		
Female	49%	52%
Male	51%	48%
Highest parental education		
No education	2%	1%
GCSEs	16%	15%
A/AS-Levels	16%	8%
Higher education below degree	21%	17%
University degree	36%	48%
No data	9%	12%
Books in the home		
0-10 books	20%	16%
11-25 books	18%	15%
26-100 books	27%	29%
101-200 books	15%	14%
201-500 books	10%	13%
More than 500 books	6%	11%
No data	4%	2%
Parental occupation		
Least advantaged	22%	17%
Second quartile	21%	20%
Third quartile	22%	26%
Most advantaged	19%	21%
No data	16%	16%
Welsh spoken at home		
Yes	2%	38%
Attend Welsh medium school		
Yes	9%	94%
First language studied		
English	91%	0%
Welsh	7%	99%
No data	2%	0%
Observations	**3,112**	**339**

Source: PISA 2015 matched database.

Notes: Figures may not sum to 100 per cent due to rounding.

35. Table 6.7 compares average PISA test scores across these two groups of pupils. Those who took the test in Welsh tended to achieve lower average scores. There is a difference of 22 points in science which is statistically significant at the five per cent level. There is also a difference of 25 points in reading (480 versus 455) although the limited sample size means that this difference does not reach statistical significance at the five per cent level (p =0.09). In contrast, average scores are very similar in mathematics (478 versus 475).

Table 6.7 Average PISA test scores of pupils who completed the PISA test in English and Welsh

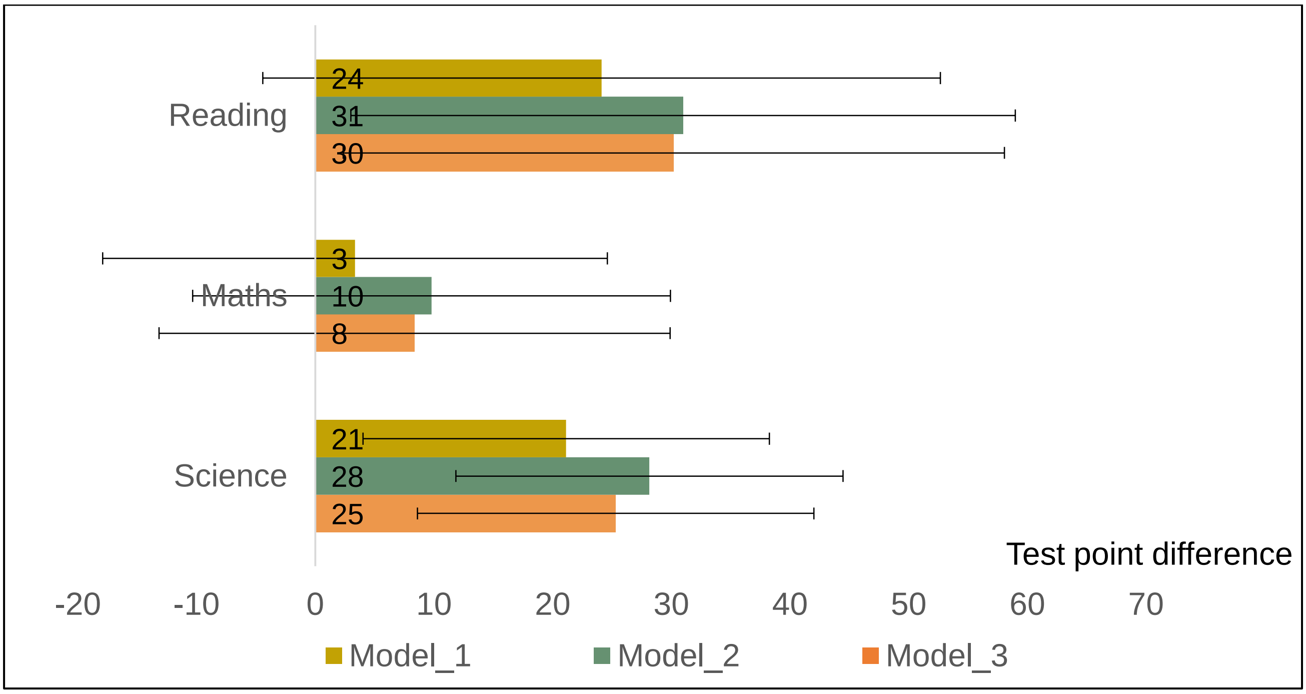
	Test in English	Test in Welsh
Science (overall)	487	465
Mathematics	478	475
Reading	480	455
Observations	**3,112**	**339**

Source: PISA 2015 matched database.

36. To what extent can this difference in achievement be explained by differences in the background characteristics of these pupils? Figure 6.7 presents results from a series of regression models where we have controlled for a number of demographic variables. Model 1 provides results from an ‘unconditional’ model (i.e. this simply illustrates differences in average scores between pupils who completed the test in English versus Welsh). Controls are then added for gender and a number of socio-economic indicators in model 2 (parental education, parental occupation and books in the home). Finally, language most often spoken by pupils at home is also included in model 3.

37. The key message to be taken from Figure 6.7 is that differences in these background characteristics cannot explain the achievement gap between pupils who completed the test in English versus Welsh. In fact, differences in average PISA science and reading scores between pupils who completed the test in English and Welsh actually increase somewhat (and reach statistical significance at the five per cent level) after such background characteristics have been taken into account. For instance, the difference between pupils who took the PISA test in English versus Welsh stands at 25 test points in science (significant at the five per cent level), 30 points in reading (significant at the five per cent level) and 8 points in mathematics (insignificant at the five per cent level) after these characteristics have been taken into account. This further strengthens the evidence that pupils who took the PISA test in Welsh achieve lower average scores than those who completed the test in English (at least in the science and reading domains).

Figure 6.7 Results from regression models investigating differences between pupils who completed the PISA test in English versus Welsh



Source: PISA 2015 matched database.

Notes: Model 1 does not include any controls. Model 2 includes controls for gender, parental education, parental occupation and books in the home. Language most often spoken at home is also added in model 3. Thin black line running through the centre of bars refers to the estimated 95 per cent confidence interval.

38. Table 6.8 takes this analysis a step further by considering the intersection between the main language pupils study in school and the language in which they chose to complete the PISA test. For brevity, we focus upon performance in science. The most notable difference in Table 6.8 is between the 337 pupils who study Welsh as a first language and who took the test in Welsh (mean score = 466) versus the 238 pupils who study Welsh as a first language but who chose to take the test in English (mean score = 495). There is a difference of 29 PISA science points (approximately a year of schooling) between these groups, with differences statistically significant at the five per cent level. A similar finding holds for reading (a difference of 30 test points between pupils who took the test in Welsh versus English) with this on the boundary of the five per cent significance threshold (p=0.06). There is no evidence of any difference in mathematics. Nevertheless, Table 6.8 suggests that pupils who study Welsh as a first language at school actually achieve higher average PISA scores in some subjects if they decide to take the test in English.

Table 6.8 Average PISA science scores of pupils by whether they study English or Welsh in school and the language in which they completed the PISA test

	Test in English	Test in Welsh
Study in English	485	N/A
Study in Welsh	495	466

Source: PISA 2015 matched database.

Notes: N/A appears in the ‘study in English’ and ‘test in Welsh’ cell due to the insufficient sample size within this group. The sample has been restricted to only those pupils for whom information on first language studied in school is available.

Key point

Pupils who took the Welsh language version of the PISA 2015 science test achieved lower scores than their peers who completed the test in English.

Chapter 7. Differences in achievement between schools

- In Wales, 90 per cent of the variation in pupils' PISA science scores occurs within schools, while 10 per cent occurs between schools. In other words, most of the variation in PISA scores occurs amongst pupils who attend the same school.

- The proportion of within-school variation in Wales is larger than the average across members of the OECD. In this respect, Wales is similar to several Nordic countries such as Sweden (83 per cent), Norway (92 per cent) and Finland (92 per cent).

- There is little difference in average PISA science, reading and mathematics test scores between pupils who attend English medium and Welsh medium schools.

- The average PISA score of pupils in green national support category schools is 497 in science, 489 in mathematics and 489 in reading. This compares to a mean score around 470 in each subject for pupils who attend an amber support category school.

- Approximately a quarter of pupils studying in schools in the amber support category lack basic skills in science. This compares to less than a fifth of pupils in green support category schools.

1. This chapter examines differences in young people's science, mathematics and reading competencies by school characteristics. It begins by decomposing the variation in PISA test scores into two components: the proportion that occurs *within* schools versus the proportion that occurs *between* schools. The distribution of PISA test scores is then reported by school type (Welsh medium versus English medium), and by National School Categorisation System category.

2. All estimates presented within this chapter need to be carefully interpreted, particularly because the number of schools in Wales participating in PISA 2015 is limited. For instance, school-level sample sizes are relatively small for certain groups, such as Welsh medium schools (24 schools with 608 pupils) and those rated as red within the National School Categorisation System (12 schools with 271 pupils). These results will therefore be subject to a relatively high degree of uncertainty due to sampling error.

3. Throughout this and the following chapter, 'Welsh medium schools' include schools within all the bilingual categories. It should also be noted that schools have been categorised according to their National School Categorisation System group at December 2015 (the approximate time of the PISA test). More generally, all figures reported in this chapter refer to descriptive associations only, and do not reveal cause and effect.

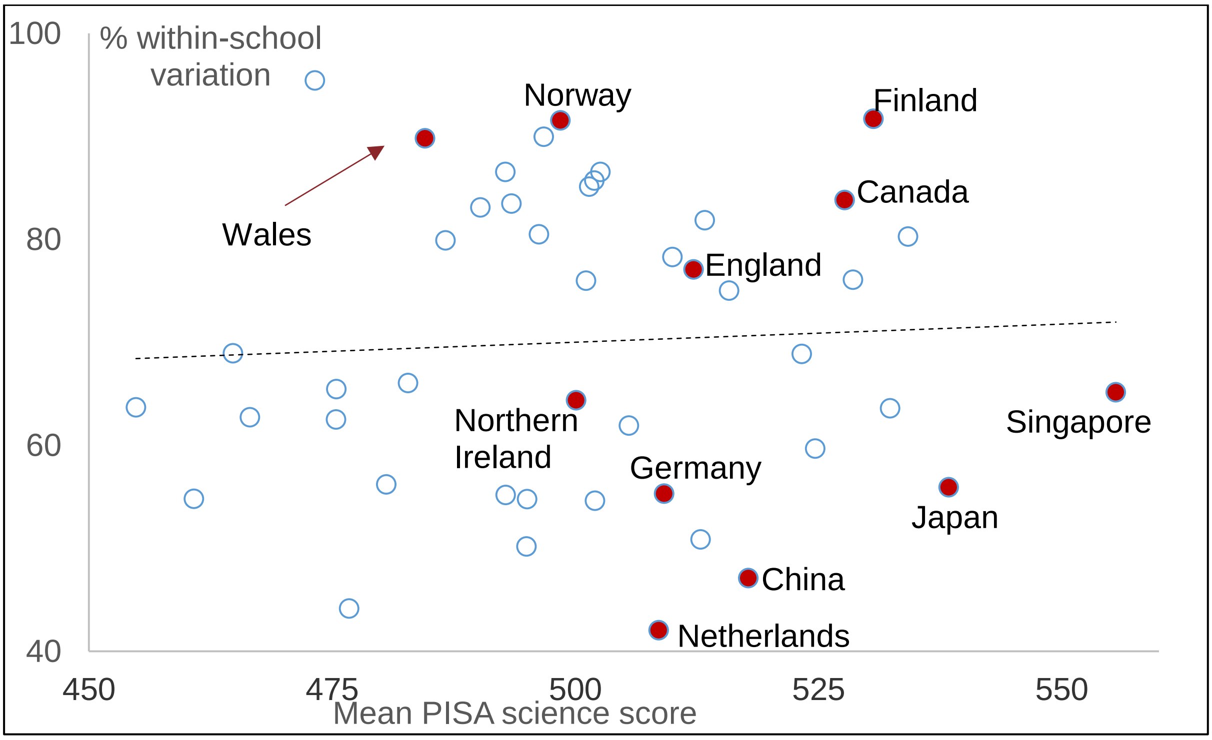
7.1 To what extent does variation in science achievement occur within schools versus between schools in Wales? How does this compare to other countries?

4. This sub-section splits the variation in 15-year-olds' PISA science scores into the portion that occurs within schools versus the portion that occurs between schools. Between school variation refers to the extent to which differences in achievement can be 'explained' (in a statistical sense) by the sorting of pupils into different schools. In contrast, within-school variation refers to the extent that PISA test scores differ, even amongst pupils who attend the same school. It is important to note that these figures do not reveal the 'importance' or 'impact' of schools per se (i.e. it is not necessarily the case that where the between school variation is higher, schools are more important). Rather, the proportion of the variance explained between schools is partially determined by 'selection effects', reflecting the fact that young people with certain characteristics disproportionately attend particular types of school. Nevertheless, previous research has suggested that a reduction '*in within-school variation is linked with an improvement in value-added, so schools embarking on the journey of reducing within-school variation can be certain that it will be*

productive on results[48]'. It is therefore important to understand the extent of within-school achievement variation that occurs in Wales, and how this compares to other countries.

5. Figure 7.1 plots average PISA science scores (horizontal axis) against the proportion of the variation in pupils' science achievement that occurs within schools (vertical axis). Note that the sample of countries in this analysis has been restricted to those with a mean science score above 450 test points. In Wales, most of the variability in 15-year-olds' science achievement occurs within schools (90 per cent), with only 10 per cent of the variance explained between schools. This suggests that there are substantial differences in 15-year-olds' science achievement, even when they attend the same school. Indeed, within-school variation is the larger of the two components in most countries. Thus, despite significant differences in the structure of secondary schooling systems across countries, within-school variation in pupils' achievement always has an important role.

Figure 7.1 The proportion of the variation in pupils' PISA science scores that occurs within schools versus mean science scores



Source: PISA 2015 database.

Notes: The sample of countries included has been restricted to those with a mean score above 450 test points.

[48] Reynolds (2007).

6. Wales sits towards the top of Figure 7.1; the proportion of achievement variation occurring within schools in Wales is larger than in most other participating countries included in this comparison. It is notable that countries with a strong tradition of academic selection into secondary schools sit towards the bottom of Figure 7.1, with a comparatively low proportion of the variance in pupils' science scores occurring within schools (and, therefore, have a high percentage of achievement variation occurring between pupils in different schools). Prominent examples include Germany and the Netherlands. In contrast, countries with a mainly comprehensive schooling system, where the use of academic selection into secondary schools is rare, are generally towards the top of Figure 7.1. Examples include Finland, Norway and Sweden, where up to 90 per cent of the variation in PISA science scores occurs within schools.

7. Further inspection suggests that, compared to other countries with a comprehensive schooling system, the proportion of within-school variation in Wales is high. For instance, 77 per cent of the variation in PISA science scores occurs within schools in England, 81 per cent in the United States and 78 per cent in Australia. In all these countries, the proportion of the variance explained within schools is less than in Wales (90 per cent). This indicates that most of the inequality in 15-year-olds' science achievement in Wales occurs amongst pupils who attend the same secondary school (and not between pupils who attend different schools).

8. Figure 7.1 also shows that there is essentially no association between the proportion of achievement variation that occurs within schools and average PISA science scores at the country level (correlation = 0.05). For instance, whereas the proportion of within-school variation is comparatively low in some of the top-performing PISA countries (e.g. Singapore, Japan) it is relatively high in others (e.g. Finland, Canada). There is hence little evidence to suggest that a low (or a high) proportion of within-school variation is a common trait amongst the leading PISA countries.

Key point

PISA scores vary more amongst pupils within the same school in Wales than they do between schools. Wales is similar to several other countries in this respect.

7.2 How do PISA test scores differ between English and Welsh medium schools?

9. Approximately a fifth of all Welsh secondary school pupils are taught in Welsh medium schools. These are schools where Welsh is a major language of instruction, sometimes in conjunction with the use of English. Throughout this report, the term 'Welsh medium school' encapsulates the following categories:

- Welsh medium = Welsh is the day to day language of the school
- Bilingual category A = At least 80 per cent of subjects are taught only in Welsh to all pupils
- Bilingual category B = At least 80 per cent of subjects are taught through the medium of Welsh but also through the medium of English
- Bilingual category C = 50-79 per cent of subjects are taught through the medium of Welsh but also through the medium of English
- Bilingual category Ch = All subjects are taught in both English and Welsh

10. Table 7.1 illustrates the proportion of Welsh medium schools that participated in PISA 2015 that belong to each of these five categories. A total of 10 schools use Welsh as the day-to-day language of the school. A further four participating schools (with 107 participating pupils) are within bilingual category A, and eight schools (with 202 pupils) within bilingual category B.

Table 7.1 Pupils in Welsh medium schools by sub-category

	Schools	Pupils
Welsh medium	10	250
Bilingual (Type A)	4	107
Bilingual (Type B)	8	202
Bilingual (Type C)	1	23
Bilingual (Type Ch)	1	26
Total	**24**	**608**

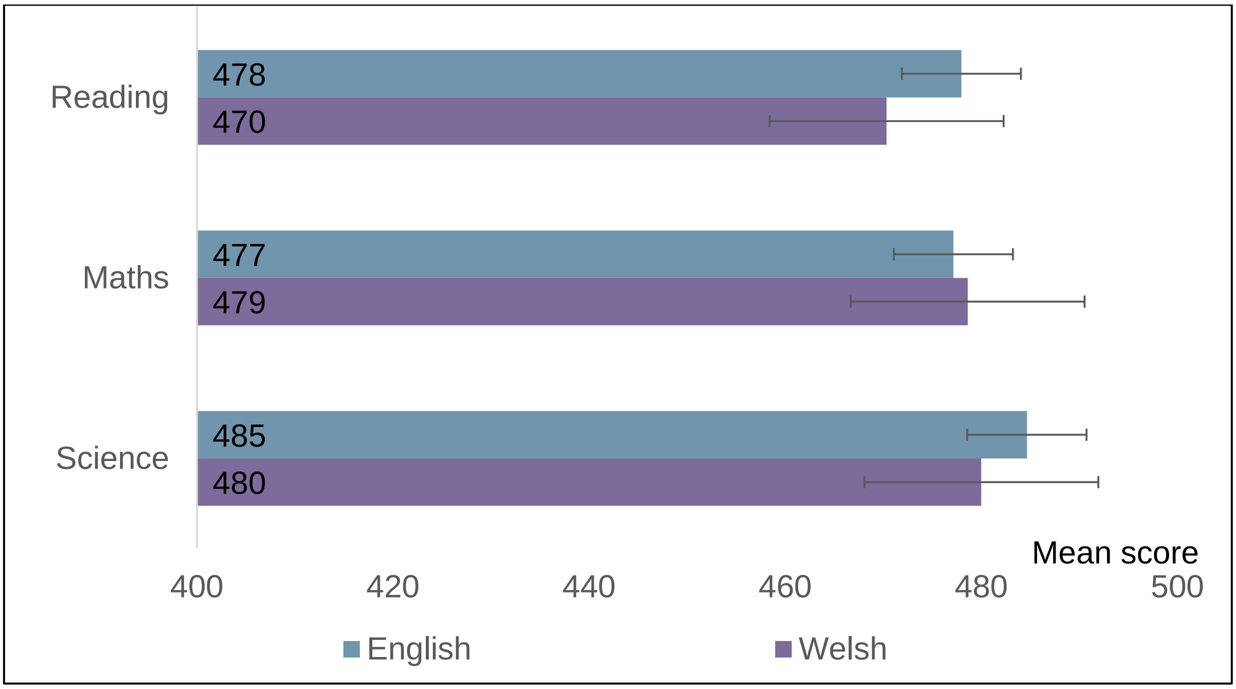
Source: PISA 2015 matched database.

11. Section 6.5 highlighted how pupils who completed the PISA test in Welsh obtained lower scores in reading and science than those who chose to complete the test in English. However, previous evidence has indicated that achievement at Key Stage 4 may be higher among young people who attend Welsh medium schools than young people who attend English medium schools[49]. Yet there remains

[49] See Estyn (2011).

relatively little evidence as to how PISA test scores differ between pupils who attend these different school types.

Figure 7.2 Mean PISA scores for pupils in English and Welsh medium schools



Source: PISA 2015 matched database.

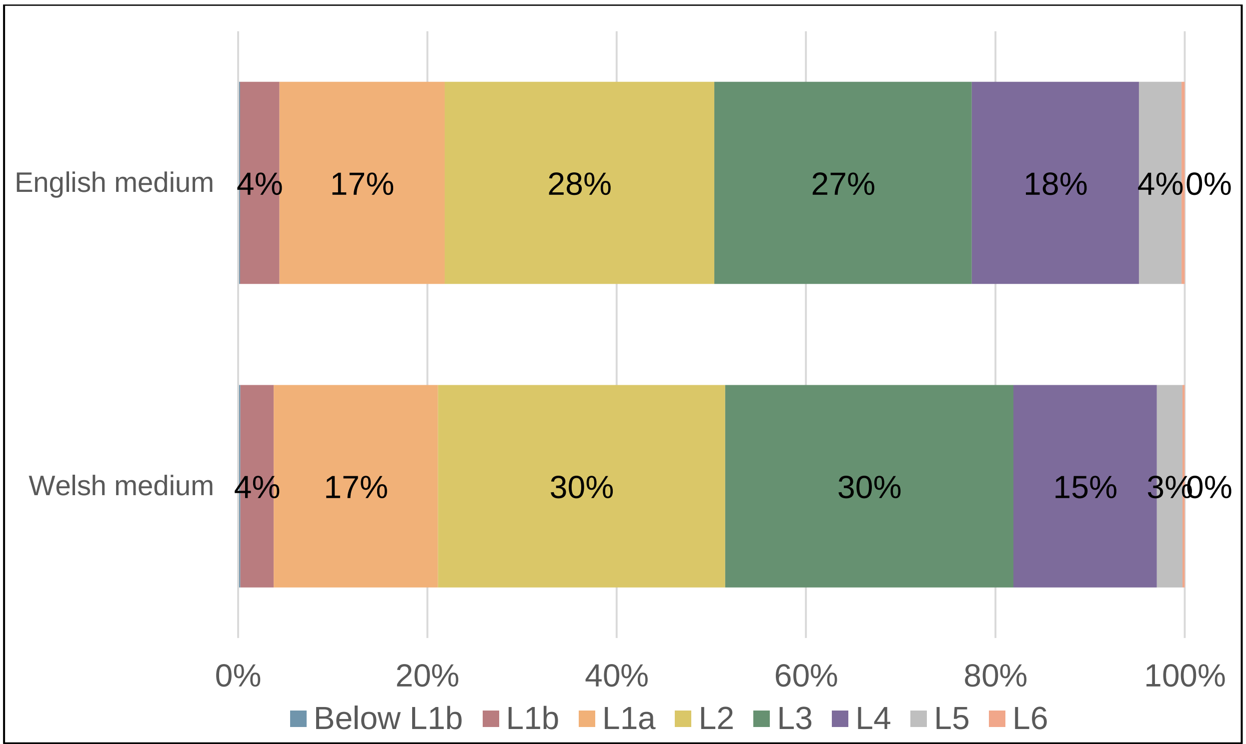
Note: 95 per cent confidence interval represented by thin black lines running through the centre of each bar.

12. Figure 7.2 documents differences in average PISA scores for pupils who attend English and Welsh medium schools. Overall, there is little evidence that average PISA test scores differ amongst pupils who attend these types of school. The average achieved by pupils studying in Welsh medium schools is 480 in science, 479 in mathematics and 470 in reading. Each of these figures is within 10 PISA test points of pupils' studying within English medium schools (485 in science, 477 in mathematics and 478 in reading). Moreover, differences between English and Welsh medium schools do not reach statistical significance at the five per cent level in any subject area.

13. Although pupils in English and Welsh medium schools achieve similar average PISA scores, the distribution of performance may differ, including in terms of the proportion of high and low-achievers. Evidence on this matter is provided in

Figure 7.3, which compares English and Welsh medium schools in terms of the proportion of pupils who achieve at each of the PISA science proficiency levels[50].

Figure 7.3 The distribution of PISA scores by the primary language of instruction (English / Welsh) used within the school



Source: PISA 2015 matched database.

14. The distribution of pupils across the PISA science proficiency levels is similar for English and Welsh medium schools. Just over a fifth of pupils lack basic science skills (i.e. fail to reach PISA Level 2) across both school types. Likewise, there is a similar share of high-performing pupils across Welsh medium (three per cent) and English medium (five percent) schools. Equivalent results hold within the mathematics and reading domains. Together, Figure 7.2 and Figure 7.3 indicate that PISA scores do not vary substantially between pupils depending upon their school's primary language of instruction.

Key point

Pupils who attend English and Welsh medium schools achieve similar scores, on average, on the PISA science, mathematics and reading test.

[50] Analogous results for reading and mathematics can be found in the online data tables.

7.3 How do PISA test scores in Wales vary by school support category?

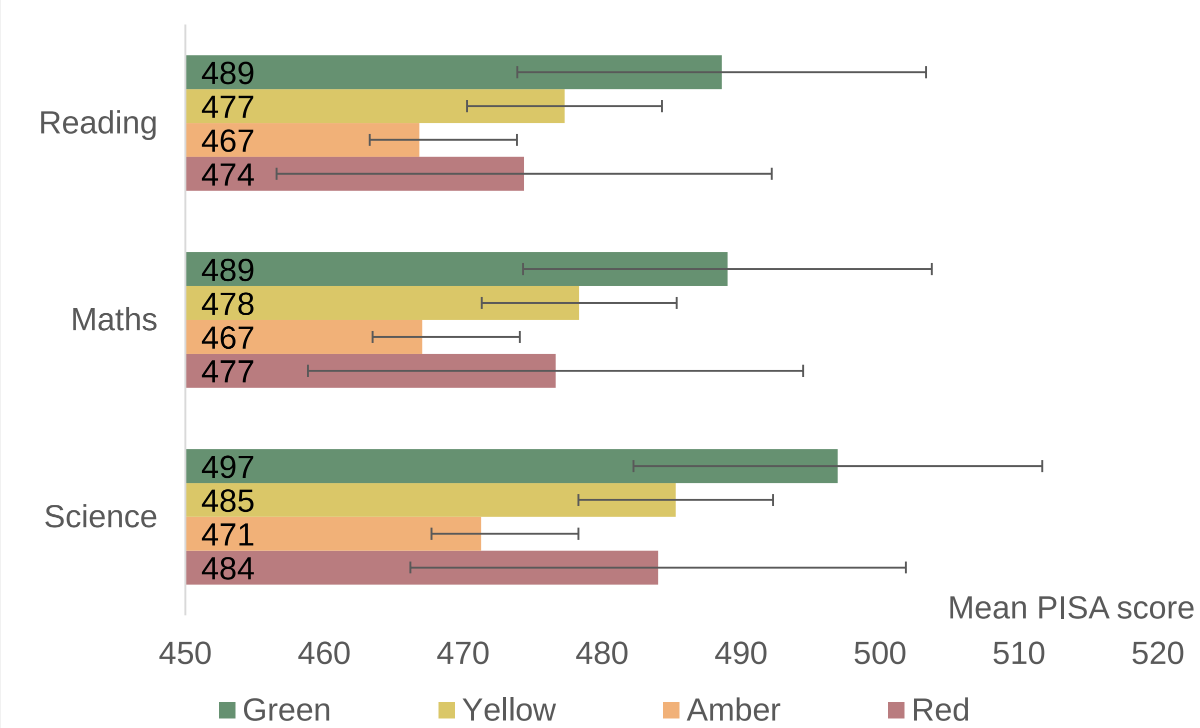
15. Secondary schools in Wales are categorised into four groups (green, yellow, amber and red) in order to identify those schools that are most in need of support, with the aim of raising standards and performance as a result. The system is designed to put schools in a position that enables them to identify the factors that contribute to their progress and achievement, or what areas to focus on to achieve further development.

16. Schools are placed into one of the four groups after a three-step procedure. The first step is data driven, and based upon the extent to which the school has met an agreed set of performance standards provided by the Welsh government. In the second step, schools self-evaluate their capacity to improve in terms of leadership, teaching and learning. Finally, in stage three, the judgements made in stages one and two leads to the assignment of the colour-coded support category, triggering a tailored programme of support, challenge and intervention[51].

17. There has previously been little research into how young people's skills in reading, mathematics and science (as measured by PISA) vary according to the support category of the school. Figure 7.4 therefore illustrates how average PISA reading, mathematics and science test scores differ between schools in the various support categories. Pupils within green-coded schools score, on average, 497 in PISA science, 489 in mathematics and 489 in reading. This is significantly higher at the five per cent level than pupils studying in schools within the amber category, where the average PISA science, mathematics and reading score is 471, 467 and 467 respectively. Schools with a yellow rating fall between these two extremes, with a mean score of 485 in science, 478 in mathematics and 477 in reading. Differences in average PISA science and mathematics scores between the 'yellow' and 'amber' categories are statistically significant at the five per cent level. On the other hand, due to the small school level sample sizes, differences in mean scores between the green and yellow categories do not reach statistical significance in either reading, mathematics or science.

[51] Welsh Government Website (2016).

Figure 7.4 Mean PISA scores by school support category in Wales



Source: PISA 2015 matched database.

Note: 95 per cent confidence interval represented by thin black lines running through the centre of each bar. Results reported for schools and pupils where data available.

18. Schools within the red category are somewhat of an outlier in Figure 7.4, with the mean score perhaps higher than one would anticipate. However, it is important to remember that the sample size is particularly small for this group (12 schools and 271 pupils), and the 95 per cent confidence interval very wide. Indeed, most comparisons between the red category and other groups are not statistically significant at the five per cent level. This indicates that sampling error is too great to draw robust inferences regarding the science, mathematics and reading proficiency of pupils who attend such schools. Readers should therefore not place too strong an interpretation upon estimates of mean PISA scores for this particular group.

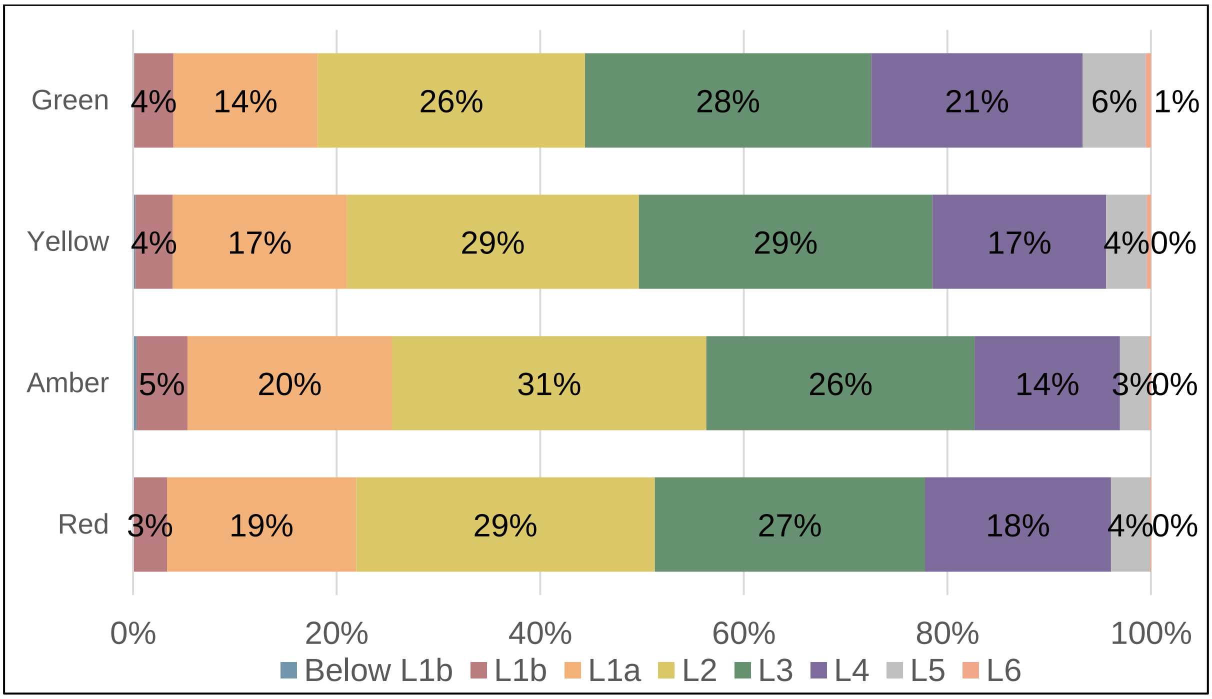
19. In additional analysis, we have investigated how these results change after accounting for differences in the socio-economic and demographic compositions of these schools[52]. Our central finding is that difference in PISA scores between the various national school support categories is reduced after accounting for the demographic characteristics of their pupils. For instance, the difference in mean

[52] These results are based upon an Ordinary Least Squares regression model, with PISA science scores as the dependent variable. Controls have been included for gender, parental education, parental occupation, immigrant status and the number of books at home.

scores between pupils in the green and amber categories is reduced from 26 to 16 points in science, from 22 to 13 points in mathematics and 22 to 13 points in reading. Nevertheless, differences in mean scores between the green and amber groups remain statistically significant at the five per cent level for science and mathematics[53]. Hence differences in demographic and socio-economic characteristics can explain some, although not all, of the achievement differential between pupils who attend green and amber category schools.

20. Figure 7.5 provides further detail on how PISA science scores vary by school support category. Specifically, it compares the distribution of pupils across the PISA proficiency levels[54]. Approximately one-in-five (18 per cent) pupils in green schools achieves a PISA science score below Level 2. This compares to approximately 21 per cent of young people within yellow schools and 25 per cent in the amber category. There is hence a moderate degree of variation across school support categories in terms of the proportion of pupils who lack basic science skills.

Figure 7.5 The distribution of PISA science proficiency levels by school support category



Source: PISA 2015 matched database.

[53] Results for reading fall just below the boundary of statistical significance at the 5 per cent level (p=0.07) once gender, parental education, parental occupation, immigrant status and the number of books at home have been controlled.
[54] Analogous results for reading and mathematics are provided in the online data tables.

21. At the other end of the spectrum, seven per cent of 15-year-olds in green coded secondary schools are classified as 'high-achievers' (reaching PISA Level 5 or 6). In comparison, four per cent reach at least Level 5 in yellow coded schools and three per cent in the amber group. Hence there is again evidence of some variation in proficiency according to the support category of the school, though this is relatively modest. Similar results hold for reading and mathematics (see the online data tables).

Key point

Pupils studying in 'green' support category schools achieve, on average, PISA reading, mathematics and science scores around the OECD mean. Pupils in 'amber' schools achieve average PISA science scores below 470 test points – comparable to the overall mean in countries like Croatia and Israel.

Chapter 8. School management and resources

- Headteachers in Wales report taking a more proactive and collaborative approach to school leadership and management than headteachers in most industrialised countries. However, within Wales, there are relatively few differences in leadership style between headteachers working in schools in different national support categories.

- A lack of good quality school infrastructure stands out as a particular concern amongst headteachers in Wales. This is especially the case for headteachers who are leading schools in the amber and red support categories.

- Headteachers in Wales are generally positive about the resources available to support science learning within their school. However, headteachers who are leading schools in amber and red support categories are less likely to report having a science department that is well-equipped.

- Headteachers in Wales are more likely to report staff absenteeism as a barrier to pupils learning than headteachers in the average OECD or high-performing country.

- Within Wales, staff absenteeism, teachers not meeting individual pupils' needs and teachers not being prepared for class are key concerns amongst headteachers leading schools in the amber and red support categories.

- Headteachers report that extensive quality assurance processes are already in place within the Welsh education system.

1. A number of factors have an impact upon the functioning of a school, and whether it provides the optimal environment to maximise pupils' well-being and attainment. This includes access to sufficient educational resources, the conduct of staff and the management approach of senior leadership teams. The aim of this chapter is to provide new evidence on such matters for Wales by drawing upon the PISA headteacher questionnaire.

2. As part of the PISA study, headteachers from all participating schools were asked to complete a questionnaire. This included questions covering a range of topics, including management styles, resources, school climate and quality assurances processes. A total of 118 headteachers completed this questionnaire in Wales, reflecting an unweighted response rate of 84 per cent amongst the participating schools.

3. Based upon headteachers' responses, this chapter seeks to answer the following questions:

- *How do headteachers in Wales manage their staff and their schools?*
- *Do headteachers in Wales believe they have access to sufficient resources in order to support pupils' learning?*
- *Are schools in Wales well-equipped to support pupils' learning in science?*
- *How do headteachers in Wales view the conduct of their staff?*
- *What quality assurance processes are used in schools in Wales?*

4. Each sub-section within this chapter will follow a similar structure. Responses of Wales' headteachers are first compared to the responses of headteachers in other countries. This focuses upon comparisons to the average across OECD members and the average across the 10 countries with the highest mean PISA science scores ('H10'). We then turn to variation within Wales, focusing upon differences between schools according to their National School Categorisation System category.

5. As with the preceding chapter, results need to be carefully interpreted. First, sample sizes remain small for particular sub-groups (e.g. schools in the 'red' national support category, Welsh medium schools). Estimates for these groups are therefore subject to a high degree of sampling error. Second, it should be remembered that the analysis presented in this chapter is based upon information reported by headteachers. Any data collected in this manner may be subject to recall bias and

measurement error. The subjective nature of some questions should also be considered when interpreting the results.

8.1 How do headteachers in Wales manage their staff and schools?

6. Effective leadership is an essential ingredient for school effectiveness, with research suggesting pupils make more academic progress in schools with better leadership[55]. There has consequently been much academic and policy interest in the development of effective leaders for schools. In this sub-section we provide new insight into school leadership styles in Wales using data from PISA 2015.

7. Headteachers across all participating countries were asked the following question as part of the school questionnaire:

'*Below are statements about your management of this school. Please indicate the frequency of the following activities and behaviours in your schooling during the last academic year'*

Table 8.1 provides the 13 statements headteachers were asked to respond to, along with the percentage who reported undertaking each activity at least once a month during the last academic year[56]. Based upon the evidence provided in Table 8.1, there are two points of particular note.

8. First, for almost every question, the percentage of headteachers who report the activity occurring at least once a month is greater in Wales than the average across OECD members and the average across the high-performing (H10) countries. This includes factors related to setting and achieving the goals of their school (e.g. ensuring professional development activities of staff are consistent with the aims of the school) and in encouraging a collaborative approach to school improvement (e.g. asking teachers to review school management practises, solving classroom problems together).

[55] Day et al. (2009).

[56] Headteachers were asked to respond to each question using a six point scale, ranging from '*did not occur*' through to occurring '*more than once a week*'. Table 8.1 presents the percent of teachers who ticked one of the top three categories ('*once a month'*, *'once a week' or 'more than once a week'*).

Table 8.1 Headteachers' management of teachers and schools

	Wales	OECD	H10
I use pupil performance results to develop the school's educational goals	61%	**23%***	**18%***
I make sure that the professional development activities of teachers are in accordance with the teaching goals of the school	50%	**33%***	**33%***
I ensure that teachers work according to the school's educational goals	75%	**53%***	**48%***
I promote teaching practices based on recent educational research	57%	**41%***	**34%***
I praise teachers whose pupils are actively participating in learning	80%	**63%***	**55%***
When a teacher has problems in his/her classroom, I take the initiative to discuss matters	77%	**68%***	**64%***
I draw teachers' attention to the importance of pupils' development of critical and social capacities	66%	**56%***	**51%***
I pay attention to disruptive behaviour in classrooms	95%	**82%***	**79%***
I provide staff with opportunities to participate in school decision-making	65%	72%	65%
I engage teachers to help build a school culture of continuous improvement	79%	73%	**66%***
I ask teachers to participate in reviewing school management practices	49%	**34%***	**36%***
When a teacher brings up a classroom problem, we solve the problem together	80%	78%	76%
I discuss the school's academic goals with teachers at faculty meetings	77%	**51%***	**49%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools where the headteacher reports undertaking the activity at least once a month over the past academic year. Bold font with a * indicates statistically significant difference from Wales at the five per cent level.

9. Second, there are certain questions where the difference between Wales and the OECD / H10 average is particularly pronounced. For instance, headteachers in Wales are much more likely to regularly use pupils' performance data to develop their school's educational goals (61 per cent in Wales versus an OECD average of 23 per cent and an H10 average of 18 per cent). Indeed, a greater proportion of headteachers in Wales use pupil performance data in setting their school's objectives than in any of the 10 highest performing countries. Other differences include headteachers in Wales being more likely to regularly praise staff when they see pupils actively engaged in learning (80 per cent versus an OECD average of 63 per cent), and being more likely to encourage staff to use an evidence-based

approach to develop their teaching practices (57 per cent in Wales versus a 41 per cent average across OECD members). It is also interesting to note that school leaders in Wales are more likely to encourage teachers to develop pupils' social skills than in the average high-performing country (66 per cent in Wales versus 51 per cent H10 average).

10. The H10 and OECD average figures reported in Table 8.1 mask the substantial variation that occurs across these countries. For instance, whereas 72 per cent of Canadian headteachers encourage the development of pupils' social skills, only 55 per cent do so in Finland and 12 per cent in Japan. Similarly, the proportion of headteachers regularly promoting the use of evidence-based teaching practices is notably higher in Canada (64 per cent) and Singapore (44 per cent) than in Estonia (25 per cent), Japan (12 per cent) and Hong Kong (13 per cent). This illustrates how school leadership and management approaches vary greatly across countries, even when we focus upon only those with the highest average PISA scores.

11. Variation in headteachers' approaches to leadership and management may also differ across different school types within Wales. Interestingly, there are relatively few differences reported by headteachers who lead schools in different support categories (red, yellow, amber, green). The questions demonstrating the greatest variation were the alignment of teachers' professional development activity with the goals of the school and the promotion of teaching practices based upon educational research. For instance, approximately 60 per cent of headteachers leading schools in the green and yellow category indicate that they ensure professional development activities are in accordance with teaching goals, compared to less than 40 per cent of headteachers in the amber and red groups[57].

12. Similarly, there are relatively few differences in leadership style between headteachers leading English and Welsh medium schools. The main exception is in the use of pupils' test scores to develop the school's educational goals. In English medium schools, 71 per cent of headteachers report that they do this at least on a monthly basis. This compares to only 28 per cent of headteachers in Welsh medium schools. Despite the small sample size, this difference of more than 40 percentage points is statistically significant at the five per cent level.

[57] Further details, providing a breakdown of the results by national support category, can be found in the online data tables (see Table 8.1b).

Key point

Headteachers in Wales report taking a more proactive and collaborative approach to school leadership and management than headteachers in most industrialised countries. However, within Wales, there are relatively few differences in leadership style between headteachers working in schools of different support categories or language mediums.

8.2 Do headteachers in Wales believe they have access to sufficient resources to support pupils' learning?

13. In order to operate effectively, schools require access to sufficient resources. This includes being able to recruit sufficiently skilled teachers and support staff, and being able to provide pupils with the educational materials that they need to succeed (e.g. textbooks, computers, equipment). Previous research has also suggested that the physical environment of a school may have an impact upon pupils' educational attainment[58]. For these reasons, it is important to consider whether headteachers in Wales feel that their schools are appropriately resourced, and how Wales compares to other countries in this respect.

Table 8.2 Headteachers' reports of resources lacking within their school

	Wales	OECD	H10
A lack of teaching staff	20%	**29%***	**31%***
Inadequate or poorly qualified teachers	15%	20%	**26%***
A lack of assisting staff	19%	**36%***	**33%***
Inadequate or poorly qualified assisting staff	13%	**19%***	**20%***
A lack of educational material	31%	34%	32%
Inadequate or poor quality educational material	28%	30%	30%
A lack of physical infrastructure	44%	36%	37%
Inadequate or poor quality physical infrastructure	48%	**35%***	**35%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools where the headteacher ticks either the 'to some extent' or 'a lot' categories. Bold font with * indicates significant difference from Wales at the five per cent level.

14. Table 8.2 details the extent to which headteachers report lacking, or only having access to poor quality, educational resources. Specifically, it provides the percentage of teachers who report that the factor in question hinders the school's capacity to provide instruction either 'to some extent' or 'a lot'. Figures for Wales are

[58] Barrett et al. (2015). Neilson and Zimmerman (2011).

compared to the average across OECD members, and the average across the 10 highest performing PISA countries in science (H10).

15. Overall, the figures for Wales are broadly in-line with the average across members of the OECD and the average across the H10 countries. In other words, in terms of resources, there are few issues which stand out as a particular concern amongst headteachers in Wales relative to headteachers in other countries. The only notable exception is with regards the physical infrastructure of schools. In the view of headteachers, almost half of Welsh pupils are taught in schools where the headteacher believes that a lack of good quality infrastructure is hindering their learning (48 per cent). This is more than 10 percentage points higher than the OECD and H10 averages (35 per cent). Headteachers in Wales are more likely to answer 'to some extent' or 'a lot' to this statement than any of the other statements presented in Table 8.2.

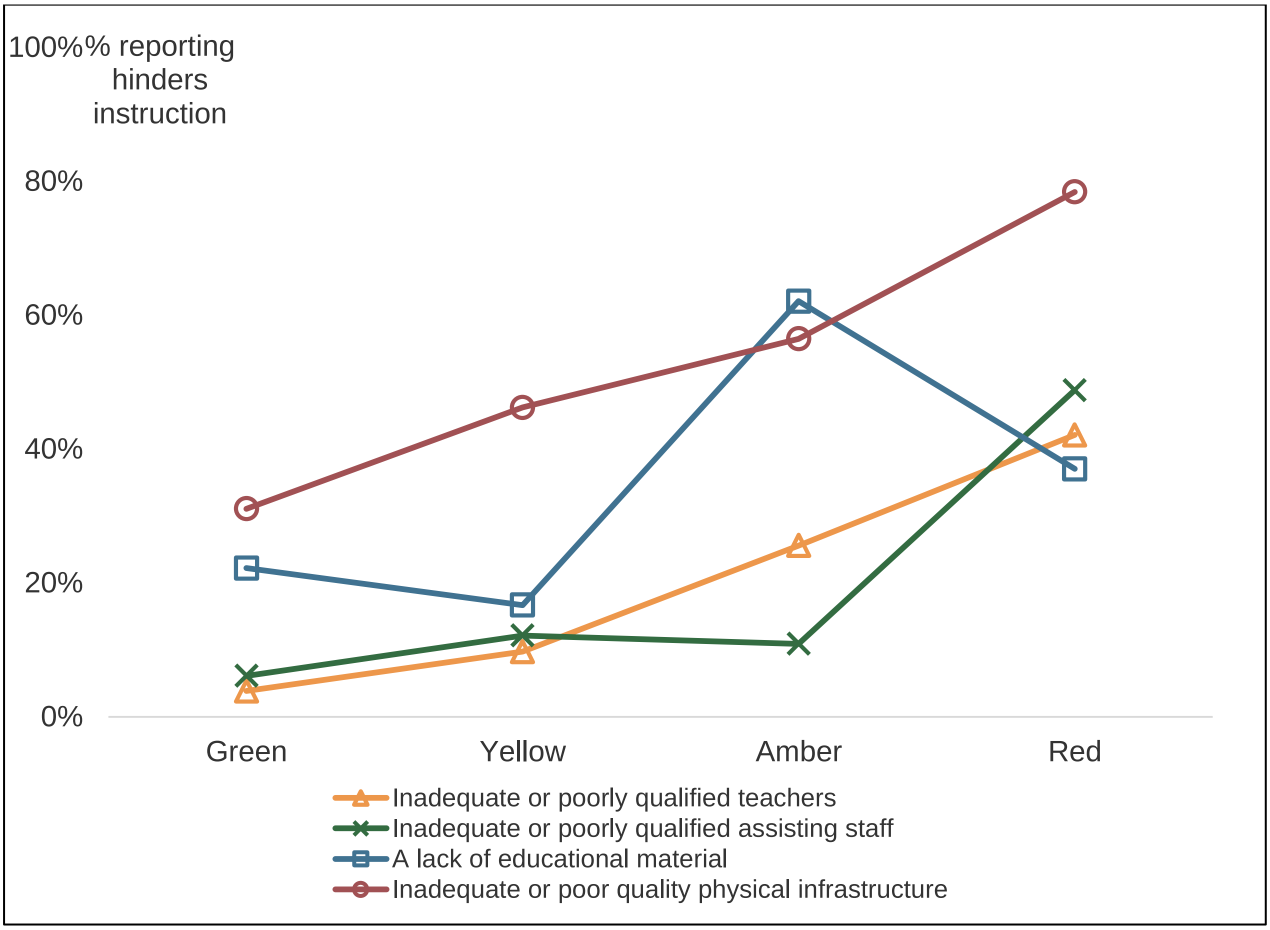
16. It is also interesting to note that, in a couple of areas, a lower proportion of headteachers in Wales report an issue than in the average OECD/H10 country. Headteachers in Wales appear generally more satisfied with their ability to hire suitably qualified staff than headteachers in the average industrialised country. For instance, 20 per cent of pupils in Wales are taught in schools where the headteacher believes a lack of teaching staff is hindering instruction, compared to an average of 29 per cent across members of the OECD. A similar finding holds true with regards to assisting / support staff (19 per cent in Wales versus an OECD average of 36 per cent).

17. Access to educational resources may also vary within countries, including between different school types within Wales. Figure 8.1 therefore explores how headteachers' views on educational resources vary by school support category. The most striking difference is in response to the statement regarding '*a lack of educational material*' (blue line with square markers in Figure 8.1). Around one-in-five headteachers agree with this statement in the green/yellow support category, compared to 62 per cent who lead schools rated as amber and 37 per cent rated as red. There is a difference between the green/yellow and the amber groups of more than 30 percentage points, with differences statistically significant at the five per cent level. This suggests that those individuals who lead lower categorised secondary schools see a lack of educational resources as a key barrier to instruction.

18. It is notable that headteachers who lead schools in lower support categories are more likely to report poor infrastructure as a barrier to pupils' learning than headteachers leading schools requiring less support (red line with circular markers). For instance, 31 per cent agree that this was a problem in schools within the green

category, compared to 56 per cent and 78 per cent in the amber and red categories. This linear trend is statistically significant at the five per cent level. Poor quality physical infrastructure therefore seems to be another key concern of headteachers who lead schools requiring more support in Wales.

Figure 8.1 Headteachers' reports of lacking resources by support category



Source: Matched PISA 2015 database

19. There is also some variation in headteachers' reports of the quality of their teaching staff by support category (orange line with triangular markers in Figure 8.1). Whereas less than four percent of headteachers leading green category schools identify inadequate or poorly qualified teachers as a barrier to pupils' learning, this increases to 10 per cent in yellow category schools, 26 per cent in the amber category and 42 per cent in the red category. This trend is also statistically significant at the five per cent level. Interestingly, the same does not seem to hold true for assisting staff (green line with cross markers), where the striking difference is between schools in the red category and all other groups. Nevertheless, it does seem that headteachers' views of the adequacy of their teaching staff vary somewhat depending upon the support category of their school.

20. There are few statistically significant differences between England and Welsh medium schools. One exception is with regards to a lack of physical infrastructure, with headteachers who lead Welsh medium schools less likely to report a lack of infrastructure as a factor hindering instruction. Specifically, 19 per cent of Welsh medium headteachers agree that this is a problem, versus 50 per cent of headteachers leading English medium schools. This is a difference of more than 30 percentage points, which is statistically significant at the five per cent level.

Key point

A lack of good quality school infrastructure stands out as a particular concern of headteachers in Wales. This is especially the case amongst headteachers who are leading schools requiring more support and of English medium schools.

8.3 Are schools in Wales well-equipped to support pupils' learning in science?

21. Whereas the previous sub-section focused upon headteachers' views of school resources in general, this sub-section pays specific attention to the availability of resources for use in the instruction of science. For instance, do headteachers in Wales believe that they have adequate laboratory equipment and appropriately trained staff to support pupils' learning in this subject? Or is it the case that when schools receive additional funds, headteachers tend to prioritise other areas? Table 8.3 provides some insight into such matters. It details how headteachers respond to a series of eight questions, each referring to a different aspect of the science resources available within their school.

22. Headteachers in Wales are generally positive about the science resources that are available within their school; more so than headteachers in the typical OECD or H10 country. This is particularly true for the availability of laboratory staff to support science teaching (90 per cent in Wales versus OECD / H10 averages of 34 per cent and 51 per cent respectively) and the availability of laboratory material (84 per cent in Wales versus 66 per cent and 72 per cent for the OECD and H10 averages).

Table 8.3 Headteachers' views on the science resources available within their school

	Wales	OECD	H10
Compared to other departments, our schools science department is well equipped	83%	**74%***	**75%***
If we ever have some extra funding, a big share goes into improvement of our school science teaching	42%	39%	47%
School science teachers are among our best educated staff members	68%	65%	62%
Compared to similar schools, we have a well-equipped laboratory	68%	62%	62%
The material for hands-on activities in school science is in good shape	79%	78%	73%
We have enough laboratory material that all courses can regularly use it	84%	**66%***	**72%***
We have extra laboratory staff that helps support school science teaching	90%	**34%***	**51%***
Our school spends extra money on up-to-date school science equipment	40%	**48%***	**49%***

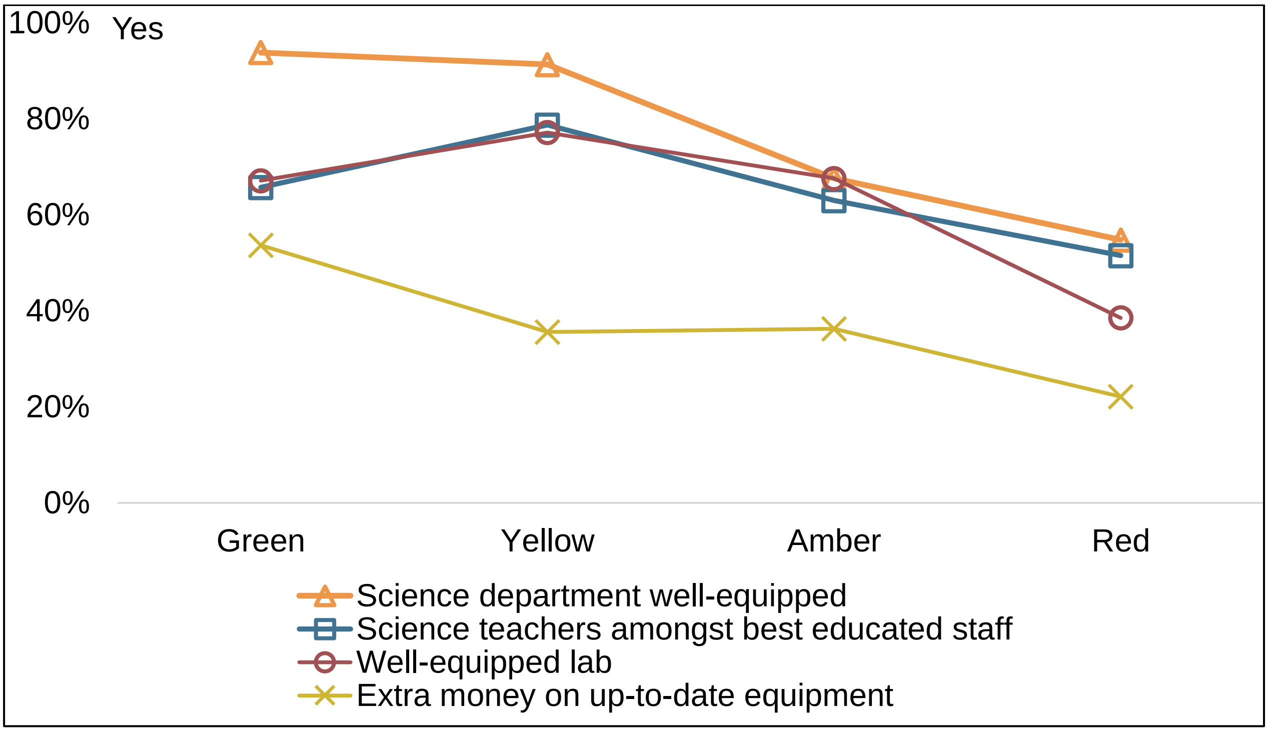
Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools where the headteacher ticks 'yes'. Bold font with a * indicates statistically significant difference from Wales at the five per cent level.

23. The two questions receiving the least positive responses from headteachers in Wales are with regards to the use of additional funding. Less than half of headteachers report that a big share of any extra funding received goes towards improving science teaching (42 per cent in Wales versus 39 per cent OECD average) and that their school spends extra money on up-to-date science equipment (40 per cent in Wales versus a 48 per cent average across OECD members). This may indicate that headteachers have other areas which take priority when additional funding is made available.

24. Figure 8.2 turns to variation within Wales, focusing upon differences in headteachers' responses by school categorisation. Two particular issues stand out. First, headteachers who lead schools requiring more support are less likely to report that their science department is well-equipped (orange line with triangular markers). Specifically, whereas more than 90 per cent of headteachers who lead schools in the green or yellow category respond positively to this statement, this falls to 68 per cent and 55 per cent for the amber and red categories.

Figure 8.2 Headteachers' reports of science resources by support category



Source: Matched PISA 2015 database

25. Second, there is also evidence of differences in headteachers' responses to the statement '*our school spends extra money on up-to-date school science equipment.*' Half of those (54 per cent) leading schools in the green category respond positively to this statement, compared to 36 per cent of those leading schools in the yellow and amber categories, and less than a quarter (22 per cent) within the red group. Together, this suggests that areas other than science may take priority for funding in schools that require more support.

Key point

Headteachers in Wales are generally positive about the resources available to support science learning within their school. However, headteachers who are leading schools requiring more support are less likely to report having a science department that is well-equipped.

8.4 How do headteachers view the conduct of their staff?

26. A successful school is likely to have teachers who are well prepared for the classes that they teach, and who are able to meet the needs of each individual pupil. On the other hand, frequent absenteeism and unprofessional behaviour of staff are

associated with lower levels of pupil attainment[59]. In this sub-section, we document the extent to which headteachers in Wales report negative behaviour of staff as hindering progress within their school.

27. Headteachers were asked the following question in the background questionnaire, with responses given on a four point scale (not at all, very little, to some extent, a lot). Table 8.4 provides the percentage reporting either ‘to some extent’ or ‘a lot’ in Wales to a series of five statements, and compares this to the average across OECD members and the 10 highest-performing countries (H10).

In your school, to what extent is the learning of pupils hindered by the following phenomena?

Table 8.4 Headteachers’ reports of factors hindering pupils’ learning: the conduct of teachers

	Wales	OECD	H10
Teachers not meeting individual pupils' needs	19%	23%	**31%***
Teacher absenteeism	24%	**17%***	**14%***
Staff resisting change	22%	**30%***	**32%***
Teachers being too strict with pupils	4%	**13%***	**16%***
Teachers not being well prepared for classes	17%	**12%***	19%

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools where the headteacher ticks either the ‘to some extent’ or ‘a lot’ categories. Bold font with a * indicates statistically significant difference from Wales at the five per cent level.

28. Around a quarter (24 per cent) of pupils in Wales are taught in schools where the headteacher believes that staff absenteeism acts as a barrier to learning. This is higher than the average across members of the OECD (17 per cent) and the average across the high-performing countries (14 per cent). However, these averages again disguise substantial cross-national variation in headteachers’ responses to this question. Specifically, whereas less than 10 per cent of headteachers report staff absenteeism to be a problem in some high-performing countries (e.g. Singapore, Japan, Canada), this is not the case in others (e.g. in China and Macao around 35 per cent to 40 per cent of pupils are taught in schools where the headteacher views this as a barrier to instruction). Nevertheless, Welsh headteachers’ negative views on staff absenteeism is rather different to the situation reported by headteachers in most of the countries with the highest average PISA science scores.

[59] Miller, Murnane and Willett (2008).

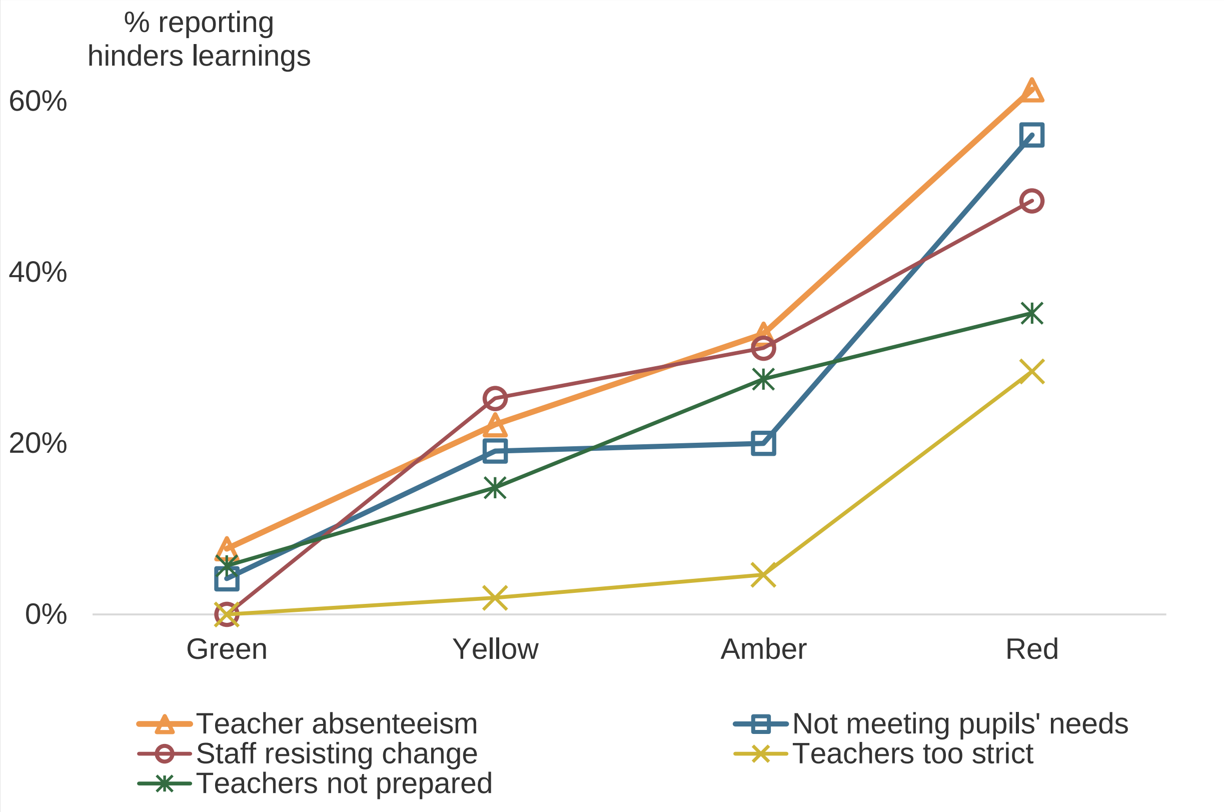
29. In contrast, headteachers in Wales are less likely to report that their staff are resistant to change (22 per cent in Wales versus an H10 average of 32 per cent). Likewise, comparatively few pupils in Wales are taught in schools where the headteacher believes that their staff are too strict (four per cent in Wales versus an average across OECD members of 13 per cent). Therefore, out of all the factors considered in Table 8.4, staff absenteeism seems to be a particularly prominent concern amongst headteachers in Wales (and more so than headteachers in the average industrialised country).

30. In terms of variation within Wales by school categorisation, two issues stand out (see Figure 8.3). First, there is a clear pattern whereby headteachers leading schools with a lower support category are more likely to report staff absenteeism as a problem (orange line with triangular markers). Whereas eight per cent of headteachers who lead a green categorised school agree that staff absenteeism hinders pupils' progression, this increases to 22 per cent in the yellow category, 33 per cent in the amber category and 61 per cent in the red category. The linear trend between school categorisation and the percentage reporting staff absenteeism as a problem is statistically significant at the five per cent level. Staff absenteeism therefore seems an issue of particular concern amongst headteachers who lead schools that require more support.

31. Second, there is also evidence of differences by school categorisation in teachers' preparation for class (green line with star markers in Figure 8.3). Only six per cent of headteachers who lead a green categorised school reported a lack of teacher preparation as an issue, compared to 28 per cent within the amber category and 35 per cent in the red category. The trend between school categorisation and the percentage reporting a lack of teachers' preparation is statistically significant at the five per cent level. A greater emphasis on staff preparation may therefore be key to improving outcomes in schools within the amber and red groups.

32. Another interesting feature of Figure 8.3 is the association between school categorisation and headteachers' views on whether their staff meet pupils' needs (blue line with square markers). This does not appear to be a problem for schools in the green support category; only around four per cent of headteachers leading such schools suggest that this is a challenge that they face. Yet this increases to more than half (56 per cent) of headteachers in the red support group. The linear trend between school support categorisation and the percentage in agreement on whether staff meet pupils' needs is statistically significant at the five per cent level.

Figure 8.3 Headteachers' reports of teachers' conduct by support category



Source: Matched PISA 2015 database

33. The main point of difference between English and Welsh medium schools is headteachers' views of staff absenteeism. Whereas only 10 per cent of headteachers in Welsh medium schools identify this as a problem hindering instruction, this increases to 27 per cent amongst headteachers within English medium schools. This difference is statistically significant at the five per cent level.

Key point

Headteachers in Wales are more likely to report staff absenteeism as a barrier to pupils learning than headteachers in the average OECD or high-performing country. Within Wales, staff absenteeism, teachers not being prepared for class and teachers not meeting pupils' needs are key concerns amongst headteachers who manage schools that require more support.

8.5 What quality assurance processes are used in schools?

34. Robust quality assurance processes are a vital part of any industry. In education, these can take several forms, including external inspections, routine recording of key data, clear specification of the school's goals, and having systems in place to be able to receive regular feedback (from both pupils and their parents).

We already know that the Welsh education system uses some of these quality assurance measures extensively; school inspections as a means of external evaluation, for example. However, less is known about the prevalence of others (e.g. to what extent do schools in Wales have systems in place to receive regular feedback from their pupils?). Table 8.5 therefore provides information on the breadth of the quality assurance processes used in secondary schools in Wales, and how this compares to other countries.

Table 8.5 Headteachers' reports of the quality assurance processes used in secondary schools

	Wales	OECD	H10
Self-evaluation	100%	**93%***	**97%***
External evaluation	97%	**75%***	**80%***
Written specification of the school's curricular profile and educational goals	94%	**89%***	95%
Written specification of pupil performance standards	97%	**79%***	**81%***
Systematic recording of data such as teacher or pupil attendance and professional development	100%	**91%***	**94%***
Systematic recording of pupil test results and graduation rates	100%	**93%***	**95%***
Seeking written feedback from pupils	97%	**69%***	**82%***
Teacher mentoring	98%	**78%***	**89%***
Regular consultation aimed at school improvement with one or more experts over a period of at least six months	91%	**48%***	**49%***
Implementation of a standardised policy for science subjects	83%	**63%***	**75%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils within schools where the headteacher reports the quality assurance process as taking place. Bold font with a * indicates statistically significant difference from Wales at the five per cent level.

35. Wales is clearly a country where extensive quality assurance processes are already in place. Almost every headteacher in Wales reports that self-evaluation, external evaluation, teacher mentoring, systematic recording of pupil data and test results, and written specification of goals and performance standards are used in their school. Indeed, the only area where less than 90 per cent of headteachers respond positively was the implementation of a standardised policy for science (83 per cent). Consequently, all 10 forms of quality assurance listed are used in most secondary schools in Wales.

36. Many of the quality assurance measures listed in Table 8.5 are also extensively used in other industrialised and high-performing countries (e.g. self-evaluation, written specification of goals, systematic reporting of pupil attendance and test scores). Yet there is also evidence of greater use of certain measures in Wales, relative to other countries. This includes more widespread use of consultation with external experts than the average across OECD members (91 per cent versus 48 per cent), greater use of external evaluations (97 per cent versus 75 per cent) and written specification of pupil performance standards (97 per cent versus 79 per cent). It is therefore the breadth of the quality assurance processes used in Welsh schools that is the standout feature of Table 8.5.

37. As Table 8.5 illustrates, external evaluations (such as those conducted by Estyn) are a prominent feature of the quality assurance process used in Wales. However, to what extent do headteachers in Wales use the results from these inspections to drive change? Moreover, do headteachers perceive these inspections to have a lasting impact upon their school?

38. Headteachers were asked to respond yes or no to the following five statements:

- *The results of external evaluations led to changes in school policies*
- *We used the data to plan specific action for school development*
- *We used the data to plan specific action for the improvement of teaching*
- *We put measures derived from the results of external evaluations into practice promptly*
- *The impetus triggered by the external evaluation "disappeared" very quickly at our school*

39. There was near universal agreement amongst headteachers in Wales that school inspections lead to a specific plan of action for school development (97 per cent) and improving teaching (93 per cent), with the measures being put into place promptly (95 per cent). However, around a quarter of headteachers report no change in school policies as a result of the inspections (24 per cent), while around one-in-eight thinks the impetus the inspection triggered disappeared quickly (13 per cent).

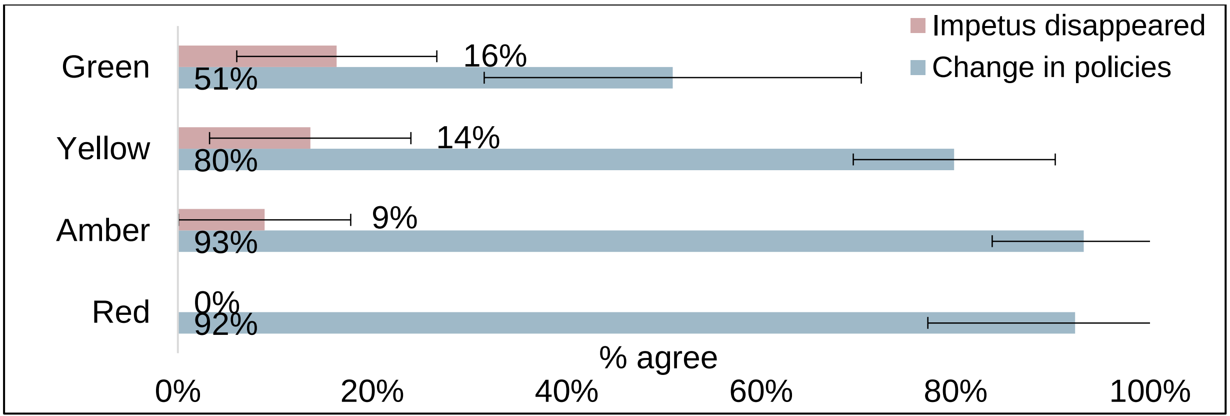
40. Do these figures vary by the support category of the school? This is important as one would hope that the results from external inspections would lead to the greatest sustained change in schools with lower levels of performance. Figure 8.4

therefore illustrates how headteachers' responses vary by National School Categorisation System group.

41. Headteachers of schools in the green support category are less likely to report a change in policy due to external inspection (51 per cent) than the other three groups (where the percentage was around 80 per cent to 90 per cent). This difference is statistically significant at the five per cent level.

42. No school in the red support category reports that the impetus of their last inspection disappeared quickly, while nine per cent of headteachers report this to be the case in the amber group. The figures are slightly higher for headteachers leading schools in the top two support categories (14 per cent for yellow and 16 per cent for green). The linear trend is statistically significant at the five per cent level, though this is mainly being driven by differences between schools in the red category and all other groups. This suggests that, according to headteachers, the impetus triggered by external evaluations does not disappear quickly within Welsh schools (particularly in schools requiring greater levels of support).

Figure 8.4 The reaction of schools in Wales to their last external inspection



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils within schools where the headteacher responds 'yes'. Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval.

Key point

Headteachers report that extensive quality assurance processes are already in place within the Welsh education system.

Chapter 9. Pupils' aspirations and future plans

- Most pupils in Wales view science as relevant to their future, irrespective of their gender, socio-economic status, and skills in this area. There are few notable differences between Wales and high-performing countries in this respect

- The proportion of 15-year-olds who aspire to a career in science is greater in Wales than the average across OECD members.

- Welsh girls are more likely to aspire to work as a health professional than boys. On the other hand, boys are more likely to want to become an engineer than girls.

- The proportion of pupils in Wales who expect to obtain an undergraduate degree is lower than the average across OECD countries.

- Girls in Wales are more likely to expect to complete university than boys. Most 15-year-olds who are planning to apply to university want to attend a Russell Group institution.

- Over a third of 15-year-olds who indicate they are likely to apply to higher education want to study in a university outside of Wales.

1. Young people's aspirations towards future educational and occupational goals are linked to their future attainment[60]. Pupils who aspire to achieve a higher level of education are more likely to do so, even once previous achievement and family background have been taken into account[61]. This means that pupils' goals for their lives post-secondary school can have a real impact upon their outcomes. In this chapter, we investigate how pupils in Wales conceive their lives after finishing school. This includes whether they plan to attend university, what type of career they hope to enter and how this differs between different groups of pupils.

2. As part of the PISA study, pupils were asked about how they view science in relation to their future plans, what level of education they expect to attain and what job they expect to have at age 30. In England, Wales and Northern Ireland, several country specific questions were also added to the pupil questionnaire asking young people to provide further details on their plans regarding higher education. These questions allow us to gain a better understanding of how pupils in Wales view their life and goals beyond secondary school.

3. This chapter seeks to answer the following questions:

- *Do pupils connect studying science in school with future careers?*
- *What types of careers are pupils in Wales interested in? To what extent are 15-year-olds interested in pursuing a career in science?*
- *What are the characteristics of pupils who plan to attend university? What factors are associated with their plans?*

9.1 Do pupils connect studying science with future careers?

4. School forms an important part of the context in which pupils shape their aspirations and expectations[62]. Pupils learn different subjects and make decisions about how their enjoyment of and ability in these subjects might translate into a future career. There is evidence that fewer pupils are interested in 'STEM' (science, technology, engineering and mathematics) careers compared to other fields[63]. For instance, a recent study in the United Kingdom found that pupils aged 10-14 have '*high aspirations, just not for science*'[64]. In this sub-section, we investigate this issue

[60] See Gutman and Akerman (2008) for an overview of the literature on the determinants of aspirations and attainment.
[61] Strand and Winston (2008).
[62] Lupton and Kintrea (2011).
[63] Archer et al. (2013).
[64] Archer et al. (2013: 1).

by considering whether pupils in Wales believe that the material they are taught about science in school is relevant for their future careers.

5. In the background questionnaire, pupils were asked several questions about how important they think school science subjects will be later on in their lives. The results in Table 9.1 show the percentage of pupils who either 'strongly agree' or 'agree' with each statement. For all four questions, the proportion of pupils in agreement is usually similar between Wales and the average across the H10 countries. For instance, 78 per cent of 15-year-olds in Wales agree or strongly agree that school science is something that will *'improve career prospects'*, compared to an H10 average of 76 per cent. On the other hand, pupils in Wales are somewhat more likely to report that school science will help to improve their career prospects than the average across OECD countries (78 per cent for Wales versus 67 per cent OECD average) and will help them to get a job (74 per cent versus 61 per cent). Interestingly, the questions where there are the greatest differences between Wales and the OECD average all explicitly mention words like 'career', 'work' and 'job'. This perhaps indicates that 15-year-olds in Wales make a particularly strong connection between what they learn in school science and their future careers.

Table 9.1 Percentage of pupils who connected school science subjects with future careers

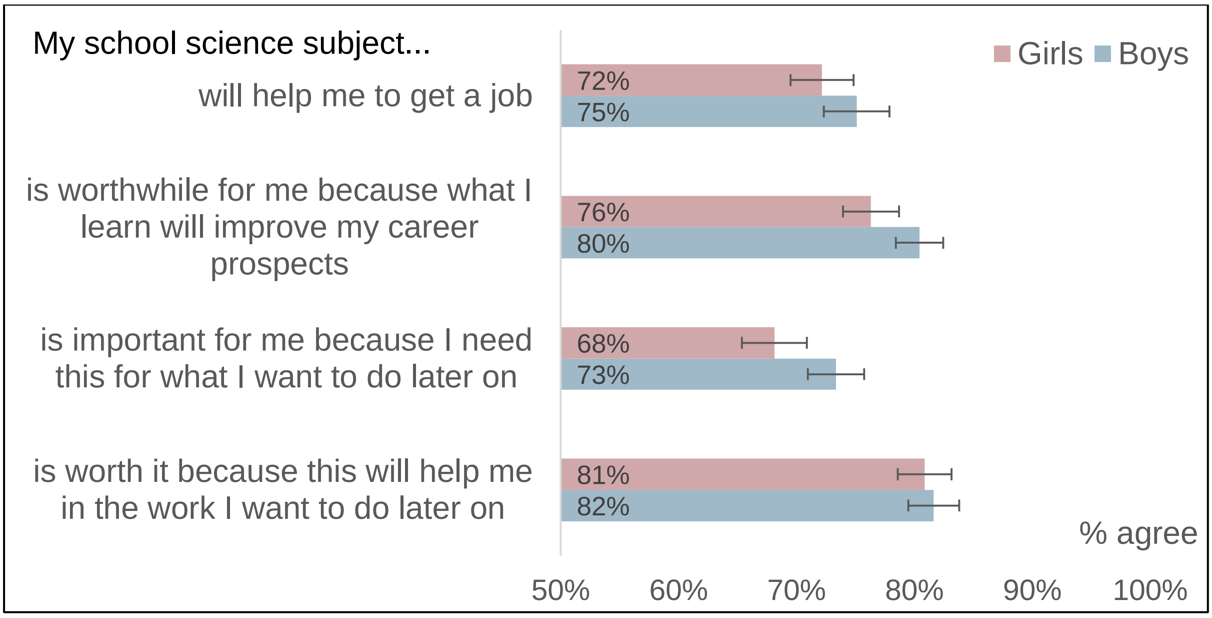
	Wales		OECD		H10	
	2006	**2015**	**2006**	**2015**	**2006**	**2015**
Making an effort in my school science subject(s) is worth it because this will help me in the work I want to do later on	75%	81%	**63%***	**69%***	-	**77%***
What I learn in my school science subject(s) is important for me because I need this for what I want to do later on	57%	71%	56%	**63%***	-	**74%***
Studying my school science subject(s) is worthwhile for me because what I learn will improve my career prospects	75%	78%	**61%***	**67%***	-	**76%***
Many things I learn in my school science subject(s) will help me to get a job	70%	74%	**56%***	**61%***	-	**69%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who either 'strongly agree' or 'agree' with the associated statements. Bold font and * denotes statistically different from Wales at the five per cent significance level. 'H10' refers to the 10 highest performing countries/economies in the PISA science domain. The OECD average for 2006 is the 'OECD-30' (includes 30 OECD members as of 2006) and the OECD average for 2015 is the 'OECD-35' (includes all 35 OECD members as of 2015). We do not calculate the H10 average for 2006 since different countries were the top science performers in that PISA cycle. In 2006, the second statement was worded slightly differently: "What I learn in my school science subject(s) is important for me because I need this for what I want to *study* later on" [emphasis added].

6. The PISA 2006 cycle included the same questions, which provides an opportunity to investigate how pupil's responses have changed over time. For every statement, both pupils in Wales and the average OECD country have become more likely to view science as important to their future since 2006. In Wales, pupils in 2015 were approximately 10 percentage points more likely to respond to these statements with 'agree' or 'strongly agree' than in 2006. In 2006, pupils Wales were also still more likely to answer these questions with 'agree' or 'strongly agree' than their peers in the average OECD country. Overall, it therefore seems that similar findings emerge for Wales regarding pupils' views on the relevance of school science subjects in 2015 as occurred in 2006.

Figure 9.1 Percentage of pupils who connect school science subjects with future careers: by gender



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who either 'strongly agree' or 'agree' with the associated statements. Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval.

7. Figure 9.1 turns to whether responses to these questions in 2015 differed by gender. Overall, boys in Wales were more likely to agree with each of these statements than girls; however, the differences are only statistically significant on two occasions. The first is that a total of 73 per cent of boys 'agree' or 'strongly agree' that science is something they need for what they want to do later on, compared to 68 per cent of girls. The second is that 80 per cent of boys also felt that studying science is worthwhile because it will improve their career prospects, while 76 per cent of girls felt the same way. However, it should be noted that these results are not

specific to Wales; gender differences in pupils' responses to these statements are also relatively small in terms of magnitude (three to four percentage points) for the average across OECD countries. Nevertheless, there is some limited evidence that boys in Wales are slightly more likely to make the connection between school science subjects and future careers than girls.

8. It is notable how the majority of Welsh pupils who lack basic science skills still believe that what they learn in their science classes is relevant for their future employment prospects. Indeed, even amongst pupils with low science skills, over two thirds responded positively to each of the statements. In additional analysis, we have also found little evidence that pupils' responses to these questions differ markedly by either socio-economic status or language of instruction in school English or Welsh medium. There are, however, some interesting differences between pupils who achieved different scores on the PISA science test. Table 9.2 indicates that the top performing pupils (Levels 5 and 6) are 17 percentage points more likely than their low achieving peers (below Level 2) to think that science is worthwhile for improving career prospects (90 per cent versus 73 per cent). Similarly, they are 17 percentage points more likely to think that what they learn in their school science subjects will help them get a job (87 per cent versus 70 per cent).

Table 9.2 Percentage of pupils who connected school science subjects with future careers by science proficiency level

	Below Level 2	**Levels 2-4**	**Levels 5 or 6**
Making an effort in my school science subject(s) is worth it because this will help me in the work I want to do later on	78%	81%	**90%***
What I learn in my school science subject(s) is important for me because I need this for what I want to do later on	71%	70%	**85%***
Studying my school science subject(s) is worthwhile for me because what I learn will improve my career prospects	73%	**79%***	**90%***
Many things I learn in my school science subject(s) will help me to get a job	70%	74%	**87%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in Welsh schools who either 'strongly agree' or 'agree' with the associated statements. 'Levels' refer to PISA science proficiency levels. 'Below Level 2' includes Levels 1a, 1b and those pupils below Level 1. Bold font and * indicates significantly different from pupils below Level 2 at the five per cent level.

Key point

Most pupils in Wales view school science as relevant to their future, irrespective of their gender, socio-economic status, and proficiency in this area. There are few notable differences between Wales and high-performing countries in this respect.

9.2 What types of careers interest pupils? To what extent are 15-year-olds interested in a career in science?

9. Adolescence and the end of secondary school represent an important transitional period in an individual's life. Pupils have to make important career-related decisions about the direction in which their lives will go. They will decide whether to enter vocational training, pursue a university degree or enter directly into the labour market. There is evidence that pupils who set and pursue goals are better equipped to master this transition[65]. The pupils who take PISA find themselves in this crucial period, and have been asked the following question about their future occupational goals: *What kind of job do you expect to have when you are about 30 years old[66]?*

10. The most popular future occupation amongst 15-year-olds in Wales is 'engineer'; six per cent expect to be working in this role at age 30. The second most popular occupation is 'medical doctor', with approximately four per cent of pupils, followed by 'creative and performing artists' in third place, also with approximately four per cent. 'Other health professionals' also made it into the top 10 with three per cent of pupils, as did 'nursing and midwifery' and 'architects, surveyors, planners, and designers', each with approximately three per cent of pupils. It is notable that five of the 10 most aspired to careers in Wales are science careers. Pupils in Wales exhibit some uncertainty in their future career aspirations; 18 per cent of 15-year-olds either did not answer the question or answered with 'do not know' or something vague as their response.

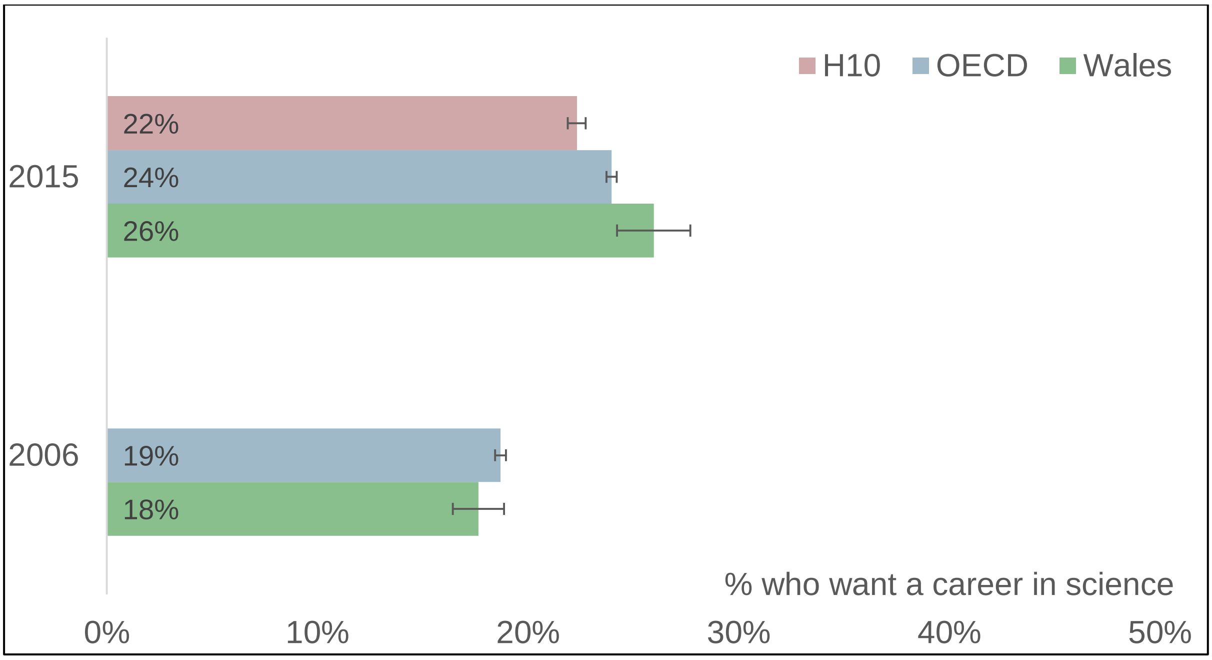
11. Figure 9.2 illustrates that, in total, just over a quarter of pupils in Wales (26 per cent) expect to work in a STEM ('science, technology, engineering and mathematics') career[67]. This is around two percentage points above the average across OECD members (24 per cent) and the average across H10 countries (22 per cent). Interestingly, Figure 9.2 also reveals that this is somewhat different to the

---

[65] See Weiss et al. (2014) for an overview of the motivational, personal and contextual factors affecting the completion of secondary school and the transition to life post-secondary school.
[66] Pupils provided a free text answer, with these then converted by the survey organisers into International Standard Classification of Occupations 2008 (ISCO-08) codes.
[67] We follow the OECD's definition of a career in science. See Annex A10 in the PISA International Report Volume 1, Chapter 3 for a list of the included occupations.

situation in PISA 2006, when science was last the focus of PISA[68]. For instance, only one-in-five (18 per cent) Welsh pupils aspired to a science career in 2006, which was little different to the average across OECD countries (19 per cent) [69]. It therefore seems that there has been a notable increase in the proportion of Welsh pupils who are interested in pursuing a STEM career over the last decade.

Figure 9.2 The percentage of pupils who aspire to a career in science: a comparison between PISA 2006 and 2015



Source: PISA 2015 database

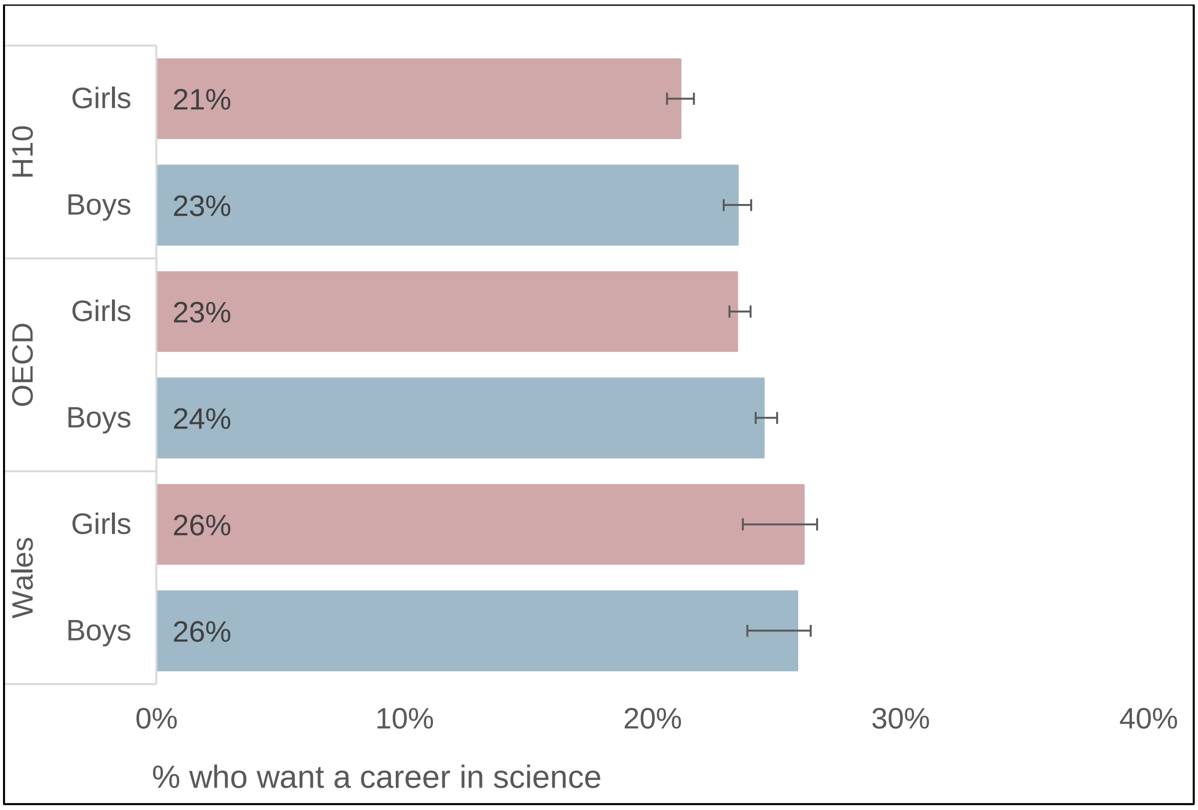
Notes: Figures refer to the percentage of pupils in schools who aspire to career in science at age 30. We do not compute the H10 average for 2006 since the high-performers in that year were different from the high-performers in 2015. The OECD average for 2006 is the 'OECD-30' (includes 30 OECD members as of 2006) and the OECD average for 2015 is the 'OECD-35' (includes all 35 OECD members as of 2015). Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval. It should be noted that the 2015 figure presented here for Wales differs slightly from the OECD international results Table I.3.10. This is because the United Kingdom initially submitted ISCO-08 three digit codes to the OECD for use in their international report, while we were able to use recoded data that included four digit codes in this national report. This is why they report 28 per cent of pupils aspiring to a science career while we report 26 per cent.

[68] For the PISA 2006 survey, the older ISCO-88 classification of occupations was used, not the ISCO-08 as in 2015. The ILO has linked the ISCO-88 and the ISCO-08, so that they are comparable, and the OECD has taken this into account in the construction of the science career variable for 2006 and 2015.

[69] The OECD average for 2006 is the 'OECD-30' (which includes the 30 OECD members as of 2006) and the OECD average for 2015 is the 'OECD-35' (which includes all 35 OECD members as of 2015).

12. Are there significant socio-economic differences in aspirations towards a career in science in Wales? Our analysis shows that pupils from disadvantaged backgrounds in Wales are 15 percentage points less likely to aspire to a STEM career than their peers from advantaged backgrounds (20 per cent versus 35 per cent). This gap exists amongst OECD countries on average as well, where there is a 13 percentage point difference between pupils from socio-economically advantaged and disadvantaged backgrounds (18 per cent versus 31 per cent). These results indicate that socio-economic disadvantage translates into different career aspirations and a decreased desire to pursue a career in science; this is despite low socio-economic status pupils being no less likely to believe that science is relevant for their future (recall sub-section 9.1).

Figure 9.3 Gender differences in aspirations towards a science career



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who aspire to a career in science at age 30. Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval. It should be noted that the figures presented here for Wales differ slightly from the OECD international results Table I.3.10. This is because the United Kingdom initially submitted ISCO-08 three digit codes to the OECD for use in their international report, while we were able to use recoded data that included four digit codes in this national report.

13. Figure 9.3 illustrates differences in aspirations to be working in a STEM career at age 30 between boys and girls[70]. There is no evidence of a gender gap in science aspirations in Wales. Specifically, 26 per cent of boys and girls aspire to a career in science. Although a similar finding holds for the OECD and H10 averages, there are some important exceptions within these groups. In Taiwan, for example, boys are 10 percentage points more likely to express interest in a science related career than girls (26 per cent versus 16 per cent). A similar sized gender gap of eight percentage points exists in Singapore (32 per cent of boys versus 24 per cent of girls). In high-performing Western countries, there tends to be no gender gap or a small gender gap in favour of girls. For example, there is around a five percentage point difference in science aspirations in Canada, but this is in favour of girls (31 per cent of boys versus 37 per cent of girls).

14. In Table 9.3 we break down the type of science career pupils aspire to into four broad groups: scientist/engineer, health professional, ICT professional and technician. Twenty per cent of Welsh girls are interested in a career as a health professional, compared to six per cent of boys. On the other hand, Welsh boys are more likely to aspire to become a scientist/engineer than girls (15 per cent versus five per cent). The magnitude of these gender differences is similarly large for the average across OECD members; there is an 11 percentage point difference between boys and girls with regards to working as a health related professional, for instance. There are hence pronounced gender differences in the specific types of scientific career 15-year-olds in Wales hope to enter.

Table 9.3 Gender differences in aspirations towards different STEM careers

	Wales			**OECD**			**H10**		
	Total	**Boys**	**Girls**	**Total**	**Boys**	**Girls**	**Total**	**Boys**	**Girls**
Scientist/engineer	11%	15%	**5%***	9%	12%	**5%***	8%	11%	**4%***
Health professional	13%	6%	**20%***	11%	6%	**17%***	11%	7%	**16%***
ICT professional	3%	4%	**1%***	3%	5%	**0%***	3%	5%	**1%***
Technician	0%	1%	0%	1%	2%	**1%***	1%	1%	**1%***

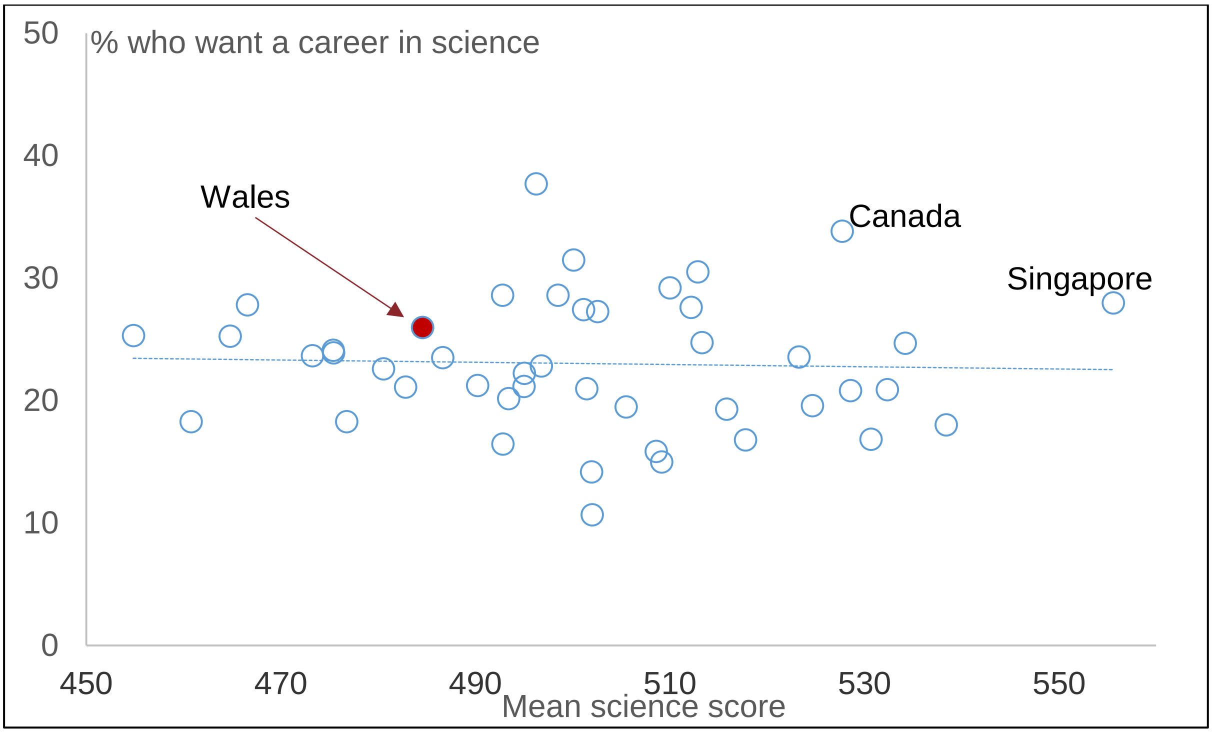
Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who aspire to a career in science in one of these four categories at age 30. It should be noted that the figures presented here for Wales differ slightly from the OECD international results Table I.3.10. This is because the United Kingdom initially submitted ISCO-08 three digit codes to the OECD for use in their international report, while we were able to use recoded data that included four digit codes in this national report. Bold font and * denotes girls statistically different from boys at the five per cent significance level.

[70] See Mau (2003) and Sadler et al. (2012) for an overview of evidence on STEM career choice and gender.

15. Do the countries with the highest average scores also have the greatest proportion of pupils who want to become scientists? Figure 9.4 provides the answer by plotting average PISA science scores (horizontal axis) against the percentage of pupils who aspire to a career in science (vertical axis). The flat trend line in Figure 9.4 indicates a weak correlation of -0.04; countries with the strongest performance in PISA do not necessarily have the highest percentage of pupils who want to work in a STEM career. In fact, of the 10 countries with the highest average PISA science scores, only Canada and Singapore have greater proportions of 15-year-olds who aspire to a science career than Wales.

Figure 9.4 PISA science performance and STEM aspirations



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who aspire to a career in science at age 30 and the country average score in the PISA science domain. It should be noted that the figure presented here for Wales differs slightly from the OECD international results Table I.3.10. This is because the United Kingdom initially submitted ISCO-08 three digit codes to the OECD for use in their international report, while we were able to use recoded data that included four digit codes in this national report.

Key point

15-year-olds in Wales are more likely to aspire to a science career than pupils in the average OECD or average top performing country. Girls are more likely to aspire to work in a career as a health professional, while boys are more likely to want to become an engineer.

9.3 What are the characteristics of pupils who plan to attend university? What factors are associated with their plans?

16. In this sub-section we gain further insight into the university aspirations and plans of 15-year-olds in Wales. There is evidence that although access to university in the United Kingdom has increased over time, enrolment rates for pupils from advantaged backgrounds remain much higher than for those from disadvantaged backgrounds, especially within higher status degree programmes[71]. One mechanism that has been proposed to explain this is the university application process, with young people from disadvantaged backgrounds being much less likely to apply to university than their academically equal but more advantaged peers[72]. We use data from the PISA background questionnaire to look at who intends to apply to university and the factors that are associated with their plans.

Table 9.4 The percentage of 15-year-olds who expect to obtain at least an undergraduate degree

	Wales	OECD	H10
Overall	35%	45%	52%
Boys	30%	40%	49%
Girls	**40%***	**49%***	**56%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who expect to obtain at least an undergraduate degree. Bold font and * indicates girls are significantly different from boys at the five per cent level. Due to lack of data for Slovakia and Vietnam and inconsistencies in the data for Finland and Taiwan, we have excluded them from the calculation of the H10/OECD averages.

17. As part of the background questionnaire, pupils were asked what level of education they expect to complete. Table 9.4 shows that 35 per cent of pupils in Wales expect to obtain at least a bachelor's degree[73]. This is lower than the average across the OECD (45 per cent) and across the top-performing countries (52 per cent). Still, there is a lot of variation between countries; less than one-in-five German 15-year-olds expects to complete university compared to around three-quarters in the United States (76 per cent). Amongst the high-performers, there are also countries such as Canada (63 per cent), where a much larger proportion of 15-year-olds expect to obtain an undergraduate qualification than in others, such as China (38 per cent).

[71] Boliver (2011).
[72] Anders (2012).
[73] This corresponds to International Standard Classification of Education (ISCED) level 5A or 6, which is a framework created by the United Nations Educational, Scientific and Cultural Organisation (UNESCO) to standardise education levels across countries. Level 5A or 6 is at least a bachelor's degree, but also includes master's degrees, doctorates and other graduate degrees.

18. Table 9.4 also illustrates how girls are 10 percentage points more likely to say they will complete university than boys (30 per cent for boys versus 40 per cent for girls). This difference is statistically significant at the five per cent level, and is consistent with the 2013/14 Higher Education Initial Participation Rate[74], where there is a nine percentage point difference in university enrolment between boys (42 per cent) and girls (51 per cent). The gender gap in university expectations is also of a similar magnitude to the average across OECD members (nine percentage points) and the average across high-performing countries (seven percentage points).

19. Similarly, we also find differences in university expectations by pupils' socio-economic background. Specifically, over half (56 per cent) of Welsh pupils from the most advantaged backgrounds expect to complete university, compared to 21 per cent of their peers from disadvantaged backgrounds. This is a difference of over 30 percentage points, and is similar in size to the equivalent difference in the top performing countries (33 per cent of disadvantaged pupils versus 78 per cent of advantaged pupils) and the average across OECD members (27 per cent of disadvantaged pupils versus 66 per cent of advantaged pupils). There are no statistically significant differences in pupils expecting to complete university based on language of tuition, whether English or Welsh medium.

20. Pupils in Wales also answered a series of questions on the university application process (see Table 9.5)[75]. Only pupils who stated that they were likely to apply to university were given the opportunity to respond to these questions. A total of 65 per cent of the full sample indicated that they were 'fairly likely' or 'very likely' to apply to university. The remaining 35 per cent of the sample was divided between pupils who said they were 'not very likely' or 'not likely at all' to attend university (19%) and pupils who skipped this question (16%). This should be kept in mind when interpreting the following results.

21. Course / course content (98 per cent), employment prospects after graduation (97 per cent) and realistic entry requirements (93 per cent) are the three most important factors in 15-year-olds' higher education plans. This holds true for both boys and girls. On the other hand, factors related to social life are somewhat less important to the plans of 15-year-olds, as are university costs. For instance, around 15 per cent of pupils in Wales do not view cost to be an important factor in their

[74] This is the sum of age specific initial participation rates in the age range of 18-30. Since most people first start university in the UK at age 18, this is the age group that dominates the statistic (Department for Business, Innovation and Skills, 2015).
[75] These questions were only posed to pupils in England, Wales and Northern Ireland, and not in other countries.

higher education plans. Finally, the least important issue is distance from home, with 60 per cent of 15-year-olds in Wales saying this would not be an important factor in determining which higher education institution they will apply to. Young people in Wales therefore seem to take a pragmatic approach when thinking about which university to apply to, focusing upon the practicalities of the course and the application process, as well as eventual employment outcomes. Nevertheless, for all factors more than half of the pupils who responded report the factor to be either 'fairly' or 'very' important, highlighting how pupils in Wales take into account a wide range of factors when forming their higher education plans.

Table 9.5 Percentage of pupils who feel certain factors matter for university application decisions

	Percentage who feel it is important				
	Total	**Boys**	**Girls**	**Bottom 25% SES**	**Top 25% SES**
Course / course content	98%	97%	**98%***	98%	98%
Employment prospects afterward	97%	95%	**98%***	95%	**98%***
Realistic entry requirements	93%	93%	94%	93%	92%
Challenging entry requirements	87%	86%	87%	87%	85%
Local employment prospects while a student	84%	81%	**86%***	88%	**77%***
Costs (as affected by fees, scholarships and bursaries)	85%	83%	**87%***	90%	**78%***
Academic ranking / 'league table' ranking	84%	83%	84%	83%	86%
Social life	82%	83%	81%	82%	83%
Fitting in	76%	76%	76%	76%	77%
Distance from home	60%	59%	61%	64%	**55%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in Welsh schools who responded to these questions, not the entire sample, and feel that these factors are either 'very important' or 'fairly important'. Bold font and * indicates significantly different from boys when in the column for girls or significantly different from the bottom quartile of socio-economic status when in the column for the top quartile of socio-economic status at the five per cent level.

22. There is surprisingly little difference in how pupils from different socio-economic backgrounds plan their university applications. The only exceptions are with respect to 'distance from home' (64 percent of disadvantaged pupils versus 55 per cent of advantaged pupils reported this to be an important factor), cost (90 per cent of disadvantaged pupils versus 78 per cent of advantaged pupils) and local employment prospects (88 per cent of disadvantaged pupils versus 77 per cent of

advantaged pupils). This result suggests that financial considerations are likely to have more of an influence upon the decision of pupils from disadvantaged backgrounds. Nevertheless, for the most part, differences in responses by socio-economic status were relatively muted.

23. Pupils were also asked to list three universities to which they might apply[76]. Just over half (57 per cent) of pupils in Wales who are planning to apply to university listed a Welsh higher education institution as their first choice. Cardiff University was listed by 16 per cent of pupils, followed by Swansea University (five per cent). Over a third of pupils in Wales who are planning to apply to university would like to attend a university somewhere else in the UK. There is no evidence of gender differences in pupils' responses.

24. The majority of pupils who plan to apply to university aspire to attend a Russell Group university[77]; 58 per cent of pupils who answered this question list a Russell Group university as their first choice. As a point of comparison, in 2014/15, only 23 per cent of undergraduate pupils in the UK were enrolled in such an institution[78]. It is therefore clear that many more 15-year-olds aspire to the top universities than the proportion who will go on to attend them.

25. There is also evidence of a socio-economic gap in terms of the type of the institution 15-year-olds hope to attend. Specifically, young people from the most advantaged socio-economic backgrounds are nine percentage points more likely (62 per cent) to aspire to attend a Russell Group university than their peers (53 per cent) from disadvantaged socio-economic backgrounds.

Key point

The proportion of pupils in Wales who expect to obtain a bachelor's degree is below the OECD average.

Over a third of 15-year-olds who are likely to apply to higher education want to study in a university outside of Wales.

[76] These answers were entered as free text, so pupils had to draw on their own knowledge of universities to answer these questions. Again, pupils only provided answers to these questions if they stated they were planning on applying to university.
[77] The Russell Group is a network of 24 universities in the United Kingdom committed to 'maintaining the very best research, an outstanding teaching and learning experience and unrivalled links with business and the private sector' (Russell Group, 2016).
[78] Based on authors' calculation using Higher Education Statistics Agency (HESA) data on undergraduate university enrolments from 2014/15 (HESA, 2016).

Chapter 10. Pupils' experiences of learning science in school

- 15-year-olds in Wales report spending more time studying science in school per week than young people in other OECD countries.
- The total amount of time 15-year-olds in Wales report spending on additional study is above the average across OECD members.
- Pupils in Wales feel they have similar opportunities to express themselves during science lessons and to draw conclusions from experiments as their peers in OECD countries. However, they report spending less time constructing arguments and engaging in debates during science lessons than pupils in the average industrialised country.
- The reported frequency of low-level disruption in Welsh science classrooms is greater than the OECD average and the average across the high-performing countries.
- Science teachers in Wales are reported as providing more regular feedback to pupils on their strengths and weakness, including specific areas they can improve, than teachers in many of the countries with the highest average PISA scores.
- Pupils in Wales generally perceive their science teachers to be supportive. However, low achieving pupils report that their science teacher is less willing to provide individual help and adapt their lessons than their high achieving peers.

1. The time pupils spend in school, learning and interacting with their teachers and their peers, plays a critical role in determining their learning outcomes[79]. There remain important gaps in our knowledge about pupils' experiences whilst in school, including the activities they complete in the science classroom. For instance, how much time do pupils in Wales spend studying science relative to other subject areas per week? Do they receive regular feedback from their teachers as part of their science lessons? Is the environment in the classroom conducive to learning, or do pupils feel that their progress is being hampered due to frequent occurrences of low-level disruption? The aim of this chapter is to provide new evidence on these issues for Wales, and whether the experiences of learning science in school for 15-year-olds in this country are similar to those of young people in other parts of the world. Specifically, this chapter seeks to answer the following questions:

- *How much time do pupils spend studying science in-school and out-of-school per week? How does this compare to other subject areas?*
- *What kind of activities take place in science classrooms in Wales? Does this differ markedly from other countries?*
- *Is low-level disruption in science classrooms a more common occurrence in Wales than in other countries?*
- *How do pupils in Wales perceive the feedback that they receive from their science teachers?*
- *Do pupils in Wales feel that they receive sufficient support from their teachers during their science classes?*

2. It should be noted that we attempt to answer these questions by drawing upon information reported by the 15-year-olds who responded to the PISA background questionnaire. The subjective nature of their views, and limitations in their ability to accurately recall and report information, should be considered when interpreting the results.

[79] See Sacerdote (2011) for an overview of how pupils may have an impact upon the learning of their peers.

10.1 How much time do pupils spend studying science per week? How does this compare to other subject areas?

3. It has been suggested that increasing instruction time in school can, up to a point, improve pupils' learning outcomes (particularly for those from disadvantaged socio-economic backgrounds) [80]. At the same time, certain forms of out-of-school study, such as intensive one-to-one tuition, are thought to be particularly effective in raising pupils' attainment[81]. It is therefore important to know how much time pupils in Wales spend studying different subjects, both within their compulsory timetable at school and beyond. In this sub-section we therefore explore the amount of time pupils report spending on a selection of subjects (a) within their core timetable and (b) in additional time, before and/or after school.

4. Table 10.1 documents the average number of hours pupils report spending on a selection of subjects as part of their core timetable per week. Figures are provided for science, home language, mathematics and 'other' subject areas[82].

Table 10.1 The average number of in-school instruction hours per week

	Wales	OECD	H10
Science	5.0 hours	**3.5 hours***	**4.0 hours***
English/test language	4.1 hours	**3.6 hours***	**4.1 hours**
Mathematics	3.9 hours	**3.6 hours***	**4.3 hours***
Other	14.1 hours	**16.6 hours***	**15.9 hours***
Total	26.6 hours	**26.9 hours***	**28.0 hours***

Source: PISA 2015 database

Notes: Figures refer to the average weekly hours of in-school instruction time, as reported by pupils. 'Other' is the difference between the sum of reported subjects and the reported total. Due to missing values, the reported subjects and the 'other' category do not sum to the reported total. Bold font and * denotes statistically different from Wales at the five per cent significance level. Data not available for Vietnam, which has therefore been excluded from the calculation of the H10 average.

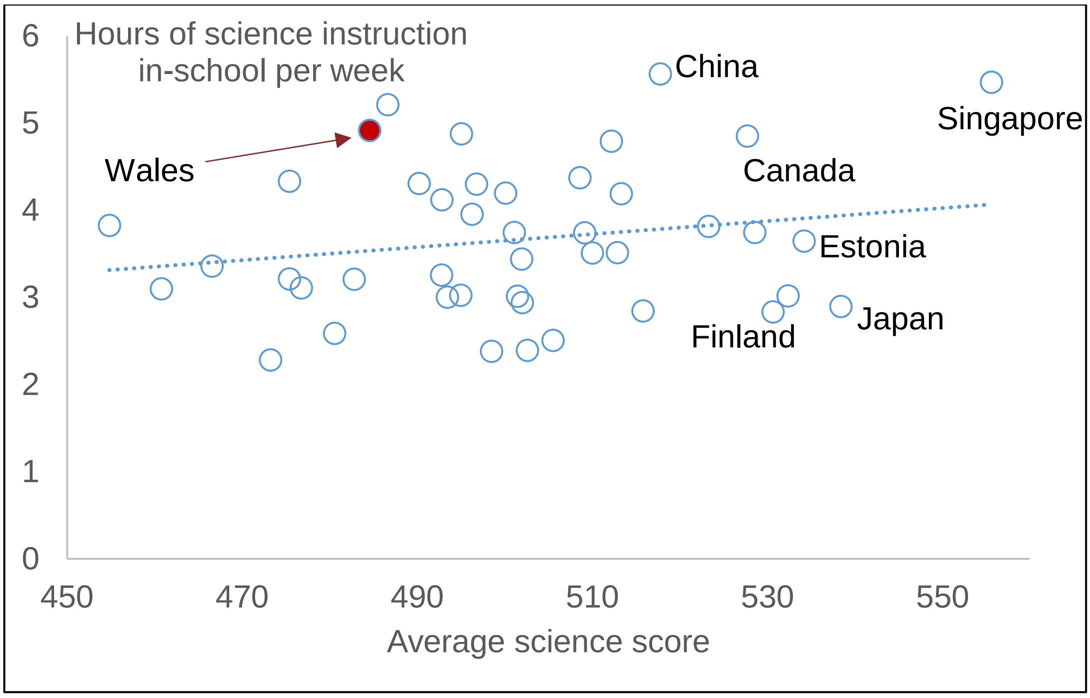
5. Pupils in Wales report receiving, on average, five hours of science instruction per week. This equates to approximately one fifth of their 27 hour weekly timetable. This is greater than the amount of time for either test language (English/Welsh, almost one hour less) or mathematics (over an hour less). This is not the case for the

[80] See Hanushek (2015) for an overview of the evidence on instruction time and pupil performance.
[81] Higgins et al. (2014).
[82] The online data tables provide additional estimates based upon the median number of hours reported, rather than the mean. These results are less likely to be affected by a small number of pupils who report very large values in response to the questions regarding the time they spend studying in-school and out-of-school.

average across OECD and H10 countries, where the average number of hours is roughly the same for science, test language and mathematics.

6. Overall, pupils in Wales report having a similar amount of total timetabled hours in-school per week as pupils in the average industrialised country, but fewer timetabled hours as compared to top performing countries (one and a half hours less). There are also some differences in how these hours are distributed across the various subject areas. For example, pupils in Wales report spending 90 minutes more per week studying science in-school as compared to the OECD average, and 60 minutes more than the average across the top performing science. Pupils in Wales also report receiving 30 minutes more weekly instruction in the language of the test than the OECD average. On the other hand, pupils in Wales report spending substantially less time in-school learning 'other' subject areas (14 hours versus an OECD/H10 average of around 16 hours).

Figure 10.1 The relationship between hours of science instruction in-school and average PISA science scores



Source: PISA 2015 database

Notes: The sample of countries has been restricted to those with an average science score above 450 points. Data not available for Vietnam and Malta.

7. Although PISA is not directly linked to the curriculum, the amount of time pupils spend learning science in-school may nevertheless be associated with their achievement. Figure 10.1 therefore investigates whether in-school instruction time in science is linked to performance in this subject at the country level.

8. There are two noteworthy features of this graph. First, Wales is at one of the uppermost points of the graph. This illustrates how there are few countries where average weekly science instruction time in-school is higher than in Wales. Indeed, 15-year-olds in most other countries typically spend at least an hour less time learning science in school per week. Second, as illustrated by the dashed regression line, the relationship between in-school instruction hours and average PISA test scores in science is relatively weak at the country level (correlation = 0.19). For instance, in some high-performing countries, pupils report as little as three hours of timetabled science lessons per week (e.g. Japan, Finland), while in others (e.g. Canada, Singapore) the average amount of time spent is similar to the five hours in Wales. Consequently, there is little evidence that countries with more timetabled hours for science necessarily achieve higher average PISA scores.

Table 10.2 Average hours spent on additional learning per week

	Wales	OECD	H10
Science	3.9 hours	**3.1 hours***	**3.4 hours***
English/test language	3.6 hours	**3.1 hours***	**3.2 hours***
Mathematics	4.0 hours	**3.8 hours***	**4.3 hours***
Foreign language	1.3 hours	**3.1 hours***	**3.1 hours***
Other subjects	5.1 hours	**3.9 hours***	**3.8 hours***
Total	17.9 hours	**17.1 hours***	17.8 hours

Source: PISA 2015 database

Notes: Figures refer to the average hours of additional learning time per week, as reported by pupils. Data not available for Vietnam, which has therefore been excluded from the calculation of the H10 average. Due to missing values, the reported subjects do not necessarily sum to the reported 'total' category. Bold font and * denotes statistically different from Wales at the five per cent significance level.

9. It is of course possible for pupils to increase the amount of time they spend studying per week via out-of-school learning. This information has also been captured in the PISA background questionnaire, with pupils asked: '*approximately how many hours per week do you spend learning in addition to your required school schedule*?' Pupils were instructed to include time spent upon homework, additional instruction and private study in their responses. Table 10.2 presents the average amount of time pupils report spending on science, mathematics, first language (English/Welsh), foreign language and 'other' subject areas[83]. Results using the median are provided in the online data tables.

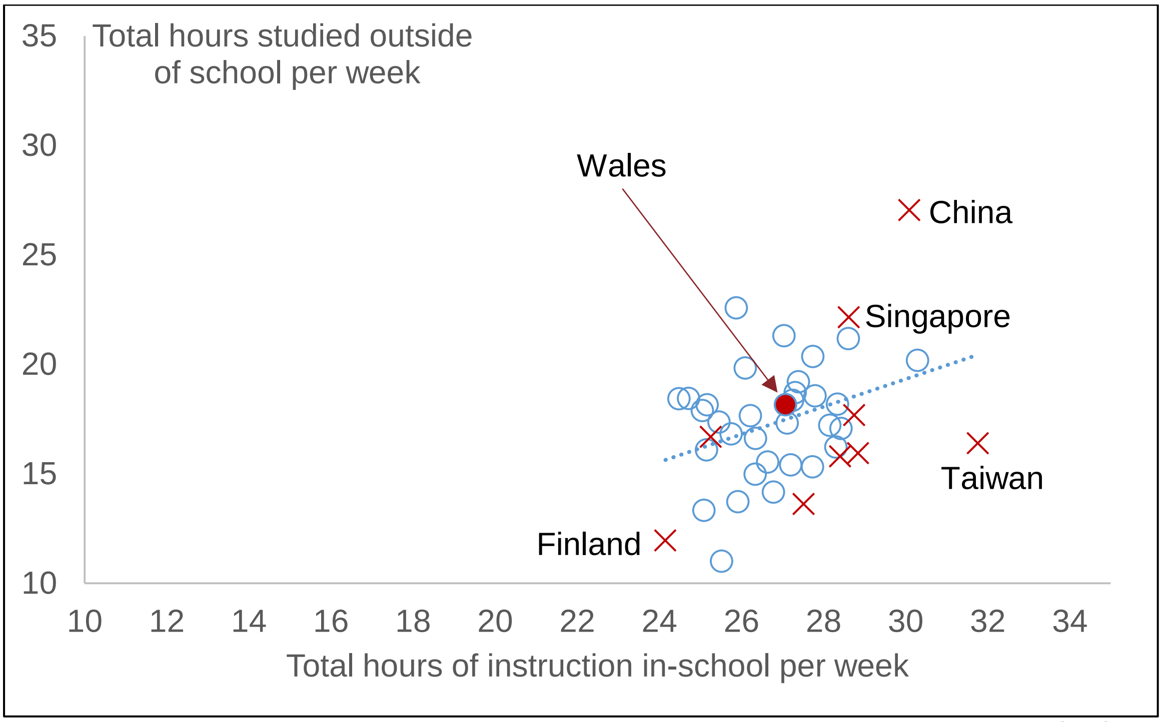
[83] Any pupil who reported spending more than 70 hours per week on additional study is treated as reporting an illogical value, and therefore excluded from this part of our analysis.

10. There are some key points of difference between the figures for Wales and the average across OECD / H10 countries. Specifically, the average number of additional learning hours is higher for Wales than the H10/OECD average in science (around 30 minutes or more per week) and in the ‘other’ category (over an hour higher). In contrast, less additional time in Wales is spent on learning foreign languages (one hour and fifty minutes less per week). Therefore, although the total number of additional learning hours is similar for the average pupil in Wales and the average across H10 countries (approximately 18 hours), there are some important differences in how this is distributed across various subject areas.

11. Do pupils report spending less time on additional study in countries with a longer school day? In other words, is there evidence of a substitution effect, whereby more hours in the school timetable is offset by less time spent on additional study? Figure 10.2 provides the answer by plotting the total timetabled hours per week for the average pupil (horizontal axis) against the total additional learning hours (vertical axis). The sample has been restricted to countries with an average PISA science score above 450 points, with the 10 countries with the highest average PISA science scores highlighted using a red cross.

12. All countries sit towards the bottom right hand corner of Figure 10.2. This indicates how, in every country, the average pupil reports spending more time studying in-school than they do on additional instruction outside of regular school hours. However, there is also substantial cross-national variation in these figures, including across the high-performing countries. At one extreme sits China, where the average pupil reports spending 30 hours per week studying in-school, accompanied by 27 hours of additional study. This is notably higher than the 27 hours (in-school) and 18 hours (additional instruction) in Wales. Weekly hours are, on the other hand, much lower in Finland, where the average 15-year-old spends 24 hours learning in school and 12 hours on additional instruction. There are also some notable outliers, such as Taiwan, where in-school instruction time is reported higher than any other country included in the comparison (32 hours), though with additional study time around the international average (16 hours). When these facts are brought together, they highlight two important points for Wales: (a) the 18 hours of additional instruction time reported by the average 15-year-old in Wales does not stand out as particularly high or low relative to pupils in most other countries and (b) China and Singapore are the only high-performing countries where total additional study hours are reported to be much higher than in Wales.

Figure 10.2 The relationship between in-school and out-of-school learning hours per week



Source: PISA 2015 database

Notes: Figures refer to the total number of weekly hours of in-school instruction (horizontal axis) and the total number of additional hours of study (vertical axis) as reported by the average pupil. Sample restricted to countries with a mean science score above 450 points. Data not available for Vietnam and Malta. Red crosses denote the top 10 performing countries on the PISA science domain.

13. The other key conclusion to be drawn from Figure 10.2 is that there is little evidence of a trade-off between in-school and additional learning hours at the country level. In fact, the cross-country correlation is weakly positive (0.35), indicating that the average pupil spends slightly more time on additional study in countries with more hours in the weekly timetable.

14. There was little evidence of gender differences in additional study time in any subject area, or for total hours overall. Similarly, the only statistically significant difference with respect to socio-economic status is that those from disadvantaged backgrounds in Wales report spending approximately 30 minutes less per week on additional science learning than their peers from advantaged backgrounds (3.5 hours versus 4.0 hours).

Key point

15-year-olds in Wales report spending 90 minutes more studying science in school per week than the average pupil across OECD countries. Pupils in Wales also report spending more time studying science outside of school than the average across OECD members and the average across high-performing countries.

10.2 What activities take place in science classrooms in Wales? Is this similar to other countries?

15. The science curriculum in Wales is designed to help pupils develop their understanding of the relationships between 'data, evidence, theories and explanation' and 'evaluate enquiry methods and conclusions'[84]. Science teachers have a critical role in helping young people to reach these goals, including through the activities that take place in their classrooms. Yet what are the activities that actually take place in school science lessons in Wales? Do pupils regularly design and conduct their own experiments? Or is more time spent on activities that require reasoning and constructing an argument, such as class debates? PISA provides us with an opportunity to take a glimpse inside science classrooms in Wales, allowing us to better understand the types of tasks that pupils complete.

16. Table 10.3 illustrates the extent to which a series of different practices and activities are used in science classroom in Wales, and how this compares to other parts of the world. This includes the opportunities pupils have to explain their ideas, to design their own experiments, and the extent to which pupils believe that their teacher clearly explains the relevance of science concepts to their lives. All figures refer to the proportion of 15-year-olds who stated that the activity or practice happens in 'every' or in 'most' science lessons (as opposed to 'some' or 'never').

17. There are some important similarities between Wales and the average across OECD countries. First, pupils in Wales (66 per cent) typically report being given the same opportunities to explain their ideas in science lessons as pupils across the OECD (69 per cent). Similar findings emerge for the statements regarding the opportunity to draw conclusions from an experiment (45 per cent in Wales versus 42 per cent OECD average), teachers explaining how an idea from science can be applied to a range of phenomena (56 per cent versus 59 per cent), and whether pupils are asked to conduct investigations to test an idea (28 per cent versus 26 per cent). It therefore seems that pupils in Wales report having similar experiences of

[84] Department for Children, Education, Lifelong Learning and Skills (2008: 11)

linking data to theory and drawing conclusions as in classrooms within many industrialised countries, at least in these particular ways.

Table 10.3 Percentage of pupils who report the use of different activities and teaching practices within school science classes

	Wales	OECD	H10
Pupils are given opportunities to explain their ideas	66%	69%	**63%***
Pupils spend time in the laboratory doing practical experiments	16%	**21%***	17%
Pupils are required to argue about science questions	18%	**30%***	**21%***
Pupils are asked to draw conclusions from an experiment they have conducted	45%	**42%***	**35%***
The teacher explains how a school science idea can be applied to a number of different phenomena	56%	**59%***	**53%***
Pupils are allowed to design their own experiments	12%	**16%***	13%
There is a class debate about investigations	17%	**26%***	17%
The teacher clearly explains the relevance of broad science concepts to our lives	45%	**50%***	**47%***
Pupils are asked to do an investigation to test ideas	28%	**26%***	**19%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils who reported that the corresponding activity or practice happens in ‘every’ or in ‘most’ of their science lessons. Bold font and * denotes statistically different from Wales at the five per cent significance level.

18. There are also some pronounced, statistically significant differences between science classrooms in Wales and the average across OECD members. Pupils in Wales report being less likely to argue about science questions (18 per cent in Wales versus 30 per cent OECD average) and less likely to debate about science investigations (17 per cent versus 26 per cent). Both of these activities involve applying reasoning to scientific fact and constructing arguments. This therefore suggests that there may be less of an atmosphere of debate in Welsh science classrooms relative to the average across OECD countries, even though pupils in Wales generally report having regular opportunities to explain their ideas.

19. We have also investigated whether pupils' experiences of learning science in school varies within Wales, according to school type. Welsh medium school pupils report being more likely to design their own experiments (18 per cent) than their peers at English medium schools (10 per cent). They also report being significantly more likely to have a class debate (22 per cent in Welsh medium schools versus 16 per cent in English medium). Together, this suggests that pupils in Welsh medium schools may have more opportunity to lead their own learning, at least in these particular respects.

Key point

Pupils in Wales report feeling that they have similar opportunities to express themselves and draw conclusions from experiments as their peers in OECD countries, but report spending less time constructing arguments and engaging in debate.

10.3 Is low-level disruption in science classrooms a more common occurrence in Wales than in other countries?

20. Low-level disruption and its 'wearing down effect' are thought to be problems in Welsh schools[85]. This is important as the school learning environment is linked to pupils' attainment, with evidence suggesting that interventions which aim to improve pupil behaviour can also lead to increases in academic achievement[86]. The PISA background questionnaire allows us to consider the frequency of low-level disruption within school science lessons in Wales, and how this compares to other countries.

21. The results for Wales in Table 10.4 show that low-level disruption is reported as a problem in most or in every lesson for approximately 40 per cent of pupils. Forty one per cent of pupils reported that their peers do not listen to their teacher and almost half of pupils reported that there is noise and disorder in 'every' or in 'most' science lessons. This is in contrast to pupils from across the H10 countries, where less than a quarter of pupils report that there is noise and disorder. However, there is an important variation that occurs even within the H10 countries. For instance, issues such as 'noise and disorder' are a lot less common in the high-performing East Asian countries (e.g. 11 per cent in Japan, 20 per cent in China) than in high-performing Western countries, with the situation in Canada (36 per cent) and Finland (38 per cent) more similar to the situation in Wales (45 per cent). Nevertheless, Table 10.4 indicates that pupils in Wales experience more frequent occurrences of

[85] Cole (2007)
[86] EEF (2016)

low-level disruption during their science lessons than pupils report in many other countries (including many of those with high average PISA science scores).

Table 10.4 Percentage of pupils who report low-level disruption occurring frequently during their school science classes

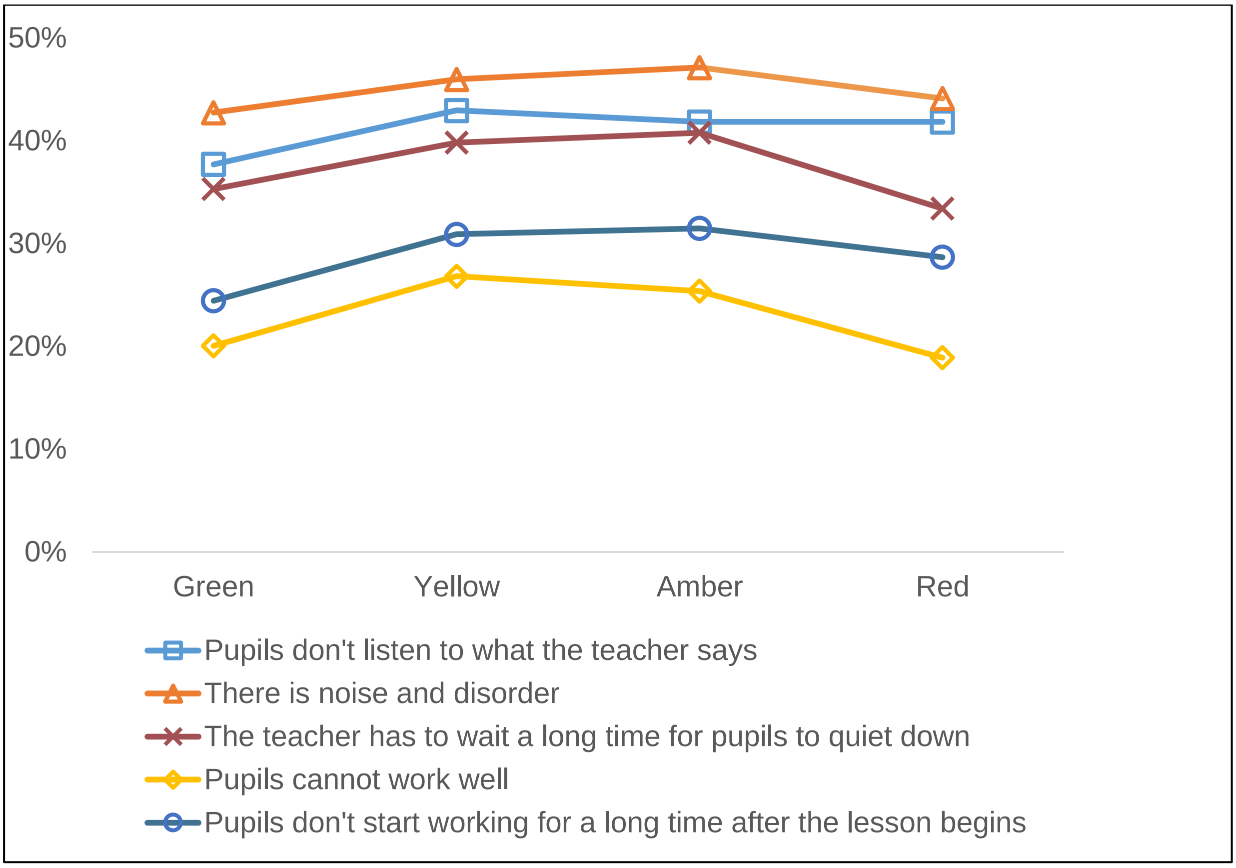
	Wales	OECD	H10
Pupils don't listen to what the teacher says	41%	**32%***	**21%***
There is noise and disorder	45%	**33%***	**22%***
The teacher has to wait a long time for pupils to quiet down	38%	**29%***	**18%***
Pupils cannot work well	24%	**22%***	**15%***
Pupils don’t start working for a long time after the lesson begins	29%	**26%***	**17%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils who reported that this form of disruption occurred in ‘every’ or in ‘most’ of their school science lessons. Bold font and * denotes statistically different from Wales at the five per cent significance level.

22. We have also examined how low-level disruption varies between English/Welsh medium schools and by school support category. There is no evidence of low-level disruption being any more of a problem in either English medium or Welsh medium schools. Moreover, as Figure 10.3 illustrates, the prevalence of low-level disruption does not vary to a great extent between the four support category groups. This is perhaps a surprising result, given that inspectors take into account classroom behaviour and low-level disruption in the categorisation of schools. However, as the sample size is relatively small for schools in the red support category (12 schools and 271 pupils), estimates for this particular group should be interpreted with caution.

Figure 10.3 Percentage of pupils who report low-level disruption in the science classroom by school support category



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils who reported that this form of disruption occurred in 'every' or in 'most' of their school science lessons. The sample size is particularly small for the red school support category group (12 schools and 271 pupils), as such, estimates for this category should be interpreted with caution.

Key point

Pupils report low-level disruption occurring more frequently in Welsh science classrooms than do pupils in many other OECD and high-performing countries.

10.4 How do pupils in Wales perceive the feedback they receive from their science teachers?

23. An important part of a teacher's role is to evaluate the strengths and weaknesses of their pupils, and provide feedback as to how they might improve. Indeed, there is evidence that pupils who receive regular, constructive feedback from their teachers perform better at school[87]. Statistics published in the Estyn 2014-15 annual report showed that in 50 per cent of Welsh secondary schools, teachers give useful written and oral feedback to pupils[88]. It was also found, however, that marking '*is often superficial and does not give enough guidance to pupils on how to improve*'[89]. This was based upon information collected by school inspectors, which raises the question, do pupils in Wales feel the same way? Is there any evidence that science teachers provide more feedback to certain types of pupil (e.g. those with the weakest skills)? Moreover, how does Wales compare to other countries in terms of pupils' perceptions of the feedback they receive from their science teachers?

Table 10.5 Percentage of pupils who receive feedback from their teachers

	Wales	OECD	H10
The teacher tells me how I am performing in this course	33%	**28%***	**26%***
The teacher gives me feedback on my strengths in this school science subject	36%	**25%***	**26%***
The teacher tells me in which areas I can still improve	40%	**30%***	**30%***
The teacher tells me how I can improve my performance	39%	**32%***	**35%***
The teacher advises me on how to reach my learning goals	38%	**32%***	**36%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils who reported that the corresponding activity or practice happens in 'every' or in 'most' science lessons. Bold font and * denotes statistically different from Wales at the five per cent significance level.

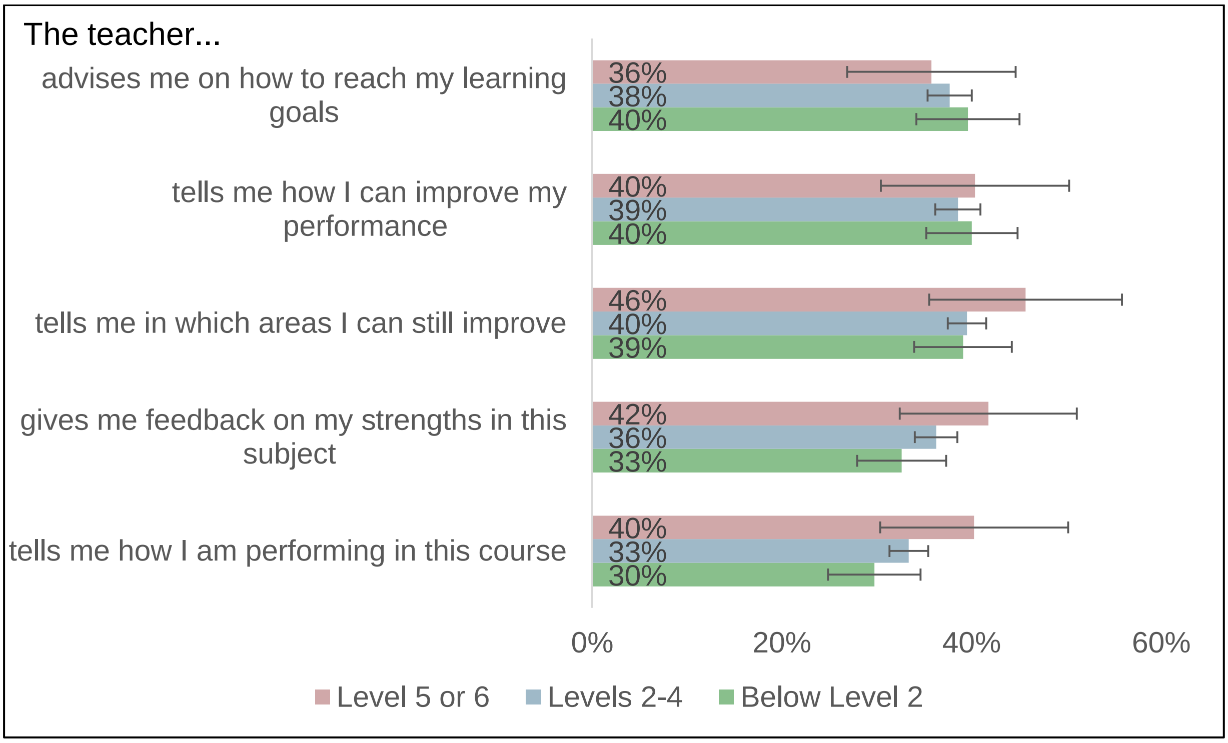
24. Table 10.5 starts to answer some of these questions by illustrating the percentage of pupils who report that they are given various different types of feedback in 'every' or in 'most' lessons (as opposed to in 'some lessons' or 'never'). For each of the five statements, around one third of pupils in Wales report receiving regular feedback. For the second and third statements, the OECD and H10 averages are at least 10 percentage points below the value for Wales. For instance, pupils in Wales are more likely to say that their science teacher advises them on their areas of strength (36 per cent versus 25 per cent) and tells them where they might improve

[87] See Airasian (2000) for an overview of the literature on assessment, feedback and pupil performance.
[88] Estyn (2016).
[89] Estyn (2016: 53).

(40 per cent versus 30 per cent). Between the H10 countries there are also differences. Pupils in Canada report similar levels of feedback to young people in Wales. On the other hand, 15-year-olds in Finland and Japan report much less regular feedback (e.g. less than 20 per cent said they receive feedback on their strengths in 'every' or in 'most' science lessons).

25. Do pupils with low-level science skills receive the most input from their teachers about how they can improve? Or do teachers tend to provide more feedback to average or higher performing pupils? Figure 10.4 provides the results. For three of the five statements, a greater proportion of pupils with high-level science skills report receiving more feedback from their science teacher than pupils with low-level science skills. However, differences between the low-achievers and top-performers are only statistically significant on one occasion, for the statement *'the teacher tells me how I am performing in this course'* (40 per cent versus 30 per cent). From this we draw the conclusion that science teachers in Wales do not give feedback any differently to high or low performing pupils.

Figure 10.4 Percentage of pupils who receive regular feedback from their teachers by science proficiency level



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils who reported receiving the feedback in 'most' or in 'every' science lesson. 'Level' refers to PISA science proficiency level. Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval.

26. Does pupil perception of teacher feedback differ by gender or socio-economic status? Boys in Wales are seven to nine percentage points more likely than girls to report that they receive each type of feedback, with these differences statistically significant at the five per cent level. The same pattern also emerges for the average across OECD members and the average across high-performing PISA countries. This finding could be driven by (a) boys perceiving the level of feedback they receive to be more frequent and/or (b) actual differences in how regularly teachers provide feedback to girls or boys. Unfortunately, the data available within the PISA background questionnaire are not sufficiently detailed to allow us to disentangle these two potential explanations.

27. Finally, there is some evidence that the type and regularity of the feedback pupils receive differs depending upon their socio-economic background. The pupils in Wales from socio-economically disadvantaged backgrounds say they are six to 13 percentage points less likely to receive feedback on each statement than their peers from advantaged backgrounds. The most pronounced difference is in terms of feedback on how they are performing in their science course (27 per cent) as compared to their peers from the most advantaged backgrounds (40 per cent).

Key point

Science teachers in Wales are reported to provide more regular feedback to pupils on their strengths and weakness, including specific areas they can improve, than teachers in many of the countries with the highest average PISA scores.

10.5 Do pupils in Wales feel that they receive regular support from their teachers during their science classes?

28. Pupils spend a considerable amount of time in the classroom, interacting with their peers and their teachers. Yet how exactly do teachers influence their pupils' learning outcomes? Previous research on this matter has been somewhat mixed, and unable to directly identify measures of teacher 'quality'[90]. However, one channel that has not been fully explored is the support that teachers provide to pupils during their time in class. To conclude this chapter, we therefore investigate how 15-year-olds in Wales interact with their science teachers. This includes whether pupils in Wales believe that their science teacher is supportive, and is able to adapt their lesson to meet the needs of those that they teach.

[90] See Hanushek and Rivkin (2010) for further discussion on the 'teacher value-added' literature and existing evidence.

29. Table 10.6 begins by exploring the extent to which a series of classroom practices (e.g. whether whole class discussion takes place) are reported as used in 'every' or in 'most' science lessons. These classroom practices are used to support learning and focus on explanation, demonstration and discussion. More than half of pupils in Wales report that their science teacher regularly explains scientific ideas (57 per cent), demonstrates an idea (53 per cent) and discusses pupils' questions (51 per cent). On the other hand, whole class discussions occur somewhat less frequently; 35 per cent of pupils in Wales report that they take place in most or every lesson. This result is consistent with pupils' reports of infrequent classroom debates (see sub-section 10.2).

Table 10.6 The extent to which teachers use different classroom practices

	Wales	OECD	H10
The teacher explains scientific ideas	57%	55%	59%
A whole class discussion takes place with the teacher	35%	**40%***	**41%***
The teacher discusses our questions	51%	**55%***	**54%***
The teacher demonstrates an idea	53%	54%	**57%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who said this happened in 'every' or in 'most' of their science lessons. Bold font and * denotes statistically different from Wales at the five per cent significance level.

30. There are relatively few substantial points of difference between the results for Wales and the OECD / H10 averages. For instance, 35 per cent of pupils in Wales report whole classroom discussion regularly taking place, compared to an H10 average of 41 per cent. Yet there is a bigger difference between Wales and other Western countries with high average PISA scores, such as Canada (51 per cent), Estonia (49 per cent) and Finland (46 per cent). Nevertheless, on the whole, pupils' perception of their science teacher's use of supportive classroom practices is similar in Wales to many other countries.

31. Table 10.7 presents further evidence as to whether pupils in Wales believe that their science teacher is supportive. Here pupils were asked to state how often their teacher engaged in supportive classroom practices, including providing help,

showing interest and making sure all pupils understand the subject matter. Again, there is little substantial difference between Wales and the OECD/H10 averages. One notable exception is that pupils in Wales are 10 percentage points more likely to say that their teachers *'help pupils with their learning'* than in the average OECD country (81 per cent versus 71 per cent). Welsh pupils are also five percentage points more likely to say that their teachers *'give extra help'* than the average across OECD countries (78 per cent versus 73 per cent). However, despite these exceptions, the overall indication from Table 10.7 is that Wales does not typically stand out from the average OECD or average high-performing country in the amount of support science teachers provide to their pupils.

Table 10.7 Percentage of pupils who perceive their teachers as supportive

	Wales	OECD	H10
The teacher shows an interest in every pupil's learning	70%	69%	72%
The teacher gives extra help when pupils need it	78%	**73%***	79%
The teacher helps pupils with their learning	81%	**71%***	80%
The teacher continues teaching until the pupils understand	72%	**69%***	72%
The teacher gives pupils an opportunity to express opinions	60%	**68%***	**72%***

Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who said this happened in 'every' or in 'most' of their science lessons. Bold font and * denotes statistically different from Wales at the five per cent significance level.

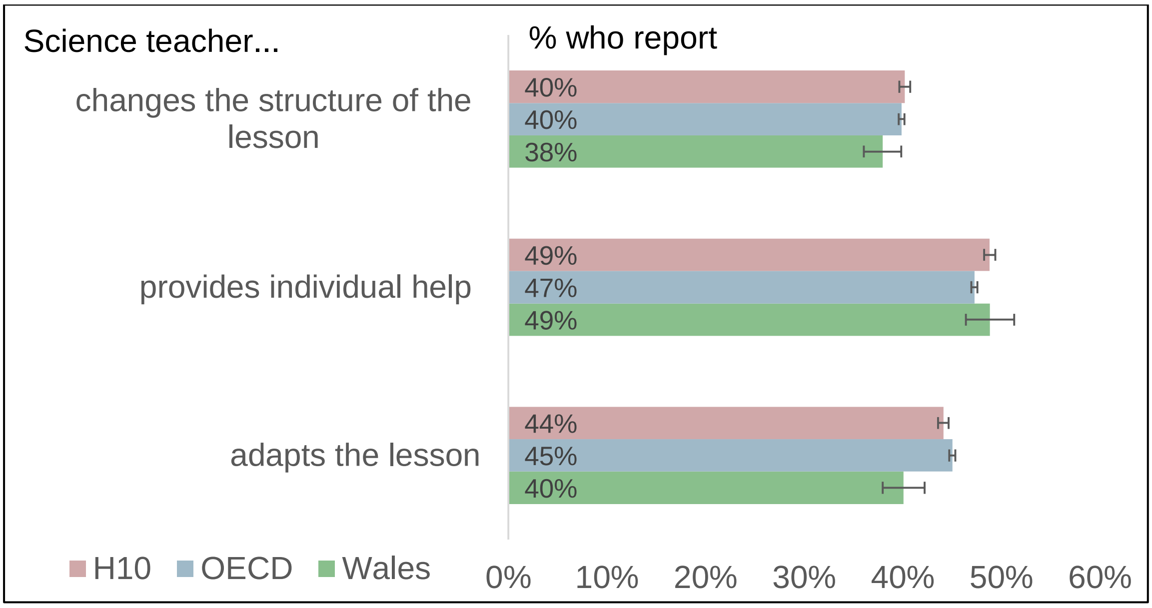
32. In order to better support their pupils, teachers may adapt their approach in the classroom depending upon the needs of those that they teach. Within the background questionnaire, pupils were asked whether they felt their science teacher did indeed adapt their lessons when needed. They were asked to say how frequently the following occurred:

- *The teacher changes the structure of the lesson on a topic that most students find difficult to understand*
- *The teacher provides individual help when a student has difficulties understanding a topic or task*
- *The teacher adapts the lesson to my class's needs and knowledge*

33. Figure 10.5 indicates that pupils in Wales are not significantly more likely to report that their science teacher adapts their lessons depending upon pupils' needs

than the average across OECD and H10 countries. Science teachers in Wales are reported as less likely to adapt their lessons to the class's knowledge and needs (40 percent in Wales versus H10 average of 44 percent and OECD average of 45 percent).

Figure 10.5 Pupils' perception of teachers' ability to adapt



Source: PISA 2015 database

Notes: Figures refer to the percentage of pupils in schools who said this happened in 'every' or 'most' of their science lessons. Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval. Data for Vietnam missing and therefore not included in the H10 average.

34. Do pupils' views of whether their science teacher is able to adapt their lessons vary depending upon their background characteristics? In additional analysis, we have found little evidence that pupils' responses to the questions above differ substantially by gender or the categorisation of the school that they attend. However, there are some striking differences between higher and lower achieving pupils, as measured by their science proficiency level. Sixty eight per cent of high achieving pupils (scoring at Level 5 or 6) report that their science teacher provides individual help during most lessons. This is over 30 percentage points higher than pupils who obtain PISA test scores below Level 2 (35 per cent). However, this finding is not unique to Wales; a similar difference also arises in England, for example. Welsh pupils who lack basic science skills are also much less likely to agree that their teachers '*adapt[ed] the lesson to [their] class's needs and knowledge'* (30 per cent) relative to pupils with high-level skills (62 per cent), while there is also an 18 percentage point difference in pupils' views of how willing their science teacher is to change the structure of the lesson on a challenging topic (32 per cent for low proficiency pupils versus 50 per cent for high proficiency pupils). Overall, these results may indicate that low achieving pupils in Wales feel left behind during some

of their science lessons, and do not perceive their science teachers as able to adapt to their needs.

Key point

Pupils in Wales generally perceive their science teacher to be supportive. However, low achieving pupils' report that their science teacher is less willing to provide individual help and adapt their lessons than their high achieving peers.

Chapter 11. PISA in the UK

- The average PISA science score is highest in England (512) and lowest in Wales (485). Scotland (497) and Northern Ireland (500) fall in-between.

- Differences in average PISA mathematics scores between England (493), Northern Ireland (493) and Scotland (491) are not statistically significant. On the other hand, the average PISA mathematics score is significantly lower in Wales (478) than the rest of the UK.

- There is no statistically significant difference in average PISA reading scores across England (500), Northern Ireland (497) and Scotland (493). However, the average reading score is significantly lower in Wales (477) than the rest of the UK.

- There has been a sustained decline in average PISA science scores in Wales, from 505 points in 2006 to 485 points in 2015.

- Since 2006, the science skills of the highest achieving pupils in Northern Ireland, Scotland and Wales have steadily declined.

- Around one-in-four pupils in the UK lack basic skills in mathematics. Moreover, around one-in-five lack basic skills in science and reading.

- The comparatively low reading skills of girls stands out as a particular challenge facing Wales.

- Headteachers' views on the factors hindering instruction within their school are generally similar across the UK. However, a lack of well-qualified teaching staff stands out as a particular concern amongst headteachers in England.

- Across the UK, 15-year-olds report spending more time studying science than English and mathematics. Scottish, Welsh and Northern Irish pupils report spending over an hour more time studying outside of school per week (on average) than their English peers.

1. The United Kingdom is a prime example of how school systems and education policies can vary markedly within a country. For instance, although comprehensive, mixed ability schools are common in England, Wales and Scotland, this is not the case in Northern Ireland, where almost half of 15-year-olds are taught in grammar schools. On the other hand, England takes a somewhat different approach to accountability than the rest of the UK, through its annual publication of school 'league tables'. Other more recent policy developments, such as the academies programme, are specific to England and have not been introduced elsewhere. These are just a handful of examples of how education policy and provision varies significantly across England, Northern Ireland, Scotland and Wales.

2. At the same time, many of the issues that complicate international comparisons are (arguably) less of a concern when looking across the four constituent countries of the UK. There are, for instance, important similarities in terms of culture, language, economic development and political systems, as well as a shared history. Although some of these factors (e.g. culture) may help to explain differences in achievement between the UK and other parts of the world (e.g. Asia), it is arguably less likely that they will explain differences between England, Northern Ireland, Scotland and Wales.

3. As noted by Taylor, Rees and Davies (2013), within-UK comparisons are therefore interesting from both an academic and education policy perspective. Yet, due to a lack of accessible and comparable national examination data, relatively few 'home international' comparisons have been conducted[91]. PISA is an important exception. By drawing separate samples for England, Northern Ireland, Scotland and Wales, PISA provides a three-yearly update of how academic achievement, pupils' attitudes and headteachers' concerns vary across different parts of the UK.

4. In this concluding chapter, we therefore focus upon differences in PISA test scores and background questionnaire responses across these four countries. The following research questions will be addressed:

- *How do average PISA test scores compare across the UK?*
- *What proportion of 15-year-olds in the UK do not have basic science, mathematics and reading skills?*
- *How have average PISA scores changed across the UK since 2006?*

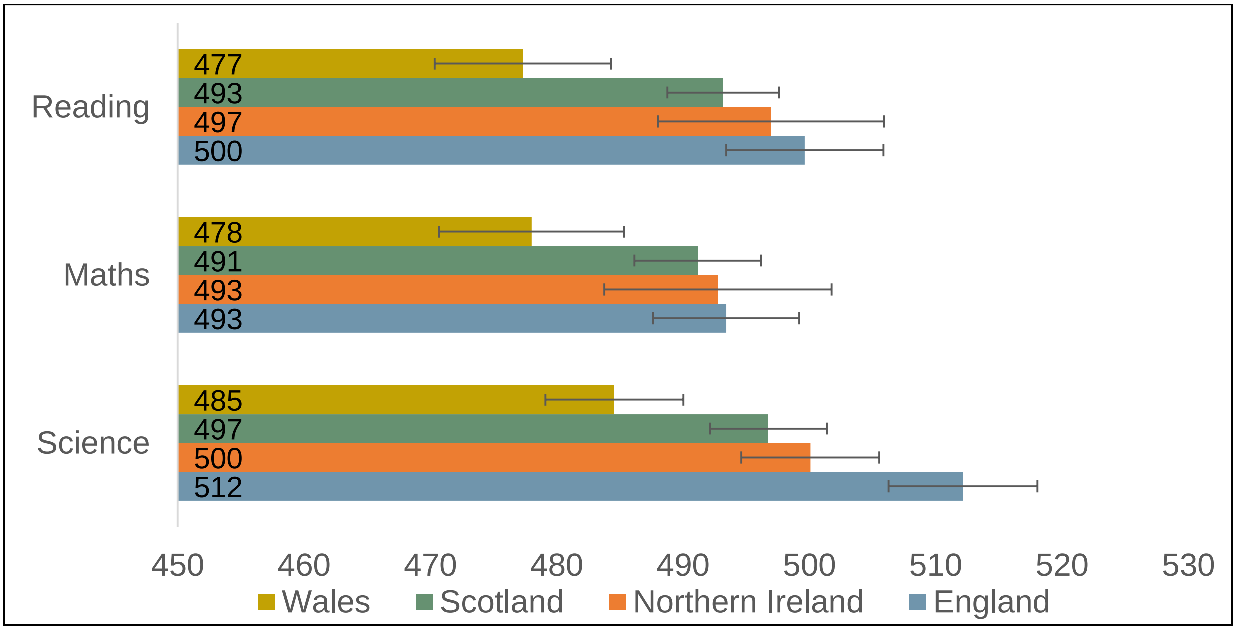
[91] Though see Taylor, Rees and Davies (2013).

- *How has the performance of the highest and lowest achieving pupils changed across the UK since 2006?*
- *Are gender gaps in achievement bigger in some parts of the UK than others?*
- *How does the relationship between socio-economic status and achievement vary across the UK?*
- *Do headteachers' views on the factors hindering instruction within their school differ across the UK?*
- *Are there differences in the amount of instruction 15-year-olds receive – both inside and outside of school?*

11.1 How do average PISA test scores compare across the UK?

5. Do 15-year-olds in certain parts of the UK achieve higher average PISA science scores than others? The answer can be found in Figure 11.1. Average science scores are highest in England (512) and lowest in Wales (485). These two countries are significantly different to both Northern Ireland (500) and Scotland (497) at the five per cent level. There is hence a clear hierarchy across the UK, with the strongest average science performance in England, the weakest in Wales, with Northern Ireland and Scotland sitting in-between.

Figure 11.1 Average PISA test scores across the UK



Source: PISA 2015 database.

Note: Thin black line running through centre of bars refers to the estimated 95 per cent confidence interval.

6. There is less variation in average scores across the UK in the PISA mathematics domain (see the middle set of bars in Figure 11.1). For instance, England (493), Northern Ireland (493) and Scotland (491) are separated by just two test points, and are statistically indistinguishable at the five per cent significance level. In contrast, the average mathematics score in Wales is 478. This is significantly lower than the mean score for the other three countries within the UK, with a difference of around 15 test points (around half a year of additional schooling).

7. Finally, the uppermost set of bars in Figure 11.1 turns to average PISA reading scores. There is little evidence of variation across England (500), Northern Ireland (497) and Scotland (493), with all cross-country differences statistically insignificant at the five per cent level. However, the mean score is again significantly lower in Wales (477).

Table 11.1 Average PISA test scores across the science sub-domains within the UK

Domain	England	Northern Ireland	Scotland	Wales
Scientific systems				
Physical	512	501	499	486
Living	512	498	497	482
Earth and Space	513	498	494	485
Scientific competencies				
Explain phenomena scientifically	512	500	498	486
Evaluate and design scientific enquiry	510	497	498	481
Interpret data and evidence scientifically	512	501	493	483
Knowledge				
Content knowledge	511	499	496	486
Procedural and epistemic knowledge	513	501	496	484
Points difference from England				
0 to 5 points				
5 to 10				
10 to 15				
15 to 20				
20 to 25				
25 or more				

Source: PISA 2015 database

8. As science was the focus of PISA 2015, we are also able to consider how achievement in this subject varies across the science sub-domains. For instance, are the comparatively high science scores of English pupils driven by a particular strength in one specific aspect of scientific literacy? Or do English pupils achieve higher science test scores than the rest of the UK across the board? Table 11.1 provides the results. In this table, darker shading refers to greater distances from the average score in England.

9. The pattern of achievement across the various science sub-domains is reasonably similar across England, Northern Ireland, Scotland and Wales; the similarities across the UK in Table 11.1 are more striking than the differences. For instance, in all four countries, scores in the living scientific system are quite similar to those in the physical and earth and space science systems. Likewise, pupils from England, Northern Ireland and Wales are no stronger (or weaker) at 'interpreting data and evidence scientifically' than at 'evaluating and designing scientific enquiry'. Finally, in all four countries, average scores for 'content knowledge' are similar to the scores for 'procedural and epistemic knowledge', with a difference of less than five points.

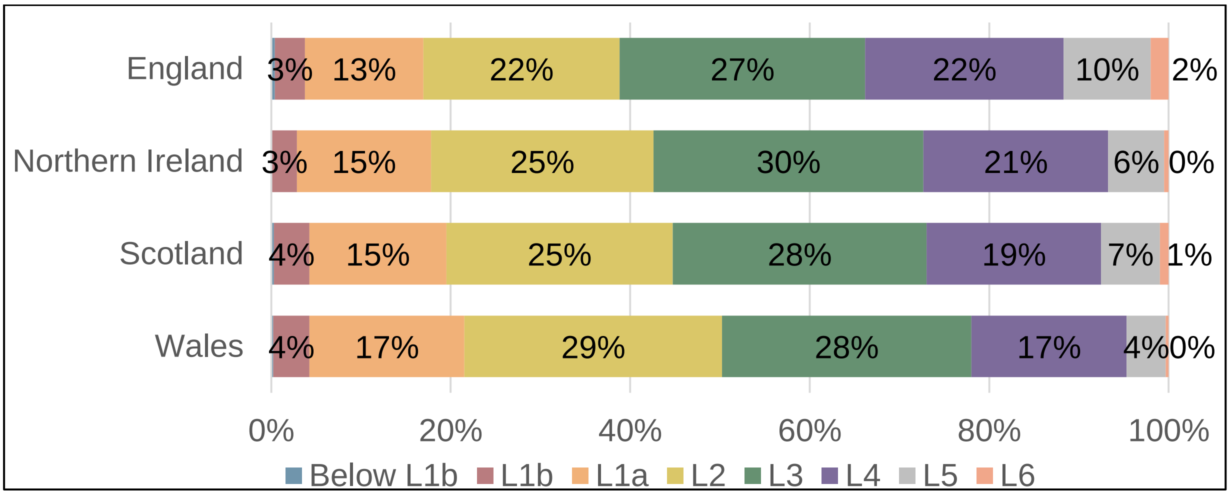
Key point

The average PISA science score is significantly higher in England than Scotland, Northern Ireland and Wales. In all three core PISA subjects, Wales has statistically significantly lower average scores than the rest of the UK.

11.2 What proportion of 15-year-olds across the UK do not have basic science, mathematics and reading skills?

10. Although average PISA test scores may be similar across most of the UK, does the same hold true for the distribution of 15-year-olds across the PISA proficiency levels? In particular, do certain parts of the UK have a greater proportion of 'low-achievers'; 15-year-olds who have not reached the OECD's baseline level of achievement? Figure 11.2 provides the answer for science. Wales has the greatest proportion of 15-year-olds operating below Level 2 (22 per cent), followed by Scotland (20 per cent), Northern Ireland (18 per cent) and England (17 per cent).Together this means that around one-in-five young people from across the United Kingdom do not have basic science skills. In terms of 'top-performers', England has the greatest proportion of young people working at PISA Levels 5 and 6 (12 per cent), compared to eight per cent in Scotland, seven per cent in Northern Ireland and five per cent in Wales.

Figure 11.2 The proportion of UK pupils reaching each PISA science level

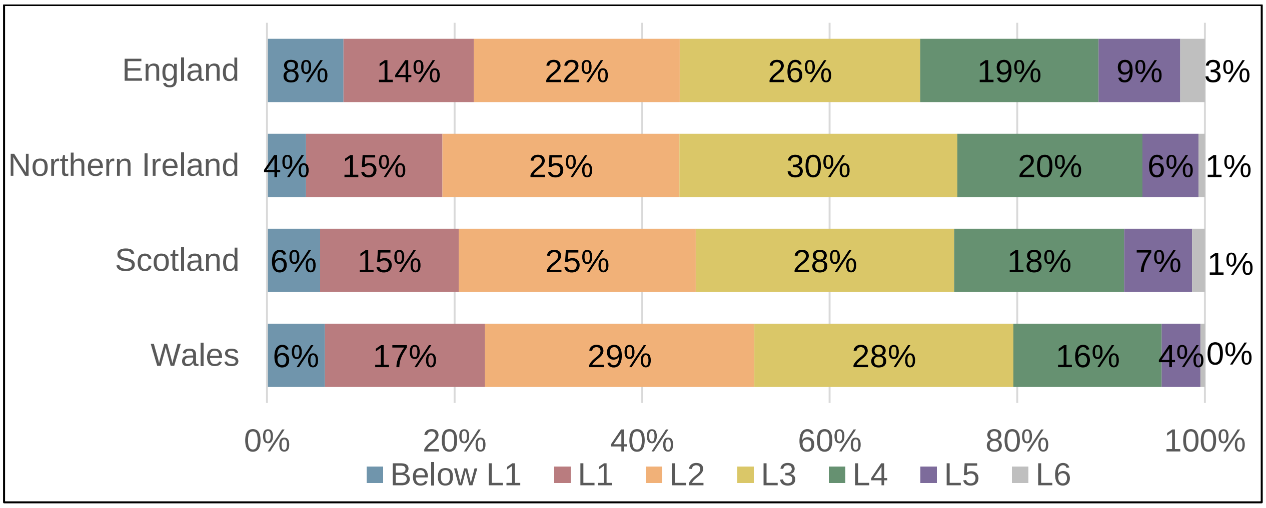


Source: PISA 2015 database.

11. Results for PISA mathematics are provided in Figure 11.3. Within the UK, England (22 per cent) and Wales (23 per cent) have the greatest proportion of low-achievers in this subject while Northern Ireland has the least (19 per cent). Consequently, across the United Kingdom as a whole, between a fifth and a quarter of 15-year-olds do not have basic proficiency in mathematics.

12. At the other extreme, Wales also has fewer 15-year-olds reaching the highest mathematics proficiency levels than the rest of the UK. Specifically, just five per cent of Welsh pupils obtain a PISA mathematics score at Level 5 or 6, compared to 11 per cent of pupils in England, nine per cent in Scotland and seven per cent in Northern Ireland. Overall, around 11 per cent of pupils across the UK are a ‘top-performer’ in mathematics.

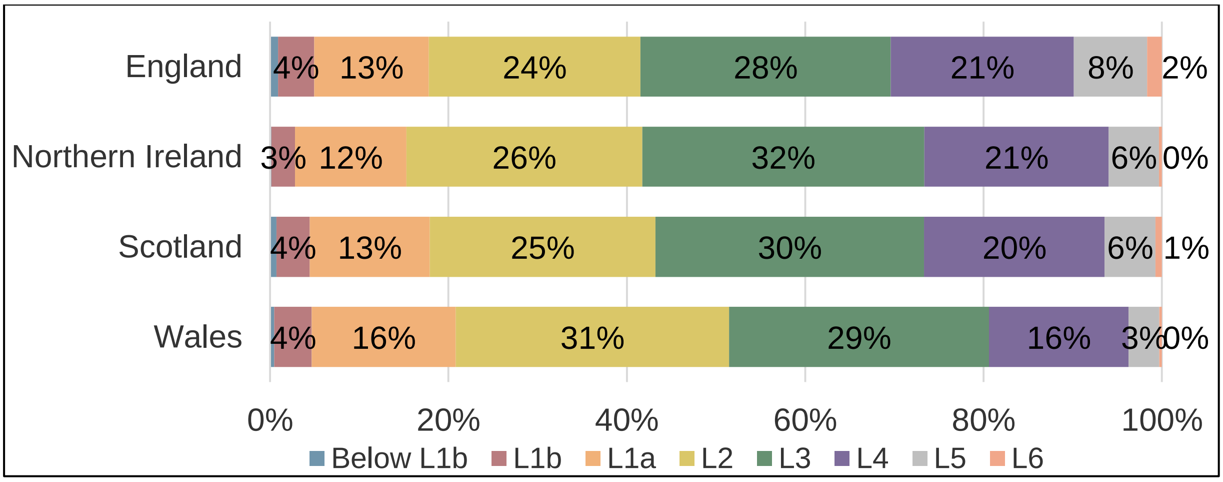
Figure 11.3 The proportion of UK pupils reaching each PISA mathematics level



Source: PISA 2015 database.

13. Finally, Figure 11.4 presents results for the distribution of PISA reading scores. The most notable difference is that Northern Ireland has slightly fewer low-performers than England and Scotland (15 per cent versus 18 per cent in England and Scotland), while England has a slightly greater proportion of the highest achievers (10 per cent versus six per cent in Scotland and Northern Ireland). Wales, on the other hand, has more 15-year-olds who lack basic reading skills (21 per cent achieve below PISA Level 2) and fewer top-performers (four per cent reaching PISA Level 5 or 6) than the rest of the UK.

Figure 11.4 The proportion of UK pupils reaching each PISA reading level



Source: PISA 2015 database.

Key point

Around 29 per cent of pupils in the UK lack basic skills in at least one PISA subject area (science, mathematics and reading). Around 10 per cent of pupils in the UK lack basic skills in all three domains.

11.3 How have average PISA scores changed across the UK since 2006?

14. Chapters 2, 4 and 5 of this report illustrated how average PISA scores in Wales have changed since 2006. Table 11.2 demonstrates how this compares to the trend observed across the rest of the UK. Two particular issues stand out.

15. There is evidence of a sustained decline in average scores during the 2006 to 2015 period for Wales in the science domain (see Table 11.4). In this country, the average science score has gradually fallen from 505 points in 2006 to 485 points in 2015. This represents a fall of 20 test points and is statistically significant at the five per cent level. There is also evidence of a fall in mathematics scores in Scotland since 2006, with the mean falling from 506 in 2006 to 499 in 2009, 498 in 2012 and 491 in 2015. The three-year average trend in Scotland is therefore downwards, and statistically significant at the five per cent level.

Table 11.2 Average PISA scores across the UK from 2006 to 2015

		2006	2009	2012	2015
Science	England	516	515	516	512
	Northern Ireland	508	511	507	500
	Scotland	515	514	513	497
	Wales	505	496	491	485
Mathematics	England	495	493	495	493
	Northern Ireland	494	492	487	493
	Scotland	506	499	498	491
	Wales	484	472	468	478
Reading	England	496	495	500	500
	Northern Ireland	495	499	498	497
	Scotland	499	500	506	493
	Wales	481	476	480	477

Source: PISA 2006 to 2015 databases.

Note: See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

16. The second notable feature of Table 11.2 is that there has been a sharp drop in average science scores in Scotland compared to previous PISA rounds. Specifically, while the mean score for Scotland remained largely unchanged between 2006 (515), 2009 (514) and 2012 (513), it dropped by around 16 test points (around half a year of schooling) in 2015. Although this is a sizeable and statistically significant difference compared to the last time science was the focus of PISA in 2006, some caution is needed when interpreting this result. As noted in chapter 1, a number of changes have been made to the administration of PISA in 2015, particularly within the science domain (e.g. the introduction of computer-based testing, alterations made to the framework and the use of interactive test questions). Furthermore, other countries have previously experienced a 'blip' in average scores in one particular wave of PISA, before quickly recovering in the following round (e.g. mean reading and mathematics scores in the Republic of Ireland dropped sharply between 2006 and 2009 before returning to their previous level in 2012[92]). Evidence from the next round of PISA, due to be conducted in 2018, is therefore needed to provide appropriate context for this result.

Key point

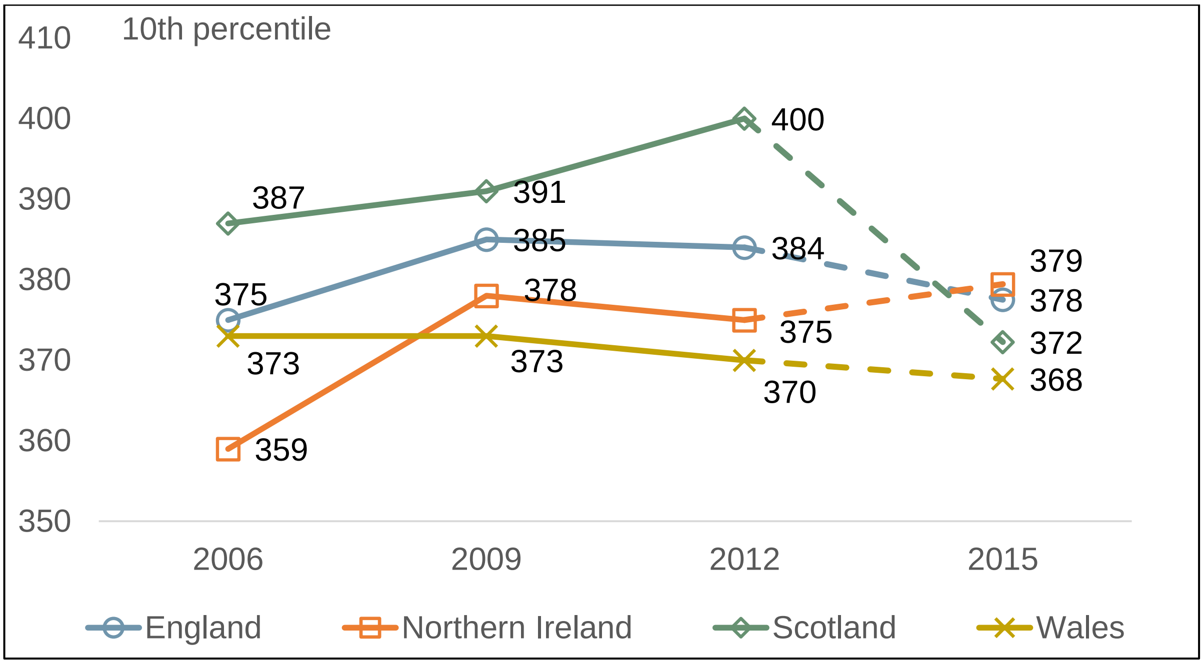
There has been a sustained decline in average PISA science scores in Wales during the last decade.

11.4 How has the performance of the highest and lowest achieving pupils changed across the UK since 2006?

17. The previous sub-section illustrated the change in *average* PISA scores across the UK over the last decade. Now we turn our attention to changes in the *distribution* of achievement over time, paying particular attention to the performance of the highest and lowest achieving pupils. For brevity, our discussion focuses upon science, with results for reading and mathematics provided in the online data tables.

[92] See Cosgrove and Cartwright (2014) for a detailed discussion of the experience of Ireland in 2009.

Figure 11.5 The 10th percentile of the science proficiency distribution between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

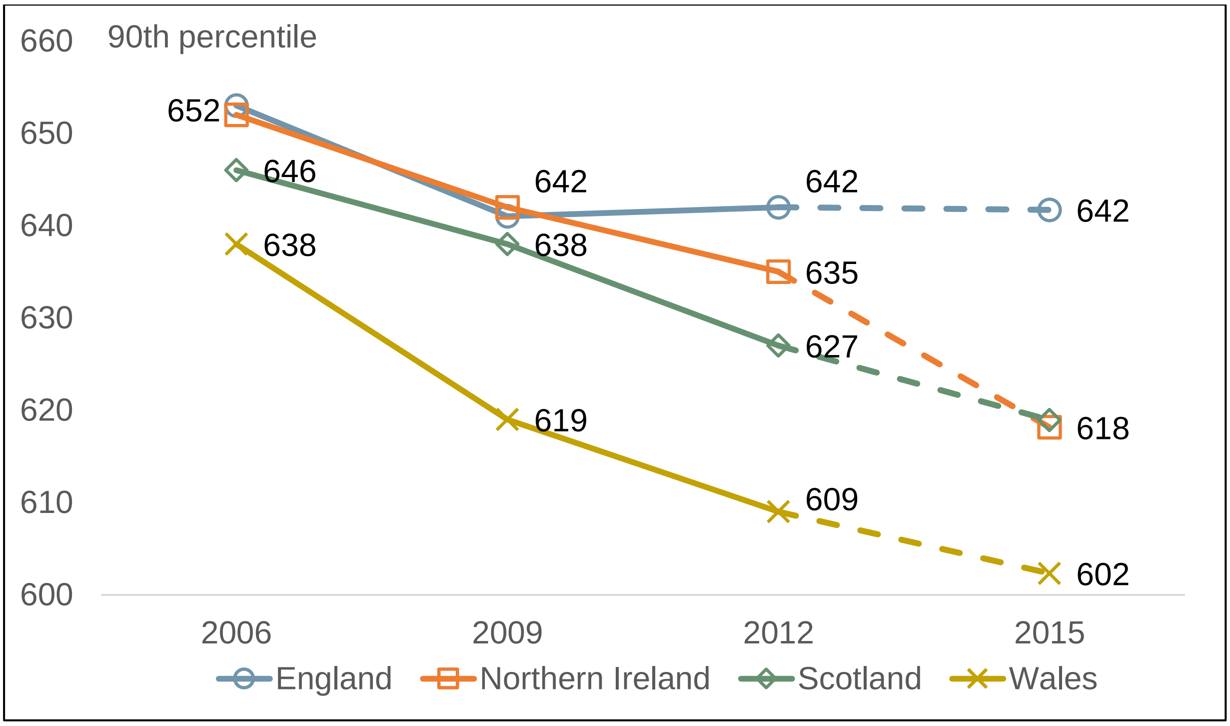
Note: Dashed line refers to the introduction of computer based testing in 2015. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

18. Figure 11.5 illustrates how the 10th percentile of the PISA science distribution has changed between 2006 and 2015. These results therefore refer to the science proficiency of the lowest achieving pupils. There are few clear consistent trends emerging for any part of the UK. Northern Ireland saw a 19 point (eight months of schooling) increase in the 10th percentile between 2006 and 2009, though this has remained at the same level ever since. Scotland, on the other hand, saw the 10th percentile improve from 387 in 2006 to 400 in 2012, before a marked decline to 372 in 2015 (a difference compared to 2012 of almost a year of schooling). Similarly, the performance of the lowest science achievers in Wales remained stable from 2006 to 2012 at around 370 PISA test points, with a slight (statistically insignificant) decline to 368 points in 2015. Meanwhile, the 10th percentile in England has remained broadly around the same level throughout this period. Overall, there seems to have been some sharp one-off movements in the 10th percentile in certain parts of the UK, though little consistent evidence of a sustained upwards or downwards trend.

19. However, the same is not true for change in the 90th percentile of the science achievement distribution, as illustrated in Figure 11.6. In Northern Ireland, Scotland and Wales there is evidence of a sustained decline in performance amongst the highest science achievers. For instance, in 2006 the 90th percentile of the science

distribution in Northern Ireland stood at 652 points. This has gradually fallen to 642 points in 2009, 635 points in 2012 and 618 points in 2015. A similar decline in the 90th percentile has been observed in Scotland (from 646 points in 2006 to 619 points in 2015) and Wales (638 points in 2006 to 602 points in 2015). Consequently, in these three countries, the highest achieving pupils in science in 2015 are around a year of schooling behind the highest achieving pupils who took the PISA test in 2006. Interestingly, the same is not true in England, where there is little evidence of sustained change in the 90th percentile of science achievement over the last decade.

Figure 11.6 The 90th percentile of the science achievement distribution between 2006 and 2015



Sources: Bradshaw et al. (2007), Bradshaw et al. (2010), Wheater et al. (2014), PISA 2015 database.

Note: Dashed line refers to the introduction of computer based testing in 2015. See Appendix F for further information on PISA 2012 scores in England, Wales and Northern Ireland.

20. A couple of additional implications of Figure 11.5 and 11.6 are also worth highlighting. First, there has been a reduction in inequality of science achievement (as measured by the difference between the 90th and 10th percentile) within certain parts of the UK over the last decade. For instance, the gap between the highest and lowest achieving pupils has fallen from 281 points in Northern Ireland in 2006 to 239 points in 2015, and from 267 points to 235 points in Wales. However, this reduction in inequality has been driven less by an increase in the performance of low-achievers, and more by a decline in achievement amongst the top-performing pupils. Second, the sizeable change in mean science scores in Scotland between 2012 and

2015 is mainly due to a decline in performance amongst lower achieving pupils. For instance, whereas the 90th percentile of the science distribtuion declined by eight points between 2012 and 2015, the 10th percentile dropped by around 28 test points. Hence it seems that certain parts of the science achievement distribution in Scotland have changed more in this short period of time than others.

Key point

The science skills of the highest achieving pupils have steadily declined over the last decade in Northern Ireland, Scotland and Wales.

11.5 Are gender gaps in achievement bigger in some parts of the UK than others?

21. Chapter 6 discussed the gender gap in 15-year-olds' PISA scores, and considered how Wales compares to the rest of the world in this respect. In this sub-section, we bring gender differences across the UK into sharper focus. This will provide an insight into whether differences in achievement between the four constituent countries of the UK are being driven by a comparatively strong or weak performance of boys or girls. Table 11.3 provides the results, with panel (a) referring to science, panel (b) to mathematics and panel (c) to reading.

22. There is no statistically significant difference in average PISA science scores between boys and girls in any country within the UK. For both genders, England has the highest average score, Wales the lowest, while Northern Ireland and Scotland fall in-between.

23. Boys achieve a higher average score than girls in the PISA mathematics test across all parts of the UK, though the gender difference only reaches statistical significance at the five per cent level in England and Wales. Nevertheless, the magnitude of the gender gap is similar across all four countries, standing at 12 test points in England, 10 points in Wales and seven points in Scotland and Northern Ireland. Thus, for both mathematics and science, the similarity of the size and direction of the gender gap across the UK is more striking than any difference.

24. Turning to the results for reading (Table 11.3 panel c), average PISA scores for girls are significantly higher than for boys across each of the four constituent countries. However, there is also evidence of some variation within the UK. In particular, the gender gap in reading is around 10 points smaller in Wales (11 point

difference between boys and girls) than England (23 point difference) and Scotland (21 point difference). This is partly the result of the particularly low reading skills of Welsh girls, who achieve an average PISA reading score around the same level as English, Scottish and Northern Irish boys.

Table 11.3 Gender differences in PISA scores across the UK

(a) Science

	Boys	Girls	Difference
England	512	512	0
Northern Ireland	501	499	3
Scotland	497	496	1
Wales	487	482	5

(b) Mathematics

	Boys	Girls	Difference
England	500	487	**12***
Northern Ireland	496	489	7
Scotland	495	488	7
Wales	483	473	**10***

(c) Reading

	Boys	Girls	Difference
England	488	511	**-23***
Northern Ireland	490	504	**-14***
Scotland	483	504	**-21***
Wales	472	483	**-11***

Source: PISA 2015 database.

Notes: Bold font with * indicates difference significantly different from zero at the five per cent level.

Key point

The comparatively low reading skills of girls stands out as a particular challenge facing Wales.

11.6 How does the relationship between socio-economic status and achievement vary across the UK?

25. Chapter 6 introduced two ways of measuring the association between socio-economic status and pupils' academic achievement. These are the 'impact' (how much test scores change per one-unit increase in the PISA Economic, Social and Cultural Status index) and the 'strength' (the amount of variation in PISA test scores explained by pupils' family background). Table 11.4 considers how these two measures of socio-economic inequality in science achievement differ across the UK[93].

Table 11.4 The 'strength' and 'impact' of socio-economic status upon pupils' science test scores

	Impact	Strength
England	38.2	11%
Scotland	36.9	11%
Northern Ireland	36.0	11%
Wales	24.8	6%

Source: PISA 2015 database.

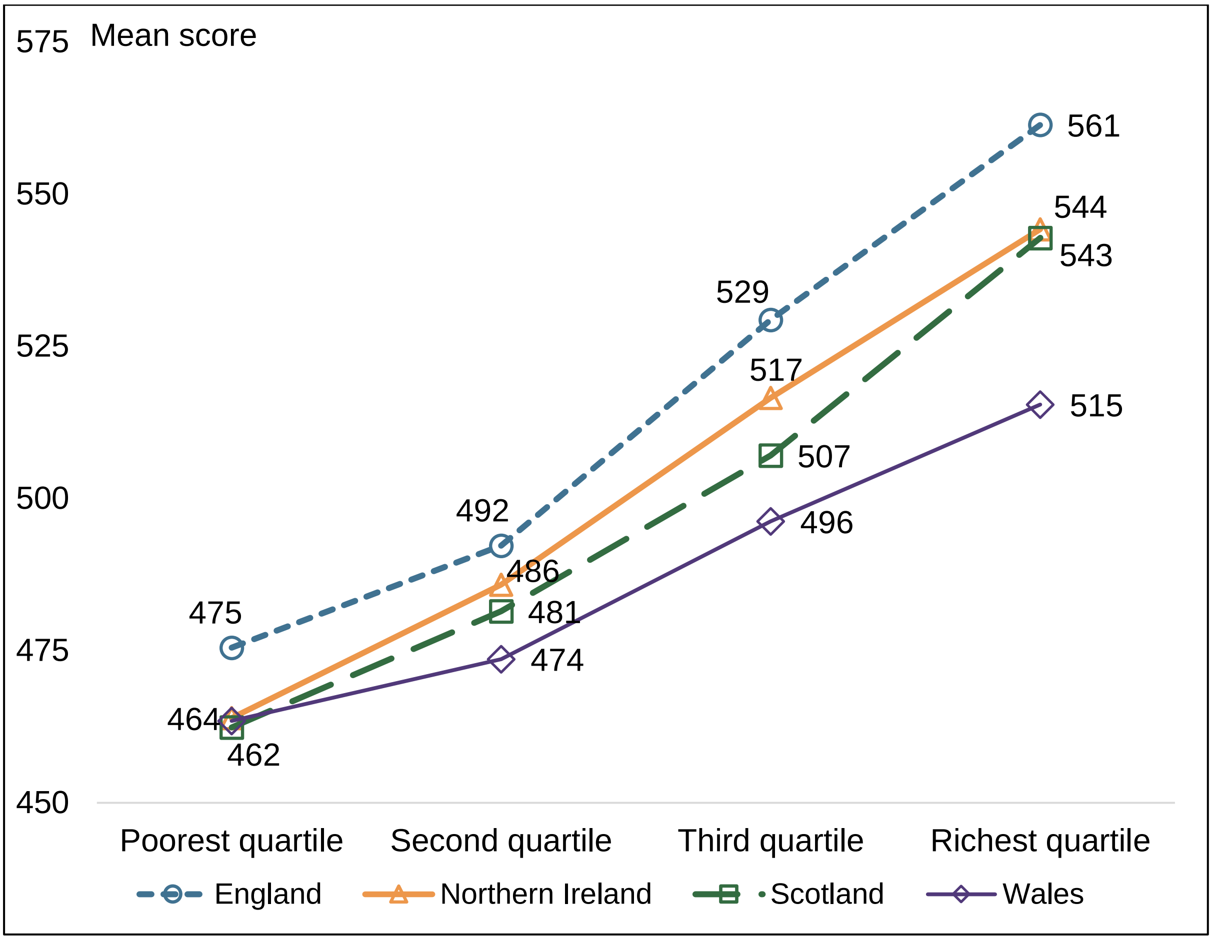
26. There is no evidence that the strength and the impact of socio-economic status varies across England, Scotland and Northern Ireland. In all three countries, a one-unit change in the ESCS index is associated with around a 35 to 40 test point increase in PISA science scores, with approximately 11 per cent of the variance in pupils' achievement explained. On the other hand, both measures are notably lower in Wales, where a one-unit increase in ESCS is associated with a 25 test point increase in PISA science scores. Moreover, in Wales socio-economic status explains only around six per cent of the variation in pupils' science scores; around half the amount that is explained in England, Northern Ireland and Scotland. Both measures suggest that socio-economic inequality in 15-year-olds' science achievement is greater in England, Scotland and Northern Ireland than in Wales. A similar, though slightly less pronounced, result holds for mathematics and reading as well (see online data tables for further details).

27. A deeper exploration of this issue is provided in Figure 11.7. Here we have divided 15-year-olds in each country into four equal groups (quartiles) based upon their ESCS index score. Average PISA science scores are then plotted along the vertical axis, with socio-economic status quartiles running along the horizontal axis.

[93] Results for mathematics and reading and provided in the online data tables.

28. A striking feature of Figure 11.7 is that differences across the four countries are much more pronounced for pupils from advantaged socio-economic backgrounds (‘richest quartile’) than for the least advantaged socio-economic group (‘poorest quartile’). For instance, socio-economically disadvantaged pupils in Northern Ireland, Scotland and Wales achieve roughly the same average science score (around 460) with those in England slightly ahead (475). Hence the four UK nations differ by around 10 to 15 test points. Yet, for the most advantaged socio-economic group, differences across the four UK countries are a lot more apparent. For instance, the average score for the top socio-economic quartile in England is around 15 points higher than in Northern Ireland and Scotland, and around 45 points higher than in Wales. Together, this suggests that England’s comparatively high mean science score relative to the rest of the UK (recall Figure 11.1) is to a certain extent being driven by the strong performance of young people from more advantaged socio-economic backgrounds. Similarly, the comparatively weak science skills of high socio-economic status pupils in Wales is a key reason why the mean score for this country lags behind the rest of the UK.

Figure 11.7 The relationship between socio-economic status quartile and average PISA science scores across the UK



Source: PISA 2015 database.

Notes: Socio-economic groups refer to quartiles of the ESCS across the UK.

Key point

There is a weaker association between socio-economic status and PISA science scores in Wales than the rest of the UK. This is driven by the most advantaged Welsh pupils not achieving as highly as their English, Scottish and Northern Irish peers.

11.7 How do headteachers' views on the factors hindering instruction differ across the UK?

29. Chapter 8 examined headteachers' views of whether their school is adequately resourced. In Table 11.5 we review their responses, and consider how Wales compares to the rest of the UK.

Table 11.5 Headteachers' reports of the resources that are lacking within their school: comparison across the UK

	England	Northern Ireland	Scotland	Wales
A lack of teaching staff	45%	27%	45%	20%
Inadequate or poorly qualified teachers	22%	4%	8%	15%
A lack of assisting staff	18%	21%	32%	19%
Inadequate or poorly qualified assisting staff	12%	5%	10%	13%
A lack of educational material	29%	26%	31%	31%
Inadequate or poor quality educational material	26%	23%	26%	28%
A lack of physical infrastructure	48%	45%	24%	44%
Inadequate or poor quality physical infrastructure	45%	45%	24%	48%

Source: PISA 2015 database.

30. For most questions, results across the four constituent countries are similar. In England, Northern Ireland and Wales, just under half of headteachers report challenges with regards to the physical infrastructure of their school, compared to around a quarter of headteachers (24 per cent) in Scotland. Likewise, just under a third of headteachers across the UK suggest that instruction was being hindered by a lack of educational material. However, one important point of difference is in respect to a lack of teaching staff. Almost half of headteachers in England (45 per cent) and Scotland (45 per cent) report this to be a problem, significantly more than in Northern Ireland (27 per cent) and Wales (20 per cent). Similarly, 22 per cent of headteachers in England agree that 'inadequate or poorly qualified teachers' were a barrier to instruction within their school, compared to 15 per cent in Wales, eight percent in Scotland and four per cent in Northern Ireland. Hence a lack of appropriately qualified teaching staff seems to be a particularly pressing concern amongst headteachers in England (compared to the rest of the UK).

31. Headteachers were also asked about the conduct of staff in their school, and the extent that this hinders learning amongst pupils. For the majority of questions, headteachers' responses are similar across the different parts of the UK (see Table 11.6). The main point of departure is in respect to the statement '*teachers not meeting individual pupils' needs*'. According to headteachers, this is a factor hindering a smaller proportion of pupils in Northern Ireland (11 per cent) than England (30 per cent) and Scotland (26 per cent), with differences statistically significant at the five per cent level.

Table 11.6 Headteachers' reports of teacher conduct hindering pupils' learning within their school: comparison across the UK

	England	Northern Ireland	Scotland	Wales
Teachers not meeting individual pupils' needs	30%	11%	26%	19%
Teacher absenteeism	24%	30%	21%	24%
Staff resisting change	17%	21%	24%	22%
Teachers being too strict with pupils	5%	4%	9%	4%
Teachers not being well prepared for classes	11%	6%	6%	17%

Source: PISA 2015 database.

Key point

Headteachers' views on the factors hindering instruction within their school are similar across the UK. However, a lack of teaching staff stands out as a particular concern of headteachers in England.

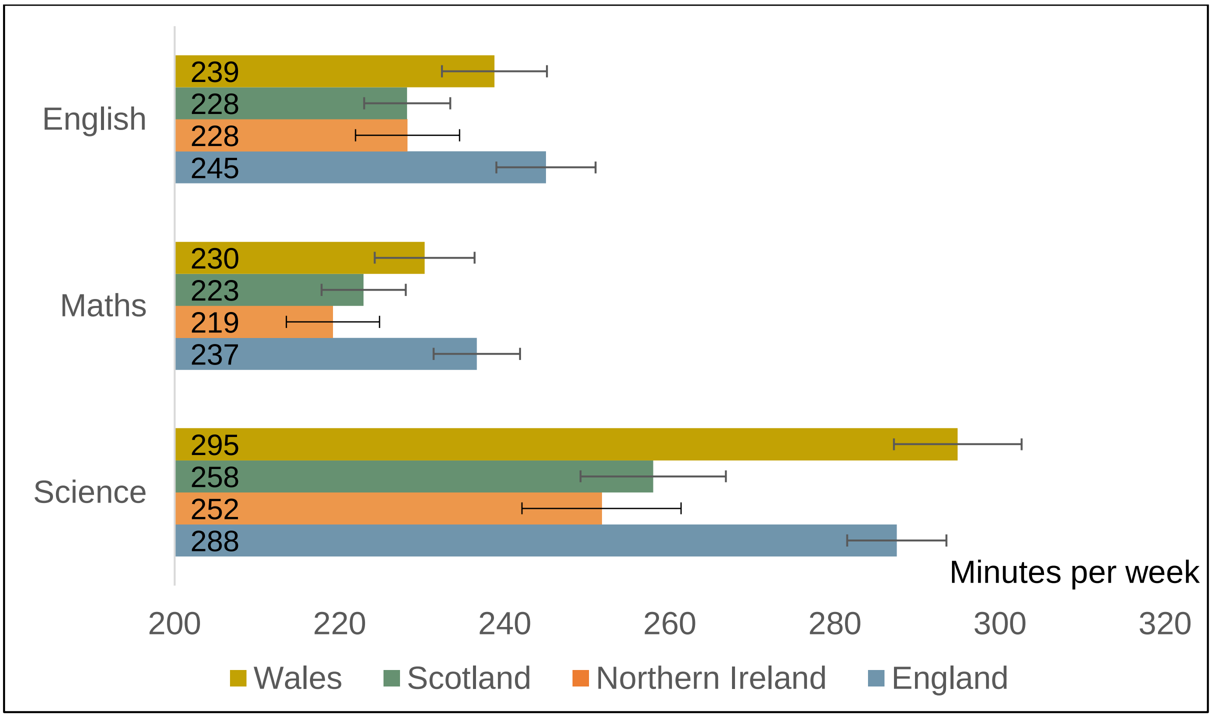
11.8 Are there differences across the UK in the amount of instruction 15-year-olds receive - both inside and outside of school?

32. Is there variation across the UK in the amount of time pupils spend learning science, mathematics and English/Welsh per week? This is important as previous research has suggested that pupils who receive more instruction time in a subject achieve higher PISA test scores[94]. Figure 11.8 therefore investigates whether the number of minutes studying science, mathematics and English differs (on average) across England, Northern Ireland, Scotland and Wales. This is based upon self-reported information from pupils.

33. In all four parts of the UK, young people report spending more time learning science in school than either English/Welsh or mathematics. The difference is typically between 30 and 60 minutes per week, with 15-year-olds in England and Wales indicating they receive around four weekly hours of in-school instruction in English and mathematics, compared to five hours of science.

[94] Lavy (2015).

Figure 11.8 The amount of time pupils report spending learning science, English/Welsh and mathematics in school: a comparison across the UK



Source: PISA 2015 database.

34. Interestingly, pupils in Northern Ireland and Scotland report significantly less instruction time per week across all three subject areas than pupils in England and Wales. For instance, Figure 11.8 indicates that they receive around 40 minutes less instruction in science per week (on average) than their peers in England and Wales. The same holds true, though the difference less pronounced, in English/Welsh (around 15 minutes less per week) and mathematics (around 15 minutes less per week).

35. The PISA background questionnaire also asked pupils how much time they spend per week learning various subjects outside of their required school schedule. This encompasses a wide range of activities, including homework, private tutoring and independent study. Table 11.7 illustrates how these average additional study hours vary across the four constituent countries.

Table 11.7 Pupils' reports of time spent learning in addition to their required schedule: a comparison across the UK

	England	Northern Ireland	Scotland	Wales
Science	3.7 hours	3.8 hours	3.9 hours	3.9 hours
Maths	3.5 hours	4.0 hours	4.0 hours	4.0 hours
English	3.0 hours	3.5 hours	3.9 hours	3.6 hours
Foreign language	1.5 hours	1.8 hours	1.5 hours	1.3 hours
Other	4.9 hours	5.2 hours	6.0 hours	5.1 hours
Mean (all subjects)	**16.6 hours**	**18.4 hours**	**19.2 hours**	**17.9 hours**

Source: PISA 2015 database.

36. On average, 15-year-olds report spending around 18 hours of additional study per week in Northern Ireland and Wales, with this increasing to 19 hours for pupils in Scotland. This is significantly more than their peers in England, who report spending, on average, around 16 and a half hours on additional study per week. Note that a similar finding holds if one considers the median number of additional hours rather than the mean (median = 14 hours in England versus 15 hours in Wales, 16 hours in Northern Ireland and 17 hours in Scotland). This finding is therefore not being driven by a small number of pupils reporting a very high number of additional hours.

37. Further inspection of Table 11.7 indicates that the additional study hours of Scottish, Welsh and Northern Irish pupils (relative to their English peers) is spread across different subject areas. However, the biggest difference seems to be in English/Welsh and mathematics. Young people in Northern Ireland spend over 30 minutes more on average per week studying these subjects in addition to their required schedule than young people in England. For both mathematics and English, additional study time is significantly lower in England than in Scotland, Northern Ireland and Wales at the five per cent threshold. Although differences between these countries tend to be smaller in other subject areas (science, foreign languages, other), point estimates still tend to be lowest in England.

Key point

Across the UK, school pupils report spending more time studying science than any other subject. Scottish, Welsh and Northern Irish pupils report spending, on average, over an hour more on additional study per week than pupils in England.

References

Airasian, P. (2000) *Assessment in the classroom: A concise approach*. New York: McGraw-Hill.

Anders, J. (2012) ‘The link between household income, university applications and university attendance. *Fiscal Studies* 33(2): 185-210.

Anders, J., Jerrim, J. and McCulloch, A. (2016) ‘How much progress do children in Shanghai make over one academic year? Evidence from PISA.’ *AERA Open.*

Archer, L., Osborne, J., DeWitt, J., Dillon, J., Wong, B. & Willis, B. (2013) ‘ASPIRES: Young people’s science and career aspirations, age 10-14.’ London: King’s College.

Avvisati, F. and Keslair, F. (2014) ‘REPEST: Stata module to run estimations with weighted replicate samples and plausible values.’ Statistical Software Components, Boston College Department of Economics.

Barrett, P., Davies, F., Zhang, Y. and Barrett, L. (2015) ‘The impact of classroom design on pupils’ learning: Final results of a holistic mutli-level analysis.’ *Building and Environment* 89: 118-133*.*

Bol, T., Witschge, J., Van de Werfhorst, H., and Dronkers, J. (2014) ‘Curricula tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries.’ *Social Forces* 92(4): 1545-1572.

Boliver, V. (2011) ‘Expansion, differentiation, and the persistence of social class inequalities in British higher education.’ *Higher Education* 61(3): 229-242.

Bradshaw, J., Sturman, L., Vappula, H., Ager, R. and Wheater, R. (2007) ‘Achievement of 15-year-olds in England: PISA 2006 national report.’ (OECD Programme for International Student Assessment). Slough: NFER.

Bradshaw, J., Ager, R., Burge, B. and Wheater, R. (2010) ‘PISA 2009: Achievement of 15-year-olds in England.’ (OECD Programme for International Student Assessment). Slough: NFER.

Bruckauf, Z. and Chzhen, Y. (2016) ‘Education for All? Measuring Inequality of Educational Outcomes Among 15-year-olds Across 39 Industrialized Nations.’ UNICEF Research Report.

Centre Forum. (2016) ‘Education in England. Annual Report 2016.’ London: Centre Forum.

Cole, T. (2007) ‘A Review of 'Attendance and Behaviour' Reports and Guidance Issued by Government and Government Agencies in the British Isles.’ Research Report. Ashford: The Social, Emotional and Behavioural Difficulties Association (SEBDA).

Cosgrove, J. and Cartwright, F. (2014) 'Changes in achievement on PISA: the case of Ireland and implications for international assessment practice.' *Large-scale assessments in education.* DOI: 10.1186/2196-0739-2-2

Coughlan, S. (2015) 'Migrants do not lower school results.' BBC Website. 17 November 2015. Accessed 09/05/2016 from http://www.bbc.co.uk/news/education-34833547

Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., Brown, E., Ahtaridou, E. and Kington, A. (2009) 'The impact of school leadership on pupil outcomes.' *Department for Education Research Report 108*.

Department for Business, Skills and Innovation (2015) *Participation Rates in Higher Education: Academic Years 2006/2007-2013/2014*.

Department for Children, Education, Lifelong Learning and Skills. (2008) 'Science in the National Curriculum.' Cardiff: Department for Children, Education, Lifelong Learning and Skills, 2-26.

Economic and Social Research Council. (2012) 'Education vital for social mobility.' Evidence Briefing. Swindon: ESRC.

Education Endowment Foundation. (2016) 'Behavioural Interventions' in *Teaching & Learning Toolkit*. London: Education Endowment Foundation.

Eleftheriou-Smith, L-M. (2014) 'Britain's schools lack capacity to deal with 'influx' of migrant children says Ofsted chief inspector of schools.' *The Independent.* 30 October 2014. Accessed 09/05/2016 from http://www.independent.co.uk/news/uk/home-news/britains-schools-lack-capacity-to-deal-with-influx-of-migrant-children-says-ofsted-chief-inspector-9827455.html

Estyn (2016) 'The Annual Report of her Majesty's Chief Inspector of Education and Training in Wales 2011-12.' Cardiff: Estyn.

Estyn (2011) 'Standards in science achieved by secondary school pupils educated in England and Welsh.' Cardiff: Welsh Assembly Government Enterprise and Learning Committee. Accessed 30.05.2016 http://www.assembly.wales/Committee%20Documents/Supplementary%20evidence%20-%20Estyn%20(PDF,%20252%20KB)-17112010-206043/supplementary_evidence_-_evidence-English.pdf

European Commission. (2001) *European Report on the quality of school education: Sixteen quality indicators*, Luxembourg: Office for Official Publications of the European Communities.

Friedman, T. L. (2005) *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus and Giroux.

Gutman, L. and Akerman, R. (2008) *Determinants of aspirations [Wider Benefits of Learning Research Report No. 27]*. London: Centre for Research on the Wider Benefits of Learning, Institute of Education, University of London.

Hanushek, E. (2015) 'Time in education: introduction.' *The Economic Journal* 125: F394-F396.

Hanushek, E. and Rivkin, S. (2010) 'Generalizations about using value-added measures of teacher quality.' *The American Economic Review* 100(2): 267-271.

HESA (2016) *HE student enrolments by HE provider, level of study, mode of study and domicile 2014/15.* Statistical First Release 224. Cheltenham: Higher Education Statistics Agency.

Higgins, S., Katsipataki, M., and Coleman, R. (2014). 'Interim Evidence Brief.' London: Education Endowment Foundation.

House of Commons Education Committee. (2014) 'Underachievement in Education by White Working Class Children.' London: Education Committee.

Jerrim, J. and Wyness, G. (2016) 'Benchmarking London in the PISA rankings.' *London Review of Education*

Jones, H. (2012) *A statistical overview of the Welsh language.* Cardiff: Welsh Language Board, 1-127.

Keskpaik, S. and Salles, F. (2013) 'Les Élèves de 15 Ans En France Selon PISA 2012 En Culture Mathématique : Baisse Des Performances et Augmentation Des Inégalités Depuis 2003'. Note d'information 13.31. MEN-DEPP. Keskpaik and Salles.

Lavy, V. (2015) 'Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries.' *The Economic Journal* 125(588): F397-F424.

Lupton, R., and Kintrea, K. (2011) 'Can community-based interventions on aspirations raise young people's attainment?' *Social Policy and Society* 10(03): 321-335.

Machin, S. and McNally, S. (2008) 'The literacy hour.' *Journal of Public Economics* 92(5): 1441-1462.

Macmillan, L., Tyler, C. and Vignoles, A. (2015) 'Who Gets the Top Jobs? The Role of Family Background and Networks in Recent Graduates' Access to High-status Professions.' *Journal of Social Policy* 44(3): 487-515.

Mau, W. C. (2003). 'Factors that influence persistence in science and engineering career aspirations.' *The Career Development Quarterly*, 51(3), 234-243.

Micklewright, J. and Schnepf, S. (2006) 'Response bias in England in PISA 2000 and 2003.' *Department for Education and Skills Research Report 771*.

Micklewright, J. and Schnepf, S. (2006) 'Inequality of Learning in Industrialised Countries.' *IZA working paper 2517*. Bonn: IZA.

Miller, R., Murnane, R. and Willett, J. (2008) 'Do Teacher Absences Impact Student Achievement? Longitudinal Evidence From One Urban School District.' *Educational Evaluation and Policy Analysis* 30(2): 181-200.

Morgan, S. (2005) *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford: Stanford University Press.

NCES Website. 'Program for International Student Assessment (PISA): Countries, Participation in PISA by Year.' Institute of Education, National Center for Education Statistics. Accessed 30.05.2016 from https://nces.ed.gov/surveys/pisa/countries.asp

Neilson, C. and Zimmerman, S. (2011) 'The effects of school construction on test scores, school enrolment and house prices.' *IZA discussion paper dp6106.* Bonn: IZA.

OECD (2001) *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.

OECD (2009) 'PISA: Who are the top performers?' In *Highlights from Education at a Glance 2009*. Paris: OECD.

OECD (2010) 'The High Cost of Low Educational Performance. The long-run economic impact of improving PISA outcomes.' Paris: OECD.

OECD (2012). *Learning beyond Fifteen*. Paris: Organisation for Economic Co-operation and Development.

OECD (2013a) 'PISA 2015 draft mathematics framework'. Paris: OECD.

OECD (2013b) 'PISA 2015 draft reading framework'. Paris: OECD.

OECD (2013c) 'PISA 2015 draft science framework'. Paris: OECD.

OECD (2013d) 'PISA 2015 integrated design'. Paris: OECD.

OECD (2014a) *Education Indicators in Focus.* 'How much time do primary and lower secondary students spend in the classroom?' Paris: OECD.

OECD (2014b) 'PISA 2012 Technical Report'. Paris: OECD.

OECD (2015) 'Helping immigrant students to succeed at school – and beyond.' Paris: OECD.

OECD (2016) 'PISA 2015 assessment and analytical framework: science, reading, mathematic and financial literacy.' Paris: OECD.

ONS (2015) 'Migration Statistics Quarterly Report: November 2015.' Statistical Bulletin. London: Office for National Statistics.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. and Schiefele, U. (2006) *PISA 2003: Untersuchungen Zur Kompetenzentwicklung Im Verlauf Eines Schuljahres*. Münster: Waxmann Verlag GmbH.

Reynolds, D. (2007) 'Schools learning from their best. The within-school variation project.' Research Report. London: National College for Teaching and Leadership.

Reynolds, G. (2008) 'The Impacts and Experiences of Migrant Children in UK Secondary Schools.' *Sussex Centre for Migration Research Working Paper Number 47.* The University of Sussex.

Russell Group (2016) 'About'. Accessed 30/05/2016 http://russellgroup.ac.uk/about/.

Rutkowski, L., von Davier, M. and Rutkowski, D. (Eds.) (2013) *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis.* Boca Raton: CRC Press.

Sacerdote, B. (2011) 'Peer effects in education: How might they work, how big are they and how much do we know thus far?.' *Handbook of the Economics of Education* 3: 249-277.

Sadler, P. M., Sonnert, G., Hazari, Z., and Tai, R. (2012) 'Stability and volatility of STEM career interest in high school: A gender study.' *Science Education* 96(3): 411-427.

Social Mobility and Child Poverty Commission. (2012) 'Fair Access to Professional Careers: A progress report by the Independent Reviewer on Social Mobility and Child Poverty.' London: Social Mobility and Child Poverty Commission.

StatsWales Website. (2016) 'Examination achievements of 15-year-olds by gender.' Accessed 30.05.2016 https://statswales.gov.wales/Catalogue/Education-and-Skills/Schools-and-Teachers/Examinations-and-Assessments/Key-Stage-4/examinationachievementsofpupilsaged15-by-gender

Strand, S. and Winston, J. (2008) 'Educational aspirations in inner city schools.' *Educational Studies* 34(4): 249-267.

Taylor, C., Rees, G. and Davies, R. (2013) 'Devolution and geographies of education: the use of the Millennium Cohort Study for 'home international' comparisons across the UK.' *Comparative Education* 49(3): 290-316.

UNEP (2012) *21 Issues for the 21st Century: Result of the UNEP Foresight Process on Emerging Environmental Issues*. Nairobi: United Nations Environment Programme (UNEP).

Weiss, D., Wiese, B. and Freund, A. (2014) 'What should I do with my life? Motivational, personal, and contextual factors in mastering the transition of graduating from high school.' In Schoon, I., and Eccles, J. S. (Eds.). *Gender differences in aspirations and attainment: A life course perspective*. Cambridge: Cambridge University Press.

Welsh Government Website. (2016) 'National School Categorisation System.' Published 21.07.2016. Accessed 01.08.2016 http://gov.wales/topics/educationandskills/schoolshome/raisingstandards/schoolcategorisation/?lang=en

Wheater, R., Ager, R., Burge, B. and Sizmur, J. (2014) 'Achievement of 15-year-olds in England: PISA 2012 national report.' (OECD Programme for International Student Assessment). Slough: NFER.

Woessmann, L. (2016) 'The Importance of School Systems: Evidence from International Differences in Student Achievement.' *Journal of Economic Perspectives* 30(3): 3-32.

World Bank. (2003) 'Strategic Approaches to Science and Technology in Development'. Research Report.

Appendix A. Background to the PISA study

1. The Programme for International Student Assessment (PISA) is a global benchmarking study of pupil performance by the Organisation for Economic Co-operation and Development (OECD). The following sections outline the development of the study, what PISA measures, how to interpret the PISA scales, how PISA is administered and details of the PISA sample in Wales. These sections outline some of the detailed international requirements that countries must meet in order to ensure confidence in the findings.

A.1. Development of the study

2. Five international contractors designed and implemented the PISA 2015 study on behalf of the OECD. These organisations were the Educational Testing Service (ETS), Westat, cApStAn Linguistic Control, Pearson and the German Institute for International Education Research (DIPF). By using standardised survey procedures and tests, the PISA study aims to collect data from around the world that can be compared, despite differences in language and culture.

3. The framework and specification for the study were agreed internationally by the PISA Governing Board, which comprises of representatives from each participating country. Both the international consortium and participating countries submitted test questions for inclusion in the assessment. After the questions were reviewed by an expert panel (convened by the international PISA consortium), countries were invited to comment on their difficulty, cultural appropriateness, and curricular and non-curricular relevance.

4. A field trial was carried out in every participating country in 2014. The outcomes of this field trial were used to finalise the contents and format of the tests and questionnaires for the main survey in 2015. A 'mode effect' study was also conducted by ETS as part of this field trial. The purpose of this aspect of the field trial was to establish how the switch from paper to computer assessment influences pupils' responses to the PISA test questions, and to ensure results from PISA 2015 can be linked to previous cycles. Further details on the design of this mode effect study are available from https://www.oecd.org/pisa/pisaproducts/2015-Integrated-Design.pdf

5. Strict international quality standards are applied to all stages of the PISA survey to ensure equivalence in translation and adaptation of instruments, sampling procedures and survey administration in all participating countries.

A. 2. What does PISA measure?

Science

6. Science was the main focus in PISA 2015, as it was in PISA 2006. Full details on the PISA 2015 science framework are available from http://www.oecd-ilibrary.org/docserver/download/9816021ec003.pdf?expires=1462366012&id=id&accname=guest&checksum=DF06918825ED39FEF30E42BB8F8BC573

7. PISA aims to measure not just science as it may be defined within the curriculum of participating countries, but the scientific understanding which is needed in adult life. This is defined as the capacity for pupils to identify questions, acquire new knowledge, explain scientific phenomena, and draw evidence-based conclusions about science-related issues. Individuals with this capacity also understand the characteristic features of science as a form of human knowledge and enquiry, are aware of how science and technology shape their lives and environments, and are willing and able to engage in science-related issues and with the ideas of science, as a reflective citizen. Therefore, PISA assessments measure not only scientific knowledge, but also scientific competencies and understanding of scientific contexts.

8. Scientific 'knowledge' in PISA constitutes the links that aid understanding of related phenomena. While the scientific concepts are familiar (relating to physics, chemistry, biological sciences and earth and space sciences), pupils are asked to apply them to the content of the test items, and not simply to recall facts. This therefore includes both knowledge of the natural world and technological artefacts (*content knowledge*), knowledge of how such ideas are produced (*procedural knowledge*) and an understanding of the underlying rationale for these procedures and the justification for their use (*epistemic knowledge*). However, the PISA 2015 test was weighted towards the first of these knowledge types. Specifically, content knowledge was targeted in 53 cent of the assessment questions, procedural knowledge in 33 per cent and epistemic knowledge in 14 per cent. The content domains can be further divided into: *living systems*, *physical systems*, and *earth and space* systems. A third of items (33 per cent) covered the physical system, 40 per cent the living system and 27 per cent earth and space sciences.

9. Scientific competencies are centred on the ability to acquire, interpret and act upon evidence. Three processes are identified in PISA. These are the ability to:

- Explain phenomena scientifically. To recognise, offer and evaluate explanations for a range of natural and technological phenomena.

- Evaluate and design scientific enquiry. Describe and appraise scientific investigations and propose ways of addressing questions scientifically.
- Interpret data and evidence scientifically. Analyse and evaluate data, claims and arguments in a variety of representations and draw appropriate scientific conclusions.

10. Among all the science test items, 48 per cent of the total test score points were targeted within the ‘explaining phenomena scientifically’ domain. A total of 30 per cent of total test score points were targeted within ‘interpreting data and evidence scientifically’, with the remaining 22 per cent within ‘evaluating and designing scientific enquiry’.

11. Scientific contexts concern the application of scientific knowledge and the use of scientific processes. This includes personal, local, national and global issues, both current and historical, which demand some understanding of science and technology. Test question contexts were spread across personal, local/national and global settings in a roughly 1:2:1 ratio, as was the case in PISA 2006 (the last time science was the focus of PISA).

12. Around a third of PISA 2015 science test items were found within each of the following three categories:

- Open constructed response. These items required pupils to provide written responses, ranging from a phrase up to a short paragraph. A small number of questions also required drawing a simple graph or diagram, using the drawing editor provided on the computer-test platform.
- Simple multiple choice. These questions required pupils to select a single response from a set of four options, or to select a ‘hot spot’ (i.e. a selectable element) within a graphic or passage of text.
- Complex multiple choice. This includes responses to a series of yes/no questions, selection of more than one option from a list, completion of sentences via drop-down choices, and responses where pupils interact with the computer-testing software to ‘drag-and-drop’. It also includes pupils’ responses to interactive tasks, such as manipulating variables in a simulated scientific experiment.

Mathematics

13. Mathematics was the main focus in the 2012 and 2003 PISA cycles. It was a minor domain in PISA 2015. Full details on the PISA 2015 mathematics framework are available from http://www.oecd-ilibrary.org/docserver/download/9816021ec005.pdf?expires=1462366094&id=id&accname=guest&checksum=0B6059225B81CAC7E6FE8CE8A02EAD1E

14. PISA aims to assess pupils' ability to put their mathematical knowledge to functional use in different situations in adult life, rather than assess what is taught in participating countries. The OECD defines this ability as:

'*an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena. It assists individuals in recognising the role that mathematics plays in the world and to make the wellfounded judgements and decisions needed by constructive, engaged and reflective citizens*'. (OECD 2013a)

15. In order to demonstrate this capacity, pupils need to have factual knowledge of mathematics, skills to carry out mathematical operations and methods, and an ability to combine these elements creatively in response to external situations.

16. PISA recognises the limitations of using a timed assessment in collecting information about something as complex as mathematics. It aims to tackle this by having a balanced range of questions that assess different elements of pupils' mathematical processing ability. This is the process through which a pupil interprets a problem as mathematical and draws on his/her mathematical knowledge and skills to provide a sensible solution to the problem.

17. PISA prefers context-based questions which require the pupil to engage with the situation and decide how to solve the problem. Most value is placed on tasks that could be met in the real world, in which a person would authentically use mathematics and appropriate mathematical tools, to solve these problems. Some more abstract questions that are purely mathematical are also included in the assessment.

Reading

18. Reading was the main focus in the first PISA study in 2000 and also in 2009. It was a minor domain in PISA 2015. Full details on the PISA 2015 reading framework are available from http://www.oecd-ilibrary.org/docserver/download/9816021ec004.pdf?expires=1462366215&id=id&accname=guest&checksum=FC03724295B8824B7A78A7560C1DCDB1

19. Reading in PISA focuses on the ability of pupils to use information from texts in situations which they encounter in their life. Reading in PISA is defined as '*understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society*' (OECD 2009).

20. The concept of reading in PISA is defined by three dimensions: the format of the reading material, the type of reading task or reading aspect, and the situation or the use for which the text was constructed.

21. The first dimension, the text format, divides the reading material into continuous and non-continuous texts. Continuous texts are typically composed of sentences which are organised into paragraphs. Non-continuous texts are not organised in this type of linear format and may require, for example, interpretation of tables or diagrams. Such texts require a different reading approach to that needed with continuous text.

22. The second dimension is defined by three reading aspects: retrieval of information, interpretation of texts, and reflection on and evaluation of texts. Tasks in which pupils retrieve information involve finding single or multiple pieces of information in a text. In interpretation tasks pupils are required to construct meaning and draw inferences from written information. The third type of task requires pupils to reflect on and evaluate texts. In these tasks pupils need to relate information in a text to their prior knowledge, ideas and experiences.

23. The third dimension is that of situation or context. The texts in the PISA assessment are categorised according to their content and the intended purpose of the text. There are four situations: reading for private use (personal), reading for public use, reading for work (occupational) and reading for education.

A.3. What do the PISA proficiency levels mean?

24. PISA uses proficiency levels to describe the types of skills that pupils are likely to demonstrate and the tasks that they are able to complete. Test questions that focus on simple tasks are categorised at lower levels, whereas those that are more demanding are categorised at higher levels. The question categorisations are based on both quantitative and qualitative analysis, taking into account question difficulty as well as expert views on the specific cognitive demands of each individual question. All PISA questions have been categorised in this manner.

25. Pupils described as being at a particular level not only demonstrate the knowledge and skills associated with that level but also the proficiencies required at lower levels. For example, all pupils proficient at Level 3 are also considered to be proficient at Levels 1 and 2. The table below shows the score points for each level in each PISA subject area.

Table A1. The correspondence between PISA test points and proficiency levels

Proficiency levels	Science	Mathematics	Reading
Level 6	>707.93	>669.30	>698.32
Level 5	633.33 to 707.93	606.99 to 669.30	625.61 to 698.32
Level 4	558.73 to 633.33	544.68 to 606.99	552.89 to 625.61
Level 3	484.14 to 558.73	482.38 to 544.68	480.18 to 552.89
Level 2	409.54 to 484.14	420.07 to 482.38	407.47 to 480.18
Level 1a	334.94 to 409.54	357.77 to 420.07	334.75 to 407.47
Level 1b	260.54 to 334.94	357.77<	262.04 to 334.75

A.4. The PISA test design

26. PISA uses a complex test design. Test questions are first separated into distinct 30 minute ‘clusters’. These clusters are then combined to generate a total of 66 test forms. Each form is made up of four clusters, and thus contains two hours of test questions. Pupils are then randomly assigned, with differing probabilities, to one of the 66 forms. Within each test form, a proportion of the questions were ones used in previous cycles. It is this that facilitates measurement of change in PISA test scores over time. A summary of the PISA 2015 assessment design is provided in Figure A1.

27. Roughly a third of pupils answered one hour of science and one hour of reading test questions (form 31 to 42). A further third of pupils answered one hour of science and one hour of mathematics questions (form 43 to 54), while just over a fifth (22 per cent) received one hour of science and one hour of Collaborative Problem Solving (CPS) questions (form 91 to 96)[95]. The vast majority of pupils (88 per cent) therefore answered test questions covering two out of the four PISA domains. The remaining 12 per cent of pupils were assigned to test forms that covered three out of the four PISA subject areas. These pupils received one hour of science questions, plus two 30 minute clusters of questions covering two out of the three other domains. These combinations were:

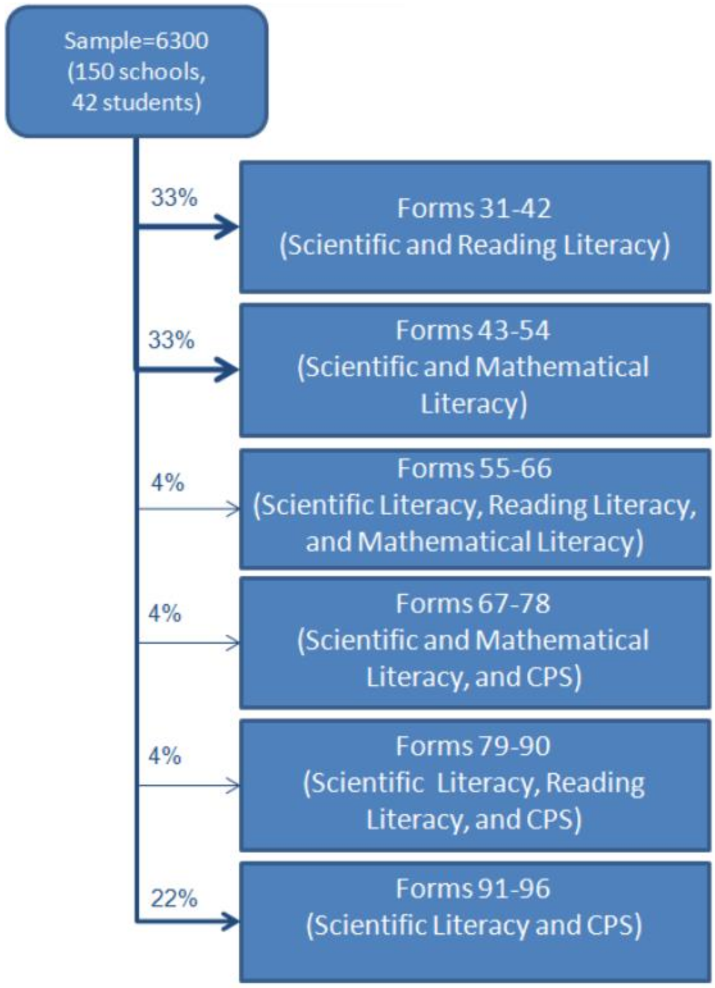
a. Form 55-66: One hour science, 30 minutes reading and 30 minutes mathematics

b. Form 67-78: One hour science, 30 minutes mathematics and 30 minutes CPS

c. Form 79-90: One hour science, 30 minutes reading and 30 minutes CPS

Figure A1. A summary of the PISA 2015 test design

[95] The hour of scientific literacy included 30 minutes of 'trend' questions (i.e. those that have been used in previous PISA cycles) with the other 30 minutes consisting of 'new' science items (not used in previous PISA cycles).



28. The main implication of this complex design is that no single pupil is presented with all PISA test questions. Instead, statistical methods are used to estimate the likelihood that the pupil would be able to answer correctly the questions which they have not actually been asked. This is executed using a complex item-response theory (IRT) model, with further details on this process available in Rutkowski, von Davier and Rutkowski (2013) and the PISA 2015 technical report (OECD, forthcoming).

A.5. Administration

29. The survey administration was carried out internationally on behalf of the OECD by a consortium of five organisations (see section A1 above). The consortium worked with the PISA National Centre within each country, through the National Project Manager (NPM). For Wales the National Centre was formed of three organisations: RM Education, World Class Arena Limited and the UCL Institute of Education.

30. National Centres were responsible for making local adaptations to test questions, manuals and the background questionnaires. They were also responsible for translation where necessary. All materials were translated into Welsh, with pupils in Wales asked to choose the language in which they wished to complete tests and questionnaires.

31. National Centres were also responsible for supplying the information necessary for sampling to be carried out. School samples were selected by the PISA consortium, while pupil samples within schools were selected by RM Education using software supplied by the international consortium.

32. In Wales, pupils sat the two-hour PISA assessment in November-December 2015 under test conditions, following the standardised procedures implemented by all countries. In Scotland, the PISA survey was carried out earlier in 2015.

33. Tests and questionnaires were generally administered in a single session. Pupils first completed the two hour PISA assessment. After a short break, they were then asked to complete the pupil background questionnaire (35 minutes), educational career questionnaire (10 minutes) and ICT familiarity questionnaire (10 minutes). The total length of an assessment session was around three and a half hours. The survey was administered by test administrators employed and trained by RM Education.

34. In each country participating in PISA, the minimum number of participating schools was 150. For countries using computer-based assessment and participating in the Collaborative Problem Solving (CPS) study, 42 pupils were then randomly selected within each school. Countries using paper-based assessment, or not participating in the CPS study, were required to randomly select 35 pupils per school. The minimum target sample size was therefore 6,300 pupils in countries involved in the CPS study (including the UK) and 5,250 in countries that were not.

35. In the case of the UK and of some other countries, slight variations on this design were allowed. Specifically, a greater number of schools across the UK were sampled than strictly required, while the number of pupils per school was slightly lower (30 pupils as opposed to 42). Consequently, the number of pupils and schools participating in PISA from across the UK exceeds the minimum requirements set by the OECD. This alternative sample design was used in the UK due to the need to over-sample certain parts of the country; for example, larger samples were drawn for Wales, Scotland and Northern Ireland than strictly required. This was to make sure it was possible to provide separate PISA results for the four constituent parts of the

UK. In some countries additional samples were drawn for other purposes, for example to enable reporting of results for a particular sub-group (e.g. indigenous pupils in the case of Australia). In very small countries with less than 150 schools, PISA was completed as a school census (meaning all eligible secondary schools were included).

36. The pupils included in the PISA study are generally described as '15-year-olds', but there is a small amount of leeway in this definition depending on the time of testing. In the case of Wales the sample consisted of pupils aged from 15 years and two months to 16 years and two months at the beginning of the testing period.

37. Countries were required to carry out the study during a six-week period between March and August 2015. However Wales was permitted to test outside this period because of the problems for schools caused by the overlap with the GCSE preparation and examination period. In Wales the study took place between November 5th and December 7th 2015. This is consistent with how PISA has been administered in Wales since 2006.

38. Each participating school in Wales was assigned a test date during this period by the National Centre. Before this date schools received two packages. One package contained the USB sticks used to deliver the PISA 2015 test (and had the PISA 2015 test questions loaded on), post-testing certificates and return materials. The second package was a list of user logins for pupils on the test day. This was issued in advance in order that the set-up on the morning of the test was as efficient as possible. Schools were then asked to conduct a system diagnostic test using one of the USB sticks provided by the National Centre. This allowed the school to run a number of checks on their hardware to ensure that the PISA test would run on the school's computers on the actual test day. Although the data gathered allowed the National Centre to determine whether the equipment at schools had the potential to run the PISA 2015 test software, it did not provide data on a number of key elements in order to plan and run test days (e.g. whether the computers to be used in the testing could all be found within a single room or were spread across the school).

39. To assist schools on the day of the PISA 2015 test, a Test Administrator (TA) was assigned to every school. Their responsibility was to help set-up the tests on the school's computers, assist in invigilating the test session(s) and help resolve any problems that may arise. All TAs were either ex-teachers or had worked within a school environment before, and were hence accustomed to the day-to-day running within a school. All received training prior to the testing period. Typically, one test

administrator was assigned per school. However, an additional TA was provided in a small number of instances where schools did not have the capacity to test all participating pupils in a single room. A member of staff within each school was also assigned as the School Co-ordinator for PISA 2015, with whom the TA and National Centre would liaise before, during and after the test day.

40. On the actual test day, TAs arrived at schools from 7.30am/8.00am to complete set up tasks. However this was reliant on the school being prepared, with their School Co-ordinator and IT Network Manager being available, and with the relevant materials (e.g. USB sticks and log-in details) to hand. On occasion this was not the case which delayed the start of the test. At schools where pupil behaviour proved to be disruptive, this was managed by the TA along with senior members of school staff. TAs worked at the school until mid-afternoon completing administrative duties, including making the packages to be returned to the National Centre by courier.

41. At the end of each test session, the test administrators were required to complete a ‘session report form’. This included the following questions:

- Were there any problems with assessment conditions? (e.g. significant disciplinary issues).
- Did you notice any pupils attend the session but not answer any test items at all? (If yes, write the number of pupils affected)
- Were there any pupils that started the test, but were unable to complete it due to computer failure? (If yes, write the number of pupils affected)
- Were there any pupils that started the test, but were unable to complete it for other reasons? (If yes, write the number of pupils students affected)
- Were there any pupils unable to start the session at all due to computer failure? (If yes, write the number of pupils affected)

42. In Wales, 161 test sessions took place across the 140 participating schools. A total of 122 schools (87 per cent) completed the PISA assessment in a single test session, while two test sessions were used in 15 schools (11 per cent) and three test sessions in three participating schools (two per cent). Test administrators reported some issues with assessment conditions in 21 (13 per cent) test sessions, though these ranged from relatively minor (e.g. some technical ICT assistance needed when setting up) through to more serious issues (e.g. poor pupil behaviour). There were 28 pupils whose test were interrupted, three for computer failure and 25 for other

reasons (e.g. pupil arrived late, challenging behaviour). Test administrators reported 18 pupils who they believed to not be answering any test questions at all.

43. Following the final day of testing at each school, a collection of the packages put together by the TA was requested by the National Centre. It was imperative that these materials were returned quickly so that these could be reconciled and any manual test uploads completed as soon as possible. As with deliveries, collections were tracked from request through to the delivery of the school package at the National Centre via an Excel spreadsheet. Once received the package was logged in and USBs reconciled. A number of schools required a revised collection date due to the school either being closed when the courier arrived, or the reception not having the package available. However these instances were minimal and on the whole the process was efficient and effective.

Appendix B. Sample design and response rates

Sample design

1. PISA requires each country to randomly recruit a minimum of 150 schools, with a minimum of 6,300 pupils completing the tests[96]. In the UK, and some other countries, the number of pupils and schools drawn exceeds this. Specifically, larger samples have been drawn from Wales, Scotland and Northern Ireland than strictly necessary to generate a representative, well-powered sample for the UK. This has been done to ensure it is possible to report robust, highly powered estimates separately for England, Scotland, Wales and Northern Ireland. Some other countries draw larger samples for other purposes, such as reporting results for particular sub-groups (e.g. Australia has traditionally oversampled indigenous pupils to ensure separate PISA results can be reported for this group). In very small countries with less than 150 schools (e.g. Iceland), PISA is essentially a school-level census, including a sample of pupils from every secondary school.

2. Throughout the national report we describe PISA as a study of 15-year-olds. There is actually a small difference in this definition, which depends upon the time of the test. In England, Wales and Northern Ireland the sample consisted of pupils aged from 15 years and two months to 16 years and two months at the beginning of the testing period.

3. The sampling frame for England, Wales and Northern Ireland was produced using lists of all schools with 15-year-olds in the 2013/14 academic year. A total of 3 per cent of pupils were excluded from the sampling frame. These were individuals who attended Hospital Schools, Special Schools, Alternative Provision Units, Pupil Referral Units and Prison Schools. After making these exclusions, 4,288 schools remained in the sampling frame.

4. Countries must follow strict international sampling procedures to ensure comparability. This process is formed of several stages. First, each country selects a set of ‘explicit stratification’ variables. Although these differ across countries, geographic region and school type are amongst the most common choices. Appendix Table B1 provides information on the explicit stratification variables that were used in Wales. This included school type, region and gender. Within each of these explicit strata, schools are then ranked by a variable (or set of variables) that

[96] This minimum number of pupils refers to countries that participated in the Collaborative Problem Solving (CPS) assessment in PISA 2015. For those countries that chose not to complete the CPS assessment, the minimum number of pupils was 5,250.

are likely to be strongly associated with PISA test scores. This is known as implicit stratification, with historic GCSE performance of the school the most important variable used for this purpose in Wales.

5. The sampling frame (a list of all eligible schools) and their populations was then sent to the international consortium, who drew the sample of schools. Schools were randomly chosen to participate from within each explicit strata, with probability proportional to size. The international consortium then sent the list of selected schools back to the national project team. In Wales this list comprised of 161 main study schools. By the time of the test, nine schools were dropped. This was mainly due to school closure, having no pupils who met the PISA population definition, or only having pupils with significant special educational needs. The final total of schools chosen and eligible to participate was therefore 152.

Appendix Table B1. The variables used to stratify the PISA sample in Wales

Explicit strata	Implicit strata
Schools Type	**GCSE school performance**
Maintained	Band 1 (lowest)
Independent	Band 2
Region	Band 3
North Wales	Band 4
Powys & South West	Band 5 (highest)
South East Wales	Band not known
Gender composition	**Local Authority**
Boys school	Varies within region
Girls school	
Mixed school	

6. The schools randomly selected into the PISA sample were then invited to participate in the study. Those that agreed were asked to supply a list of all pupils who met the PISA age definition at the start of the testing period (November 2015). The majority of these children were in Year 11.

7. Inevitably, some schools declined to participate. In such instances, PISA uses a system of ‘replacement schools’. This means that, if a school declines to participate, a substitute is entered in its place. Two replacement schools are selected by the international consortium per ‘main study’ school. These are typically the schools that follow the non-participating school on the sampling frame (which has been explicitly and implicitly stratified). This should mean that the replacement

schools are similar to the one which declined to take part (at least in terms of the variables used to stratify the sample). For further information on this process, readers are directed to the PISA technical report[97].

8. RM education then used specialist software (Keyquest), provided by the international consortium, to randomly select the 30 pupils from each participating school. These pupils, who all met the PISA age definition, were then invited to participate in the study.

Target response rates

9. PISA has strict rules surrounding school response rates. Countries are set a target of an 85 per cent school level response rate, before replacement schools have been taken into account. If a country meets these criteria, then the use of replacement schools is not strictly necessary (although, in many countries, replacements for non-participating schools are included in any case).

10. Conversely, if the response rate of initially selected schools falls below 65 per cent, the sample is deemed unacceptable by the international consortium. In such circumstances, the chance of the sample being biased (i.e. no longer nationally representative) is too great. Hence the country will be excluded from the international report, due to poor data quality.

11. If the response rate for initially selected schools is between 65 per cent and 85 per cent, then an 'acceptable' overall response rate can still be achieved through the use of replacement schools. However, the target response rate also moves upwards. For instance, if only 70 per cent of initially sampled schools are willing to participate, then a country must achieve a 94 per cent response rate after the substitute schools have been entered. If this target is achieved, results for the country will be included in the international report.

12. Finally, a country may achieve a before replacement response rate between 65 per cent and 85 per cent, but then fail to meet the revised target after replacement schools have been included. This is known as the 'intermediate zone'. If a country falls into this area, their results may still be included in the international report. However, the country is required to provide an analysis of the likely non-

[97] At the time of writing, the most recent technical report available is for PISA 2012. See OECD (2014b:76) for details on the use of replacement schools.

response bias to the international consortium. This report will then be scrutinised by referees from the international contractor, who will deem whether the data collected are sufficiently robust for meaningful cross-national comparisons to be made.

13. PISA also enforces strict rules around pupil-level response. First, in order for a school to be considered as ‘participating’, at least 50 per cent of the selected eligible pupils must take part (e.g. assuming all 30 pupils selected within a school are indeed eligible for the study, at least 15 must complete the PISA test). Second, an overall response rate of 80 per cent amongst selected students within participating schools is required.

Response rates in PISA 2015

14. A total of 140 schools and 3,451 pupils completed the PISA 2015 study in Wales. Appendix Table B2 provides further details on how the schools were distributed between initially selected schools, first replacement schools, and second replacement schools (along with non-participants[98]). The final response rate for Wales was 86 per cent of the initially sampled schools and 92 per cent after replacements were considered. This is within the ‘acceptable’ zone, and thus fully consistent with the OECD requirements.

Appendix Table B2. School response rates

	Wales
Participating main sample schools	131
Participating first-replacement schools	9
Participating second-replacement schools	0
Non-participating schools	12
Total initially sampled	**152**

Notes: Schools with less than 50 percent of eligible pupils completing the test are considered non-participants. Figures refer to the number of schools.

15. The international report produced by the OECD includes the United Kingdom as a single country, rather than in its four constituent parts. Hence it is the response rate for the United Kingdom as a whole that determines entry into the international report, and whether a non-response bias analysis is required. The overall UK response rate is weighted by the population size in each constituent country, as well as by school size. The weighted UK-wide response rate was 84 per cent of main

[98] Here a ‘non-participant’ refers to where neither the initially selected school, nor its two replacement schools, took part in the PISA study.

sample schools, and 93 per cent after replacement. This fully met the participation requirements.

16. Appendix Table B3 provides details on pupil level response. Of the 4,239 pupils initially selected to participate in Wales, 3,451 successfully completed the PISA study. A total of 315 pupils were excluded for reasons of SEN, enrolment elsewhere, or ineligibility. Finally, 473 pupils were absent on the day of the test. This represents a final response rate (among eligible pupils) of 88 per cent. This exceeds the 80 per cent threshold required by the international contractors for inclusion in the international report.

Appendix Table B3. Pupil-response rates in Wales

	Number of pupils
Assessed	3,451
Absent	473
Excluded	275
Ineligible	40
Total initially sampled	**4,239**

Source: PISA 2015 national data file.

Appendix C. Testing statistical significance in PISA across cycles

1. To test statistical significance across two independent samples (e.g. a comparison of mean test scores across countries in PISA) a two-sample t-test can be applied. For instance, if one were to compare the mean score in country A to the mean score in country B, the t-statistic to be used in statistical significance testing would be:

$$T - stat = \frac{(\mu_A - \mu_B)}{\sqrt{SE_A^2 + SE_B^2}} \quad \text{(C1)}$$

Where:

μ_A = Mean score in country A

μ_B = Mean score in country B

SE_A = Standard error in country A

SE_B = Standard error in country B

2. However, when testing for statistical significance over time in international assessments such as PISA, an extra term has to be added to the denominator of equation C1. This is known as the 'link error'. The link error attempts to capture the fact that there is a degree of uncertainty when equating (or linking) tests together from different cycles. Therefore, to compare mean scores for a country across two time points (e.g. average PISA scores in 2006 and 2015) the following formula for the t-statistic should be applied:

$$T - stat = \frac{(\mu_1 - \mu_2)}{\sqrt{SE_1^2 + SE_2^2 + LE_{1,2}^2}} \quad \text{(C2)}$$

Where:

μ_1 = Mean score at time point 1 (e.g. 2015)

μ_2 = Mean score at time point 2 (e.g. 2006)

SE_1 = Standard error at time point 1

SE_2 = Standard error at time point 2

$LE_{1,2}$ = The link error for comparisons between time point 1 and time point 2

3. In PISA, a common link error is specified which can be applied in all countries. Details on how this link error is calculated will be provided by the OECD in the PISA 2015 technical report. Appendix Table C1 provides the value of the link error to be applied when comparing estimates from PISA 2015 to previous cycles.

Appendix Table C1. The value of the link error when comparing results from PISA 2015 to previous cycles

	Science	Mathematics	Reading
2006	4.4821	3.5111	6.6064
2009	4.5016	3.7853	3.4301
2012	3.9228	3.5462	5.2535

4. We demonstrate the use of these link errors by working through an example. The mean PISA science score for Northern Ireland in 2006 is equal to 508.14 with a standard error of 3.34. In 2015, the mean science score in Northern Ireland is equal to 500.09 with a standard error of 2.79. Finally, as Appendix Table C1 illustrates, the value of the link error for comparing mean PISA 2006 and 2015 science scores is 4.4821. Using equation C2, the t-statistic for the change in the mean score for Northern Ireland between 2006 and 2015 is:

$$\frac{(500.09 - 508.14)}{\sqrt{2.79^2 + 3.34^2 + 4.48^2}} = -1.289$$

5. The correct estimate of the t-statistic is therefore -1.289. As this is smaller in absolute value than the ‘critical value’ of -1.99[99] (based upon a standard two-tailed test with a five per cent significance threshold), one should fail to reject the null hypothesis that average PISA science scores in Northern Ireland are the same in 2006 and 2015. (Note that, if one were to exclude the link error from this calculation, the estimated t-statistic would become -1.85, which is still below the critical value in absolute magnitude).

6. A 95 per cent confidence interval can also be constructed for the change between two PISA test sore statistics over time using the following formula:

[99] As the PISA sample design includes 80 replicate weights, the number of degrees to freedom is approximately 79. Consequently, the critical t-value for a two-tailed significance test at the 5 per cent level is 1.99.

$$(\mu_1 - \mu_2) \mp 1.99\sqrt{SE_1^2 + SE_2^2 + LE_{1,2}^2} \qquad \text{(C3)}$$

Where:

μ_1 = Mean score at time point 1 (e.g. 2015)

μ_2 = Mean score at time point 2 (e.g. 2006)

SE_1 = Standard error at time point 1

SE_2 = Standard error at time point 2

$LE_{1,2}$ = The link error for comparisons between time point 1 and time point 2

7. Returning to the example of the change in mean science scores in Northern Ireland between 2006 and 2015, the formula in equation C3 becomes:

$$(508.14 - 500.09) \mp 1.99\sqrt{3.34^2 + 2.79^2 + 4.48^2}$$

This results in a confidence interval spanning -4.4 and +20.5. The fact that the 95 per cent confidence interval crosses 0 confirms that the change in mean PISA science scores in Northern Ireland between 2006 and 2015 does not reach statistical significance at the five per cent level.

Appendix D. The conversion of PISA scores into years of schooling

8. The OECD has previously equated 40 PISA points into one year of additional schooling (OECD 2010:110). This was based upon an analysis investigating how PISA scores vary between pupils in different school year groups. The OECD has reviewed the evidence for the conversion between PISA points and years of schooling as part of the PISA 2015 international report (Box I.2.1). They point to the following studies in particular:

- Prenzel et al. (2006), who conducted a follow-up of the PISA 2003 cohort in Germany one year after taking the PISA test. Over this year, pupils gained about 25 score points in PISA mathematics and 21 points in science.
- OECD (2012), where the PISA 2000 cohort in Canada were re-tested at age 24. The average reading score increased by 57 points, from 541 to 598, over this nine year period.
- Keskpaik and Salles (2013), who compared PISA scores of eighth and ninth grade pupils in France. They found a score point difference of 44 points over the year of schooling, though this is recognised to be an upper-bound.
- Woessmann (2016), who states that learning gains on most national and international assessments during one year is equal to between a quarter and a third of a standard deviation.

9. Based upon this evidence, the OECD have revised their guidance, and now equate 30 PISA test points to a year of additional schooling. However, they note that this must be understood as an approximate rule of thumb, and that variation across subjects and across different countries may occur. To illustrate this point, Anders et al. (2016) highlight how PISA scores in Shanghai and Taiwan increase by very little (typically by less than 10 test points) over one particular academic year.

Appendix E. The PISA proficiency levels

Appendix Table E1. The PISA science proficiency levels

Level	Description of the science proficiency levels
Level 6	At Level 6, students are able to use content, procedural and epistemic knowledge to consistently provide explanations, evaluate and design scientific enquiries and interpret data in a variety of complex life situations that require a high level of cognitive demand. They can draw appropriate inferences from a range of different complex data sources, in a variety of contexts and provide explanations of multi-step causal relationships. They can consistently distinguish scientific and non-scientific questions, explain the purposes of enquiry, and control relevant variables in a given scientific enquiry or any experimental design of their own. They can transform data representations, interpret complex data and demonstrate an ability to make appropriate judgments about the reliability and accuracy of any scientific claims. Level 6 students consistently demonstrate advanced scientific thinking and reasoning requiring the use of models and abstract ideas and use such reasoning in unfamiliar and complex situations. They can develop arguments to critique and evaluate explanations, models, interpretations of data and proposed experimental designs in a range of personal, local and global contexts.
Level 5	At Level 5, students are able to use content, procedural and epistemic knowledge to provide explanations, evaluate and design scientific enquiries and interpret data in a variety of life situations in some but not all cases of high cognitive demand. They draw inferences from complex data sources, in a variety of contexts and can explain some multi-step causal relationships. Generally, they can distinguish scientific and non-scientific questions, explain the purposes of enquiry, and control relevant variables in a given scientific enquiry or any experimental design of their own. They can transform some data representations, interpret complex data and demonstrate an ability to make appropriate judgments about the reliability and accuracy of any scientific claims. Level 5 students show evidence of advanced scientific thinking and reasoning requiring the use of models and abstract ideas and use such reasoning in unfamiliar and complex situations. They can develop arguments to critique and evaluate explanations, models, interpretations of data and proposed experimental designs in some but not all personal, local and global contexts.
Level 4	At Level 4, students are able to use content, procedural and epistemic knowledge to provide explanations, evaluate and design scientific enquiries and interpret data in a variety of given life situations that require mostly a medium level of cognitive demand. They can draw inferences from different data sources, in a variety of contexts and can explain causal relationships. They can distinguish scientific and non-scientific questions, and control variables in some but not all scientific enquiry or in an experimental design of their own. They can transform and interpret data and have some understanding about the confidence held about any scientific claims. Level 4 students show evidence of linked scientific thinking and reasoning and can apply this to unfamiliar situations. Students can also develop simple arguments to question and critically analyse explanations, models, interpretations of data and proposed experimental designs in some personal, local and global contexts.
Level 3	At Level 3, students are able to use content, procedural and epistemic knowledge to provide explanations, evaluate and design scientific enquiries and interpret data in some given life situations that require at most a medium level of cognitive demand. They are able to draw a few inferences from different data sources, in a variety of contexts, and can describe and partially explain simple causal relationships. They can distinguish some scientific and non-scientific questions, and control some variables in a given scientific enquiry or in an experimental design of their own. They can transform and interpret simple data and are able to comment on the confidence of scientific claims. Level 3 students show some evidence of linked scientific thinking and reasoning, usually applied to familiar situations. Students can develop partial arguments to question and critically analyse explanations, models, interpretations of data and proposed experimental designs in some personal, local and global contexts.

Level 2	At Level 2, students are able to use content, procedural and epistemic knowledge to provide explanations, evaluate and design scientific enquiries and interpret data in some given familiar life situations that require mostly a low level of cognitive demand. They are able to make a few inferences from different sources of data, in few contexts, and can describe simple causal relationships. They can distinguish some simple scientific and non-scientific questions, and distinguish between independent and dependent variables in a given scientific enquiry or in a simple experimental design of their own. They can transform and describe simple data, identify straightforward errors, and make some valid comments on the trustworthiness of scientific claims. Students can develop partial arguments to question and comment on the merits of competing explanations, interpretations of data and proposed experimental designs in some personal, local and global contexts.
Level 1a	At Level 1a, students are able to use a little content, procedural and epistemic knowledge to provide explanations, evaluate and design scientific enquiries and interpret data in a few familiar life situations that require a low level of cognitive demand. They are able to use a few simple sources of data, in a few contexts and can describe some very simple causal relationships. They can distinguish some simple scientific and non-scientific questions, and identify the independent variable in a given scientific enquiry or in a simple experimental design of their own. They can partially transform and describe simple data and apply them directly to a few familiar situations. Students can comment on the merits of competing explanations, interpretations of data and proposed experimental designs in some very familiar personal, local and global contexts.
Level 1b	At Level 1b, students demonstrate a little evidence to use content, procedural and epistemic knowledge to provide explanations, evaluate and design scientific enquiries and interpret data in a few familiar life situations that require a low level of cognitive demand. They are able to identify straightforward patterns in simple sources of data in a few familiar contexts and can offer attempts at describing simple causal relationships. They can identify the independent variable in a given scientific enquiry or in a simple design of their own. They attempt to transform and describe simple data and apply them directly to a few familiar situations.

Appendix Table E2. The PISA mathematics proficiency levels

Level	Description of the mathematics proficiency levels
Level 6	At Level 6, pupils can conceptualise, generalise and utilise information based on their investigations and modelling of complex problem situations, and can use their knowledge in relatively non-standard contexts. They can link different information sources and representations and flexibly translate among them. Pupils at this level are capable of advanced mathematical thinking and reasoning. These pupils can apply this insight and understanding, along with a mastery of symbolic and formal mathematical operations and relationships, to develop new approaches and strategies for attacking novel situations. pupils at this level can reflect on their actions, and can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situation
Level 5	At Level 5 pupils can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare, and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Pupils at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They begin to reflect on their work and can formulate and communicate their interpretations and reasoning.
Level 4	At Level 4 pupils can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects of real-world situations. Pupils at this level can utilise their limited range of skills and can reason with some insight, in straightforward contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments, and actions.

Level 3	At Level 3 pupils can execute clearly described procedures, including those that require sequential decisions. Their interpretations are sufficiently sound to be a base for building a simple model or for selecting and applying simple problem-solving strategies. Pupils at this level can interpret and use representations based on different information sources and reason directly from them. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their solutions reflect that they have engaged in basic interpretation and reasoning
Level 2	At Level 2 pupils can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Pupils at this level can employ basic algorithms, formulae, procedures, or conventions to solve problems involving whole numbers. They are capable of making literal interpretations of the results.
Level 1	At Level 1 pupils can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are almost always obvious and follow immediately from the given stimuli.

Appendix Table E3. The PISA reading proficiency levels

Level	Description of the reading proficiency levels
Level 6	Tasks at this level typically require the reader to make multiple inferences, comparisons and contrasts that are both detailed and precise. They require demonstration of a full and detailed understanding of one or more texts and may involve integrating information from more than one text. Tasks may require the reader to deal with unfamiliar ideas, in the presence of prominent competing information, and to generate abstract categories for interpretations. Reflect and evaluate tasks may require the reader to hypothesise about or critically evaluate a complex text on an unfamiliar topic, taking into account multiple criteria or perspectives, and applying sophisticated understandings from beyond the text. A salient condition for access and retrieve tasks at this level is precision of analysis and fine attention to detail that is inconspicuous in the texts.
Level 5	Tasks at this level that involve retrieving information require the reader to locate and organise several pieces of deeply embedded information, inferring which information in the text is relevant. Reflective tasks require critical evaluation or hypothesis, drawing on specialised knowledge.
Level 4	Tasks at this level that involve retrieving information require the reader to locate and organise several pieces of embedded information. Some tasks at this level require interpreting the meaning of nuances of language in a section of text by taking into account the text as a whole. Other interpretative tasks require understanding and applying categories in an unfamiliar context. Reflective tasks at this level require readers to use formal or public knowledge to hypothesise about or critically evaluate a text. Readers must demonstrate an accurate understanding of long or complex texts whose content or form may be unfamiliar.
Level 3	Tasks at this level require the reader to locate, and in some cases recognise the relationship between, several pieces of information that must meet multiple conditions. Interpretative tasks at this level require the reader to integrate several parts of a text in order to identify a main idea, understand a relationship or construe the meaning of a word or phrase. They need to take into account many features in comparing, contrasting or categorising. Often the required information is not prominent or there is much competing information; or there are other text obstacles, such as ideas that are contrary to expectation or negatively worded. Reflective tasks at this level may require connections, comparisons, and explanations, or they may require the reader to evaluate a feature of the text. Some reflective tasks require readers to demonstrate a fine understanding of the text in relation to familiar, everyday knowledge. Other tasks do not require detailed text comprehension but require the reader to draw on less common knowledge.

Level 2	Some tasks at this level require the reader to locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. Others require recognising the main idea in a text, understanding relationships, or construing meaning within a limited part of the text when the information is not prominent and the reader must make low-level inferences. Tasks at this level may involve comparisons or contrasts based on a single feature in the text. Typical reflective tasks at this level require readers to make a comparison or several connections between the text and outside knowledge, by drawing on personal experience and attitudes.
Level 1a	Tasks at this level require the reader to locate one or more independent pieces of explicitly stated information; to recognise the main theme or author's purpose in a text about a familiar topic, or to make a simple connection between information in the text and common, everyday knowledge. Typically the required information in the text is prominent and there is little, if any, competing information. The reader is explicitly directed to consider relevant factors in the task and in the text.
Level 1b	Tasks at this level require the reader to locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. The text typically provides support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. In tasks requiring interpretation the reader may need to make simple connections between adjacent pieces of information.

Appendix F. Revisions made to PISA 2012 scores

1. Due to an error in the layout of the Welsh language version of the PISA 2012 pupil questionnaire, some of the information on pupil gender within the Wales sample in the PISA 2012 international database for the United Kingdom is incorrect. The error was not large enough to have a detectable impact on the UK's PISA 2012 results. However, it does have a small impact on estimates of overall scores and gender differences for Wales, Northern Ireland and England as pupil characteristics (including gender) are used in the calculations of estimated performance scores for individual pupils.

2. The tables which follow provide the mean score, variation and gender differences in mathematics, science and reading, for England, Northern Ireland and Wales, based on the corrected data. The data for Scotland is not affected by this revision as data for Scotland was collected, coded and analysed separately.

3. Appendix Table F1 compares the original scale scores at the time of PISA 2012 publication (December 2013) to the revised scores published in May 2015[100]. As the tables illustrate, in all three countries, the impact upon mean scores, percentiles and gender differences was minimal; estimates of most of these statistics differed by around one scale score point or less. None of the key substantive findings therefore changed as a result of this anomaly.

4. For consistency with previously published information, and the fact the rescaling led to minimal changes, we have chosen to present results based upon the original scale scores throughout this report.

[100] See http://www.oecd.org/pisa/keyfindings/PISA-2012-UK-revised%20scores.xlsx

Appendix Table F1. A comparison of the original and revised PISA 2012 scale scores across England, Northern Ireland and Wales

(a) England

	Science		Mathematics		Reading	
	Original	Revised	Original	Revised	Original	Revised
Mean	515.8	515.8	495.2	495.7	499.9	499.8
10th percentile	384.3	384.3	370.5	371.9	370.7	372.1
25th percentile	449.1	449.1	429.8	430.8	438.2	437.7
75th percentile	587.1	587.1	562.2	562.5	568.2	568.7
90th percentile	641.7	641.7	618.5	619.5	621.3	622.7
Results by gender						
Mean boys	522.9	522.9	501.7	502.5	487.3	487.7
Mean girls	509.0	509.0	489.0	489.2	511.8	511.3
Gender gap	13.8	13.8	12.7	13.3	-24.5	-23.6

(b) Northern Ireland

	Science		Mathematics		Reading	
	Original	Revised	Original	Revised	Original	Revised
Mean	507.2	507.2	486.9	486.9	497.6	498.0
10th percentile	374.7	374.7	365.3	364.4	373.4	373.8
25th percentile	438.1	438.1	421.8	421.1	435.8	436.9
75th percentile	577.9	577.9	552.9	550.7	565.4	564.5
90th percentile	635.2	635.2	608.5	607.8	617.6	618.6
Results by gender						
Mean boys	509.8	509.8	491.8	491.4	484.5	484.5
Mean girls	504.4	504.4	481.5	482.0	511.9	512.6
Gender gap	5.4	5.4	10.3	9.4	-27.4	-28.1

(c) Wales

	Science		Mathematics		Reading	
	Original	Revised	Original	Revised	Original	Revised
Mean	490.9	490.9	468.4	468.7	479.7	479.7
10th percentile	370.1	370.1	359.7	359.9	364.6	363.5
25th percentile	428.1	428.1	409.8	411.9	420.7	421.1
75th percentile	556.3	556.3	526.4	526.1	541.5	541.7
90th percentile	609.2	609.2	577.6	577.2	592.8	593.3
Results by gender						
Mean boys	496.2	496.2	473.0	473.9	466.4	465.4
Mean girls	485.5	485.5	463.7	463.6	493.1	493.6
Gender gap	10.7	10.7	9.3	10.3	-26.7	-28.2

Source: http://www.oecd.org/pisa/keyfindings/PISA-2012-UK-revised%20scores.xlsx

Note: Original refers to the initial scale scores before correction, as published in December 2013. Revised refers to the scale scores after correction, published in March 2015.

The views expressed in this report are the authors' and do not necessarily reflect those of the Welsh Government.

Global Education Monitoring Report

Policy Paper 24

February 2016

HR 01b
Ymchwiliad i hawliau dynol yng Nghymru
Inquiry into Human Rights in Wales
Ymateb gan: J. Jones
Response from: J. Jones

If you don't understand, how can you learn?

Key Messages:

1. Children should be taught in a language they understand, yet as much as 40% of the global population does not have access to education in a language they speak or understand.
2. Speaking a language that is not spoken in the classroom frequently holds back a child's learning, especially for those living in poverty.
3. At least six years of mother tongue instruction is needed to reduce learning gaps for minority language speakers.
4. In multi-ethnic societies, imposing a dominant language through a school system has frequently been a source of grievance linked to wider issues of social and cultural inequality.
5. Education policies should recognize the importance of mother tongue learning.
6. Linguistic diversity creates challenges within the education system, notably in areas of teacher recruitment, curriculum development and the provision of teaching materials.

Quality education should be delivered in the language spoken at home. However, this minimum standard is not met for hundreds of millions, limiting their ability to develop foundations for learning. By one estimate, as much as 40% of the global population does not have access to an education in a language they speak or understand (Walter and Benson, 2012).[1] The challenges are most prevalent in regions where linguistic diversity is greatest such as in sub-Saharan Africa and Asia and the Pacific (UNDP, 2004).

1. According to an earlier study, around 221 million children are estimated to speak a different language at home from the language of instruction in school (Dutcher, 2004).

Poverty and gender magnify educational disadvantages linked to ethnicity and language. With a new global education agenda that prioritizes equity and lifelong learning for all, the policy of respecting language rights is essential and deserves close attention.

This policy paper, released for International Mother Language Day, argues that being taught in a language other than their own can negatively impact children's learning. It shows the importance of teacher training and inclusive supporting materials to improve the learning experience of these children, and provide them with a resilient path of achievement in life.

UNESCO
United Nations Educational, Scientific and Cultural Organization

To be taught in a language other than one's own has a negative effect on learning

In many countries, large numbers of children are taught and take tests in languages that they do not speak at home, hindering the early acquisition of critically important reading and writing skills. Their parents may lack literacy skills or familiarity with official languages used in school, which can then reinforce gaps in learning opportunities between minority and majority language groups.

International and regional learning assessments confirm that when home and school languages differ there is an adverse impact on test scores. The Global Education Monitoring Report's World Inequality Database on Education (WIDE) also shows the extent of learning inequalities within countries, depending on whether children speak the language of assessment at home or not.

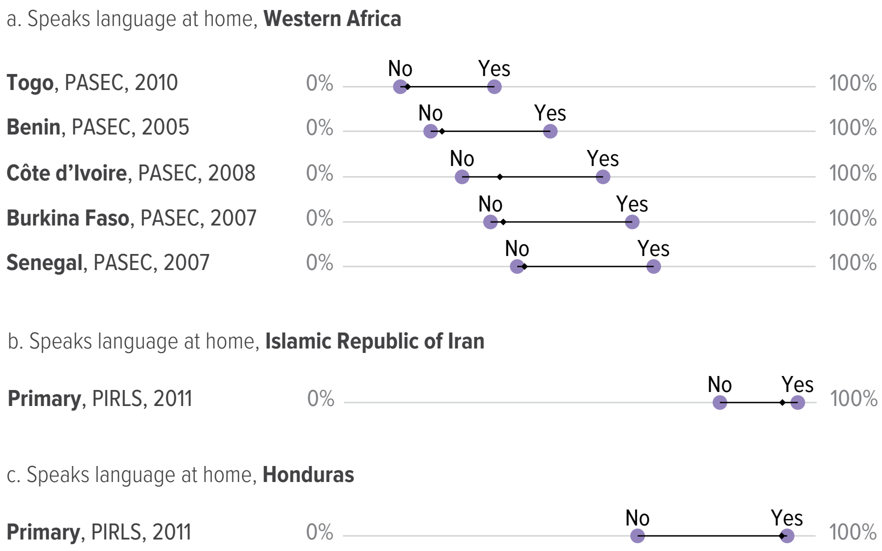
In many western African school systems, French continues to be the main language of instruction, so the vast majority of children are taught from the early grades in a language with which they have limited familiarity. This seriously hampers their chances of learning. In Côte d'Ivoire, 55% of grade 5 students who speak the test language at home learned the basics in reading in 2008, compared with only 25% of the 8 out of 10 students who speak another language (**Figure 1a**).

In the Islamic Republic of Iran, around 20% of grade 4 students taking the test in Farsi, the official language of instruction, reported speaking a different language at home. Of these, 80% reached the basics in reading, compared with over 95% of Farsi speakers (**Figure 1b**).

Similarly, in Honduras, in 2011, 94% of grade 6 students who spoke the language of assessment at home learned the basics in reading in primary school compared to only 62% of those who did not (**Figure 1c**).

Language and ethnicity can combine to produce complex patterns of compounded disadvantage. In Peru, the difference in test scores between indigenous and non-indigenous children in grade 2 is sizeable and increasing. In 2011, Spanish speakers were more than seven times as likely as indigenous language speakers to reach a satisfactory standard in reading (Guadalupe et al., 2013).

Figure 1: When home and school languages differ, there is an adverse impact on test scores
Percentage of children taking part in an assessment who achieved an international minimum learning standard in reading.



Source: World Inequality Database on Education (WIDE).

The legacy of marginalization facing indigenous communities in high income countries, though clearly visible in student assessments, has received minimal attention in international education debates. According to an analysis of TIMSS data, in Australia, approximately two-thirds of indigenous students achieved the minimum benchmark in mathematics in grade 8 between 1994/1995 and 2011, as compared with almost 90% of their non-indigenous peers (Thomson et al., 2012).

Language, ethnicity and poverty can interact to produce an extremely high risk of being left far behind. Students from poor households who speak a minority language at home are among the lowest performers. In 2012, poor 15-year-olds in Turkey speaking a non-Turkish language, predominantly Kurdish, were among the lowest performers. Around 50% of poor non-Turkish speakers achieved minimum learning benchmarks in reading, against the national average of 80%.

In 2006, of poor students in Guatemala speaking a minority language (mostly indigenous) at home, only 38% learned the basics in mathematics while 77% of rich students who speak Spanish reached that level.

Using the home language as language of instruction has a positive impact on learning across the board

"Ensure that every institution is secure and has water, electricity, gender-segregated toilets that work and are accessible, adequate and safe classrooms, and appropriate learning materials and technology."
—Natalee, teacher, Bay Islands, Honduras

To ensure that children acquire strong foundation skills in literacy and numeracy, schools need to teach the curriculum in a language children understand. Mother tongue based bilingual (or multilingual) education approaches, in which a child's mother tongue is taught alongside the introduction of a second language, can improve performance in the second language as well as in other subjects.


Credit: Nguyen Thanh Tuan/UNESCO

Nguyen, a teacher in Muong Khuong county, Viet Nam: 'There are 13 ethnic students in my class. All Hmong girls. Sometimes when you teach in Vietnamese they seem not to understand.'

In Guatemala, students in bilingual schools have higher attendance and promotion rates and lower repetition and dropout rates. Moreover, they have higher scores on all subject matters, including the mastery of Spanish. A shift to bilingual schooling would result in considerable cost-saving as a result of reduced repetition (Patrinos and Velez, 1996).

In Mali's Pédagogie Convergente programme, children starting school with mother tongue instruction ended up with better mastery of the official language, French. Between 1994 and 2000, children who began their schooling in the language they spoke at home scored 32% higher on French proficiency tests at the end of primary school than those in French-only programmes (Bühmann and Trudell, 2008). In Burkina Faso, results of tests in French carried out with children from bilingual schools in 2005 were comparable to or higher than those of children in traditional French-instruction schools (Nikiema, 2011).

The benefits of bilingual/multilingual programmes extend beyond cognitive skills to enhanced self-confidence and self-esteem. In Burkina Faso, mother tongue instruction facilitated the use of effective teaching practices in the classroom and encouraged learners to be active and become involved with the subject matter (Nikiema, 2011).

Earlier research and common practice in the past have supported earlier transitions from home to official language. However, recent evidence now claims that at least six years of mother tongue instruction – increasing to eight years in less well-resourced conditions – is needed to sustain improved learning in later grades for minority language speakers and reduce learning gaps (Heugh et al., 2007; UNESCO, 2011). Yet many countries in sub-Saharan Africa that support bilingual education continue to favour early transition to the official language, usually by grade 4 (Alidou et al., 2006).

In Cameroon, between 2007 and 2012, 12 schools in the Boyo division of the Northwest region, used a curriculum in the local language, Kom, for grades 1 to 3, switching to English instruction in grade 4. Children taught in Kom showed a marked advantage in achievement in reading and comprehension compared with children taught only in English. Kom-educated children also scored twice as high on mathematics tests at the end of grade 3. Yet, these learning gains were not sustained when the students switched to English-only instruction in grade 4. The early exit from a mother tongue environment prevented them from sustaining this performance across the curriculum (Walter and Chuo, 2012).

The Six Year Primary Project in Ife, Nigeria, used Yoruba as the medium of instruction for the six years of primary education. Evaluations of the project found that students who switched to English after six years of mother tongue instruction performed better in English and in other subjects compared with those who did so after only three years (Bamgbose, 2000, 2004; Fafunwa et al., 1989).

Ethiopia has gone further than many countries, seeking to combine mother tongue instruction with Amharic and English in grades 1 to 8. Children's participation in bilingual programmes for eight years improved their learning in subjects across the curriculum and not just in the language of instruction. Primary school children learning in their mother tongue performed better in grade 8 in mathematics, biology, chemistry and physics than pupils in English-only schooling (Heugh et al., 2007).

Language in education policies can be a source of wider grievances

Language both reflects the culture of one's community as well as an individual's ethnic identity. The language(s) one learns and speaks often create a sense of personal identity and group attachment. Yet, language can serve as a double-edged sword: while it strengthens an ethnic group's sense of belonging and social ties, it can also turn into a basis for their marginalization.

By one estimate conducted in 2009, over half of the countries affected by armed conflict are highly diverse linguistically, making decisions over the language of instruction a potentially divisive political issue (Pinnock, 2009). Whatever the underlying complexities and political dynamics, the following cases highlight the ways in which language policy in education has emerged as a focal point for violent conflict.

In multi-ethnic countries, for example, the imposition of a single dominant language as the language of instruction in schools, while sometimes a choice of necessity, has been a frequent source of grievance linked to wider issues of social and cultural inequality. The fault lines of violent conflict have often followed the contours of group-based inequality exacerbated by language policies in education. For example, disputes about using Kurdish in schools have been an integral part of the conflict in eastern Turkey (Graham-Brown, 1994; UNESCO, 2010). In Nepal, the imposition of Nepali as the language of instruction fed into the broader set of grievances among non-Nepali speaking castes and ethnic minorities that drove the civil war (Murshed and Gates, 2005). Guatemala's imposition of Spanish on schools was seen by indigenous people as part of a broader pattern of social discrimination. Armed groups representing indigenous people included the demand for bilingual and intercultural education in their conditions for a political settlement, and the country's peace agreement included a constitutional commitment to that end (Marques and Bannon, 2003).

Disputes over language often reflect long stories of domination, subordination and, in some cases, decolonization. In Algeria, the replacement of French by Arabic in primary and secondary schools after independence in 1962 was intended to build the new government's legitimacy but marginalized the non-Arabic-speaking Berber minority (Brown, 2010).

In Pakistan, the post-independence government adopted Urdu as the national language and the language of instruction in schools. This became a source of alienation in a country that was home to six major linguistic groups and fifty-eight smaller ones. The failure to recognize Bengali, spoken by the vast majority of the population in East Pakistan, was one of the major sources of conflict within the new country, leading to student riots in 1952. The riots gave birth to the Bengali Language Movement, a precursor to the movement that fought for the secession of East Pakistan and the creation of a new country, Bangladesh (Schendel, 2009).

Both countries have continued to face language-related political challenges. In Bangladesh, where Bengali is the national language, non-Bengali speaking tribal groups in the Chittagong Hill Tracts have cited a perceived injustice over language as a factor that justifies their secession demands (Mohsin, 2003). In Pakistan, the continued use of Urdu as the language of instruction in government schools, even though it is spoken at home by less than 8% of the population, has also contributed to political tensions (Ayres, 2003; Winthrop and Graff, 2010).

National education policies should recognise the importance of teaching children in their home language

Education policies seldom reflect linguistic diversity. A review of 40 countries' education plans finds that only less than half of them recognize the importance of teaching children in their home language, particularly in early grades. Cambodia and the Lao People's Democratic Republic are positive examples, encouraging the recruitment of teachers with specific language skills. Namibia encourages the production of learning materials in minority languages (Hunt, 2013).

In sub-Saharan Africa, there has been a general trend towards more widespread use of local languages. At the time of independence, only 20 out of 47 used local languages in primary education, whereas 38 now do so, largely influenced by advocacy from local actors (Albaugh, 2015).

Several Latin American countries – for example, Colombia, Guyana, Paraguay and Peru – go further than others in identifying reforms to improve the learning of disadvantaged groups, notably ethnolinguistic minorities and the poor. While such reforms mainly focus on extending access, they also include adapting curricula and pedagogical practices to the needs of particular groups. In Paraguay, this involves creating educational materials in various languages (Paraguay Ministry of Education, 2011).

Good teachers are critical for helping the most disadvantaged learners

"Without adequate knowledge in English, teachers are unable to interact with the students, and the result is a strict chalk-and-talk structure."
—Inga, teacher, Kigali, Rwanda

For mother tongue based bilingual (or multilingual) education approaches to be effective, governments need to recruit teachers from minority language groups. Language of instruction policies may be difficult to implement though, particularly when there is more than one language group in the same classroom and teachers are not proficient in one or several of the local languages (Alidou and Brock-Utne, 2011).

Children who speak minority languages not taught in the classroom often enter school with low self-esteem and learning needs that teachers may feel unable to meet. Schools can play an important role in changing this situation. Hiring teachers from minority language communities can help widen children's horizons and raise their ambitions.

Yet, because ethnic and language minorities often obtain less formal education than majority groups, fewer members of the former are available and qualified for recruitment as teachers. In India, all states have a caste-based reservation of posts to ensure that teachers are available in more disadvantaged areas and schools, but teachers with lower levels of qualifications are hired to fill the reserved positions. There are not enough teachers who speak local languages, and very few bilingual teachers belong to minorities, which compounds the disadvantage children face when their home language is not the medium of instruction (Chudgar and Luschei, 2013). In Mexico, teachers whose mother tongue is an indigenous language often have less education and training than other teachers (Luschei et al., 2013).

Teachers need to be trained to teach in two languages and to understand the needs of second-language learners.

Teachers are rarely prepared for the reality of bilingual or multilingual classrooms. In Senegal, where attempts are being made to use local languages in schools, training is given only in French, and a survey found that only 8% of trainees expressed any confidence about teaching reading in local languages. In Mali, this was the case for just 2% of teachers (Akyeampong et al., 2013). A small-scale study of mathematics teaching in Botswana indicated that bilingual teacher education was failing in its aim of preparing teachers for multilingual classrooms where pupils' home language may be different from both the national language and English, the medium of mathematics teaching (Kasule and Mapolelo, 2005).

In Peru, bilingual programmes aim to ensure that children can learn in their own language together with Spanish. Yet, children attending these programmes perform badly in both languages. By grade 4, only 1 in 10 Quechua speakers in bilingual programmes, and 1 in 20 speakers of other indigenous languages, reach a satisfactory level in their own language. Their achievement in Spanish is similarly weak. This highlights the importance not only of providing instruction in a child's own language, but also of ensuring that schools are of sufficient quality to ensure that learning takes place. A study found that half of teachers in bilingual education schools in southern Peru could not speak the local indigenous language (Cueto et al., 2012; Guadalupe et al., 2013).

Teacher education programmes need to support teachers to be able to teach early reading skills in more than one language and to use local language materials effectively. Teachers should have a good understanding of the linguistic and cultural backgrounds of children, language development, and the interdependence of mother tongue and second-language development, and the use of appropriate teaching practices (Pinnock and Nicholls, 2012).

One reason Ecuador has been able to deliver strong bilingual teaching is that it has established five specialized teacher-training colleges. Similarly, the Plurinational State of Bolivia has created three indigenous language universities to support bilingual training (Lopez, 2010).

In Australia, where teachers often mistake a language problem for a learning difficulty, the Deadly Ways to Learn project is an attempt to change how teachers view Aboriginal languages. Set up in 14 government, private and Catholic schools across rural and urban Western Australia, the project included the preparation of books to introduce teachers to the culture, identity and history that inform Aboriginal language. Indigenous Education Officers provide support and guidance to teachers in the selected schools. The project highlights the importance of all students receiving an education that is sensitive to the history, culture and language of indigenous Australians, and that also takes into account the backgrounds of people from other minority groups (Biddle and Mackay, 2009).

The effectiveness of mother tongue based bilingual education depends on inclusive learning materials

For early grade literacy and bilingual education to be successful, pupils need inclusive and relevant learning materials in a language they are familiar with. Textbooks, when available, are much less useful if learners have difficulty reading them, as was demonstrated in an experiment supplying textbooks written in

English to Kenyan classrooms. Test scores rose only among those who were already high achievers. Many pupils could not read the books, which were suited to academically strong pupils with educated parents. As a result, low achievers, mainly from poor and disadvantaged backgrounds, did not benefit from the greater access to textbooks (Glewwe et al., 2009).

Open licence educational resources and new technology can make learning materials more widely available, including in local languages. In South Africa, the Breadbin Interactive Project provides a cost-effective way of disseminating large quantities of open license digital content from a hard drive via digital dispensers. These can be connected directly to schools' computer systems or made available through electronic kiosks where materials are printed as required. Schools do not need internet connections to access the digital material. Further strategies are needed to support distribution to remote rural schools, however, as they rarely have even the most basic infrastructure for information technology. Outside the classroom, the Nal'ibali initiative supports volunteer-run reading clubs by producing colourful bilingual supplements with stories, ideas for literacy activities, and reading tips, available in English, and Xhosa or Zulu (Butcher, 2011; Nal'ibali, 2013; Project for the Study of Alternative Education in South Africa, 2013; Welch, 2012).

Teachers need to be supported with appropriate assessment strategies

To improve learning for all children, teachers need the support of assessment strategies that can reduce disparities in school achievement and offer all children and young people the opportunity to acquire vital transferable skills.

Diagnostic and formative assessments tools are crucial to improve the quality of education and make it more equitable.

Diagnostic and formative assessments can provide reliable, timely and informative information about student mastery and group progress on subject content. Such assessments are helpful in diagnosing learning difficulties, especially among low achievers. Teachers need easy-to-use and reliable assessment tools that are clearly linked to instruction, to better assess how instruction and classroom dynamics can be altered to meet the needs of all learners, even those with few or no writing skills. Teacher training in the use of diagnostic and formative assessments is crucial, so that teachers can identify weak learners and provide them with targeted support.

Early Grade Reading Assessments (EGRA) are designed to be administered orally in local languages and are sensitive to the lower end of the achievement range, capable of detecting performance on emerging skills (Gove and Cvelich, 2011). In Liberia, the EGRA Plus Project, which trained teachers in the use of classroom-based assessment tools and provided reading resources and scripted lesson plans to guide instruction, made a substantial impact, raising previously low levels of reading achievement among grade 2 and 3 pupils. The project involved several types of continuous assessment. Teachers used a simple oral assessment outlined in scripted lessons plans to check pupils' understanding during reading instruction. This allowed teachers to quickly assess responses and identify pupils requiring further assistance. Teachers also applied regular curriculum-based measures to check individual pupils' progress and calculate class averages, reporting both pupil and class progress to parents a minimum of four times a year.

Colour-coded report cards allowed parents to visualize their child's progress easily throughout the year. In addition, periodic tests were built into the curriculum to check pupils' mastery of particular skills and determine instructional needs. One challenge was to ensure that teachers understood the importance of the data gathered from assessments and used the tools consistently to inform practice. To address this, trained mentors regularly visited schools to support teachers and ensure the quality of instruction and assessment (Davidson et al., 2011).

Quality early childhood education and support for early transitions are vital

Improving the provision of good quality education through early childhood learning centres not only increases children's success in making the transition to primary school, but also improves later achievement – particularly for children facing disadvantage.

It is often particularly difficult for children who are members of linguistic and ethnic minorities to gain access to high quality early childhood education that prepares them for primary school. Culturally appropriate school-readiness programmes provided to children as they make the transition into primary school can improve learning outcomes.

One such successful programme was implemented in Viet Nam under the Primary Education for Disadvantaged Children project. While Vietnamese, the language of the Kinh majority, is the medium of instruction in primary schools, it is not the mother tongue of the other 53 ethnic groups that constitute about 15% of the population. Children in remote, single ethnic minority communities who are taught by a Kinh teacher can have difficulties coping with the classroom environment, understanding the curriculum and retaining interest in school (Harris, 2009).

The Teaching Assistants and School Readiness programme began in 2006 and has reached over 100,000 children. Over 7,000 locally recruited bilingual teaching assistants in 32 provinces were deployed to support ethnic minority children from isolated communities as they made the transition into primary school. The assistants helped children prepare for school through early childhood education activities for two months prior to grade 1 entry and provided additional instruction once they were in school, including help with learning Vietnamese.

In a two-year study completed in 2009, grade 1 pupils that participated in school-readiness activities scored between 20% and 30% higher in reading and writing, as well as shape and number identification, than children in schools not participating in the programme. In addition, parents were happier to send their children to school, knowing they would have someone who understood their language and culture. As a result, head teachers reported increased enrolment and attendance (Harris, 2009; Primary Education for Disadvantaged Children, 2010).

Second-chance accelerated learning programmes in local languages enable the disadvantaged to catch up

Where schools fail to deliver good quality education, children are more likely to drop out early. Second-chance programmes, if well-designed, can teach foundation skills through an accelerated learning cycle. Such programmes can effectively raise the achievements of disadvantaged groups and linguistic minorities.

Accelerated learning programmes are typically delivered in non-formal settings and target disadvantaged out-of-school children. The programmes often produce their own curriculum resources. The in-class timetable reflects children's and communities' realities, and trained teachers provide an inclusive atmosphere. In such programmes, teachers are generally recruited from surrounding communities, ensuring a common cultural and linguistic background and enhancing accountability to community members (Longden, 2013).

Accelerated learning programmes typically cover two or more grades of formal schooling in one year with the aim of raising participants' academic proficiency to a level that allows them to re-enter the formal system in the appropriate grade. The majority of such programmes focus on basic numeracy and literacy skills, taught in the local language, coupled with practical learning oriented to learners' lives (Longden, 2013).

The Complementary Basic Education programme in Malawi recruited young men and women under 35 with a secondary qualification who lived in or near the villages hosting the learning centres. Community leaders were closely involved in the selection process (Jere, 2012). In South Sudan, secondary school graduates are recruited from surrounding communities and provided with intensive initial teacher education and regular in-service training. Their use of the local language to clarify instruction is seen by learners as an important positive aspect of the accelerated learning programme (Østergaard, 2013).

Recommendations

1 Teach children in a language they understand. At least six years of mother tongue education should be provided in ethnically diverse communities to ensure those speaking a different language from the medium of instruction do not fall behind. Bilingual or multilingual education programmes should be offered to ease the transition to the teaching of the official languages.

2 Train teachers to teach in more than one language. To fully support the implementation of mother tongue based bilingual/multilingual education programmes, teachers should receive pre-service and ongoing teacher education to teach in more than one language.

3 Recruit diverse teachers. Policy-makers need to focus their attention on hiring and training teachers from linguistic and ethnic minorities, to serve in the schools of their own communities.

4 Provide inclusive teaching materials. Curricula need to address issues of inclusion to enhance the chances of students from marginalized backgrounds to learn effectively. Textbooks should be provided in a language children understand. Classroom-based assessment tools can help teachers identify, monitor and support learners at risk of low achievement.

5 Provide culturally appropriate school-readiness programmes. Locally recruited bilingual teaching assistants can support ethnic minority children from isolated communities as they make the transition into primary school, including by providing additional instruction to them after they have enrolled.

Global Education Monitoring Report
c/o UNESCO
7, place de Fontenoy
75352 Paris 07 SP, France
Email: gemreport@unesco.org
Tel: +33 (1) 45 68 10 36
Fax: +33 (1) 45 68 56 41
www.unesco.org/gemreport

Developed by an independent team and published by UNESCO, the *Global Education Monitoring Report* is an authoritative reference that aims to inform, influence and sustain genuine commitment towards the global education targets in the new Sustainable Development Goals (SDGs) framework.

ED/GEMR/MRT/2016/PP/24

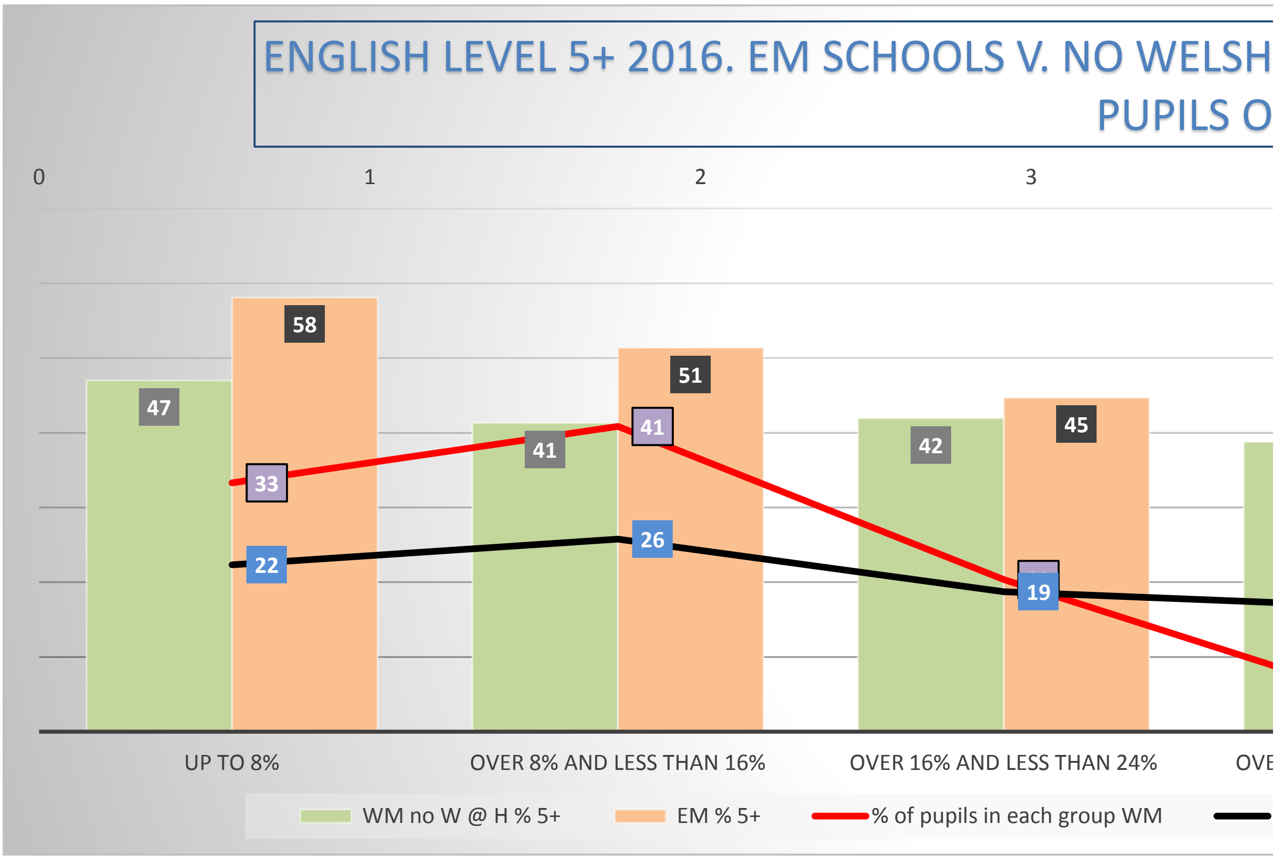
NOTES

1. References to this policy paper can be found online at the following link: https://en.unesco.org/gem-report/sites/gem-report/files/language_paper_references.pdf

HR 01c
Ymchwiliad i hawliau dynol yng Nghymru
Inquiry into Human Rights in Wales
Ymateb gan: J. Jones
Response from: J. Jones

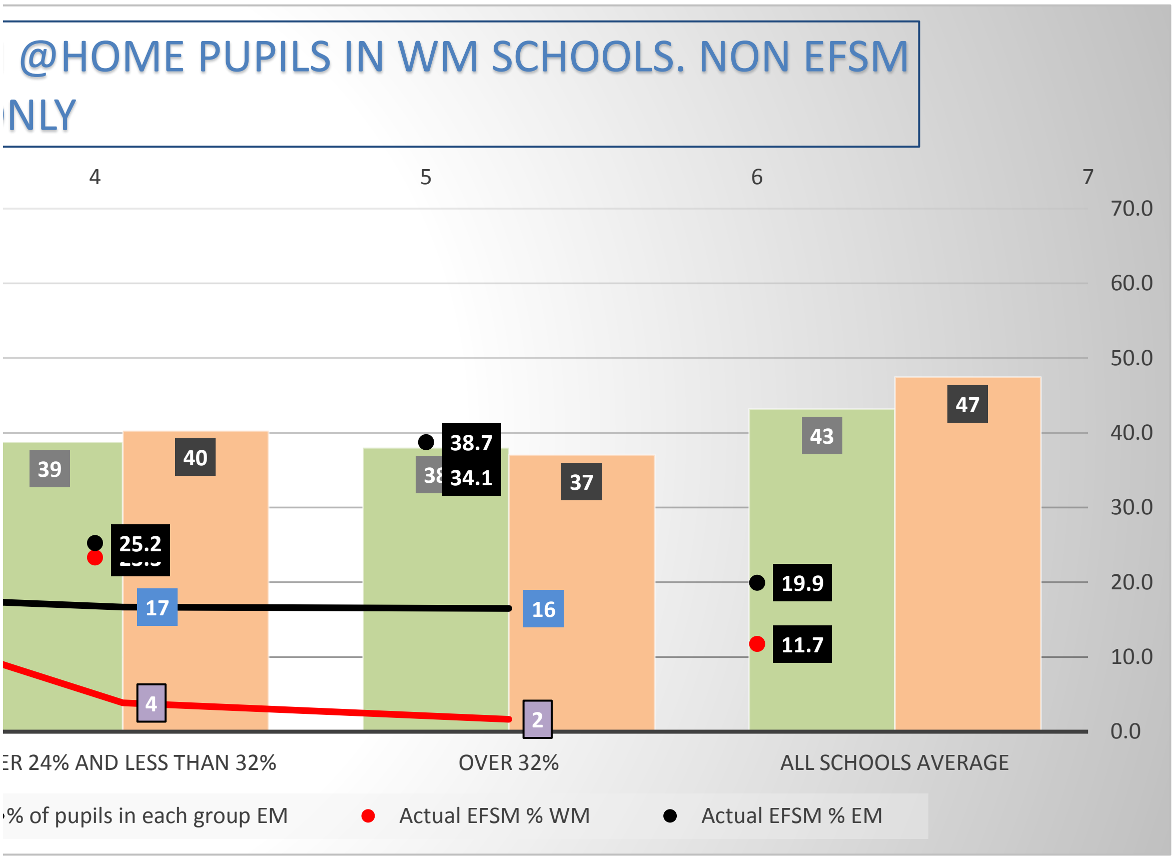
2016 KS2, COMPARISON WM PUPILS, NO WELSH @ HOME, VERSUS PUPILS IN I ENGLISH

	up to 8%	over 8% and less than 16%	over 16% and less than 24%	over 24% and less than 32%	over 32%
WM no W @ H % 5+	47	41	42	39	38
EM % 5+	58	51	45	40	37
Actual EFSM % WM				23.3	34.1
Actual EFSM % EM				25.2	38.7
% of pupils in each group WM	33	41	20	4	2
% of pupils in each group EM	22	26	19	17	16



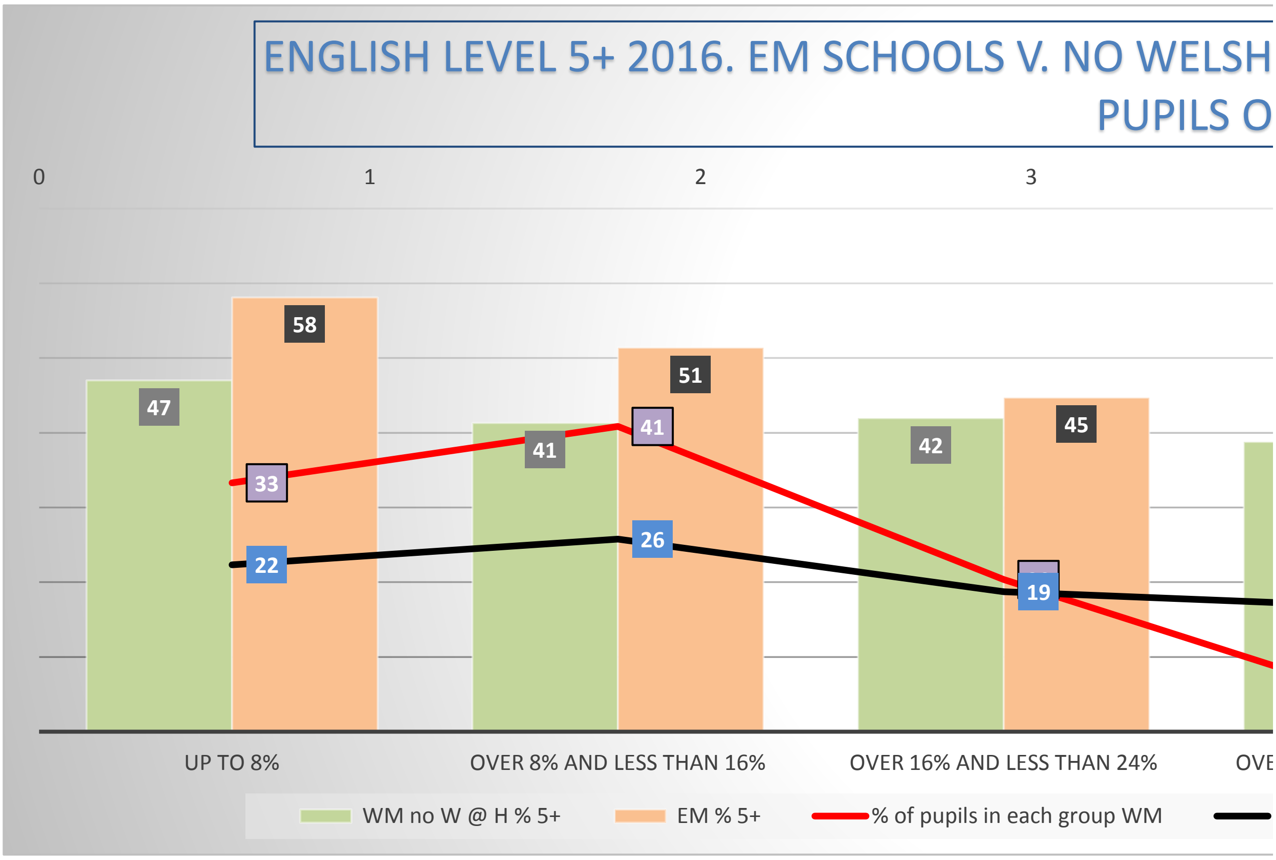
EM SCHOOLS BY SCHOOL EFSM GROUP. NON-EFSM PUPILS ONLY.

All schools average
43
47
11.7
19.9



2016 KS2, COMPARISON WM PUPILS, NO WELSH @ HOME, VERSUS PUPILS IN I ENGLISH

	up to 8%	over 8% and less than 16%	over 16% and less than 24%	over 24% and less than 32%	over 32%
WM no W @ H % 5+	47	41	42	39	38
EM % 5+	58	51	45	40	37
Actual EFSM % WM				23.3	34.1
Actual EFSM % EM				25.2	38.7
% of pupils in each group WM	33	41	20	4	2
% of pupils in each group EM	22	26	19	17	16



EM SCHOOLS BY SCHOOL EFSM GROUP. NON-EFSM PUPILS ONLY.

All schools average
43
47
11.7
19.9