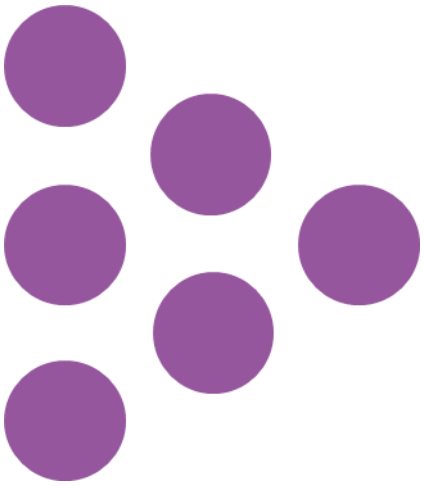
Llywodraeth Cymru
Welsh Government

PISA 2018 Additional Analyses:

Regional Performance and PISA/GCSE matching in Wales

National Foundation for Education Research (NFER)



PISA 2018 Additional Analyses: Regional Performance and PISA/GCSE matching in Wales

Geeta Gambhir, Afrah Dirie and Juliet Sizmur

Published in November 2020

By the National Foundation for Educational Research,

The Mere, Upton Park, Slough, Berkshire SL1 2DQ

www.nfer.ac.uk

ISBN: 978-1-912596-22-5

How to cite this publication:

Gambhir, G., Dirie, A. and Sizmur, J. (2020). *PISA 2018 Additional Analyses: Regional Performance and PISA/GCSE Matching in Wales*. Slough: NFER

Contents

List of tables

Executive Summary

The Programme for International Student Assessment (PISA) is a study of educational achievement run by the Organisation for Economic Co-operation and Development (OECD). Schools and pupils from 79 countries and economies participated in PISA 2018. The PISA study takes place every three years and examines the performance of 15 year old pupils in reading, mathematics and science. It also collects a range of contextual information on school and pupil backgrounds. Wales has participated in PISA since 2006.

This NFER study uses the 2018 PISA data to explore regional variation in PISA results, to compare PISA results with those of the General Certificate of Secondary Education (GCSE) and to examine differences in performance in English- and Welsh-medium schools.

Firstly we compared the four Welsh regions' PISA 2018 performance in the three core subjects of reading, mathematics and science and found no significant regional differences.

Next, we mapped PISA 2018 mean scores, proficiency levels and percentiles in the three core subjects (reading, mathematics and science) to best GCSE grades in GCSE literacy (English Language, Welsh Language, English literature or Welsh literature), GCSE mathematics and GCSE science. We found clear relationships, in that higher-achieving GCSE pupils also gained higher PISA mean scores and those with lower GCSE grades gained correspondingly lower mean scores in PISA in all subjects. Similar corresponding patterns were seen when comparing GCSE grades with proficiency levels and with percentiles across all subjects.

The third strand of these analyses compared PISA and GCSE performance in English- and Welsh-medium schools. We also compared the distribution of socioeconomic status (SES) by school medium and the language of PISA assessment. Pupils in English-medium schools significantly outperformed their Welsh-medium counterparts in PISA reading and mathematics, but not science. The largest mean score difference between English- and Welsh-medium was found in reading.

Some further analysis was undertaken to explore whether this difference in performance between the English- and Welsh-medium pupils who took PISA was reflected in their GCSE performance. It was not. In fact, pupils in Welsh-medium schools achieved more highly in their GCSEs than their English-medium PISA peers (when weighted by free school meals (FSM) eligibility and gender). These findings suggest that the large discrepancy in PISA scores was unlikely to be due to the Welsh-medium pupils being an unusually weak cohort, compared to the English-medium group. Nor were they unrepresentative of the Welsh-medium cohort; in fact both English- and Welsh-medium cohorts were higher achieving than their respective national cohorts.

We also examined the spread of pupils in English- and Welsh- medium schools by socioeconomic levels. Here we found that there were no significant differences in the proportions of pupils at each of the five PISA bands of OECD economic, social and cultural status (ESCS) when comparing the medium of the schools. The discrepancy in PISA scores cannot, therefore, be attributed to lower SES levels in Welsh-medium schools.

In reading, Welsh-medium pupils' PISA performance was lower than that of English-medium pupils in every SES quintile, from the most disadvantaged 20 per cent through to the most advantaged 20 per cent.

In mathematics and science, however, Welsh-medium pupils' PISA performance was lower than that of their English-medium peers only in the highest (most advantaged) SES quintile.

When comparing GCSE capped 9 scores in English- and Welsh-medium schools, no disparities were found at any SES level.

PISA scores differ considerably according to the language of assessment. Pupils who took the PISA assessment in Welsh had significantly lower scores in all three PISA subjects than those who took the assessment in English. However, the GCSE capped 9 average scores between these two groups was not significantly different.

The findings in this report show that:

- There were no significant differences in PISA scores between the four Regional Consortia in Wales.
- Patterns of performance are broadly similar when PISA scores are mapped against GCSE scores, with higher/lower PISA scores largely corresponding to higher/lower SES bands.
- Pupils in Welsh-medium schools gain lower PISA scores in all three subjects, significantly so in reading and mathematics, while GCSE scores and SES levels were higher among Welsh-medium pupils in the PISA sample.
- Pupils taking the PISA test in the Welsh language scored significantly lower than pupils who took the test in English, while GCSE performance showed no significant difference between the two groups.

In order to more fully investigate the reasons for the poorer performance of Welsh-medium pupils in the PISA assessments, further analysis would be required. This might include regression modelling to tease out the relative impact of school medium, home language and language of assessment; to explore performance in different categories of Welsh-medium schools, to consider the extent to which pupils who take the PISA test in Welsh

are effectively second language pupils; and whether the time taken to read the questions is longer for Welsh language questions and may, therefore, be a disadvantage in a timed assessment. Further information on any curricular differences in English- and Welsh-medium schools and any GCSE moderation comparisons may also shed light on the discrepancies noted.

1. Introduction

1.1 Background

PISA is the Programme for International Student Assessment (PISA). It is a study of educational achievement organised by the Organisation for Economic Co-operation and Development (OECD) and assesses the knowledge and skills of pupils aged 15. Pupils are assessed on their competence to address real-life challenges involving reading, mathematics and science.

Wales participates in PISA, alongside 78 other countries, to collect valuable data that enables governments to benchmark education policy and performance, to make evidence-based decisions and to learn from policies and practices in other countries.

The findings from PISA 2018 for Wales are available in *Achievement of 15-Year-Olds in Wales: PISA 2018 National Report* (Sizmur *et al.,* 2019).

In this report, we further analyse the PISA 2018 results of pupils in Wales. This analysis will partly replicate the study published by the Welsh Government in July 2017[1] following the PISA 2015 results, which examined regional performance in each of the three core subject areas (reading, mathematics and science) and mapped PISA scores to pupils' best General Certificate of Secondary Education (GCSE) grades.

Our analysis will:

- present the average (mean) PISA 2018 scores for reading, mathematics and science for each of the four Welsh regions
- map PISA 2018 reading mean scores and proficiency levels to pupils' best grade in GCSE literacy, which comprises the four literacy subjects: English language, Welsh language, English literature and Welsh literature; we also map PISA reading mean scores to best grade in English or Welsh literature for a direct comparison to the 2015 report. The greater emphasis on these subjects reflects the main domain of PISA 2018, which was reading
- map PISA mathematics mean scores and proficiency levels to pupils' best grade in GCSE mathematics (from GCSE mathematics and GCSE numeracy) and PISA science scores to their best GCSE science grade (including the single best grade from a double science award)[2]
- compare PISA performance by school medium (Welsh- and English-), pupils' SES and the language of the PISA assessment

[1] https://gov.wales/sites/default/files/publications/2018-03/Additional analysis of PISA 2015 results – regional performance and GCSE/BTEC

[2] Our analysis does not include BTEC science, as in the 2015 report, as the numbers of pupils taking this qualification have decreased substantially since 2017.

- compare GCSE achievement by school medium (Welsh- and English-), pupils' SES and the language of the PISA assessment.

The report is set out as follows:

- Section 2 - PISA results by Regional Consortia
- Section 3 - PISA 2018 scores by GCSE results
- Section 4 - PISA and GCSE performance in English- and Welsh-medium schools.

For further insight into Wales' PISA 2018 performance, it may be useful to refer to our complementary PISA Additional Analyses reports: *Welsh- and English-Medium School Results* (Classick *et al.,* 2020) and *Maths and Science Item Analysis in Wales* (Galvis *et. al.,* 2020).

Sample numbers

In the international database, the number of PISA pupils from schools in Wales was 3,165. For the analyses in this report, however, the majority of the analyses required data to be matched to National Pupil Database (NPD) data provided by Welsh Government, for example, to compare scores in regional consortia (Section 2) or to compare pupils' attainment in PISA and GCSE (Sections 3 and 4). When the PISA dataset was matched with the NPD, 28 pupils were excluded because of discrepant/duplicate Unique Pupil Numbers (UPN) identifiers, so the overall number of pupils for these analyses is 3,137.

Of this number (3,137), 257 pupils were classified as 'missing'. These included 114 pupils from maintained schools (probably pupils who had changed schools since the Register of Schools was last updated), and 143 pupils from independent schools who do not have UPNs and, therefore, their GCSE results are not recorded. All independent schools in the PISA sample were English-medium.

Accordingly, all analyses will be based on 3,137 pupils (with 'missing' as described), unless otherwise stated in the text.

Data suppression

NFER abides by the Office for National Statistics' guidance on Statistical Disclosure Control, which maintains the confidentiality of data that has been used. Accordingly, in some of the tables in this report, data has been removed, as signified by an asterisk. These tables feature the note "Indicates that N is suppressed to preserve pupil anonymity" underneath.

2. PISA Results by Regional Consortia

Key findings

- Pupils in ERW/West Wales gained the highest mean scores in all three subjects and GwE/North Wales pupils had the lowest mean scores. However, differences between regions were not statistically significant.
- The widest range of scores was seen in reading. This was the major domain and in 2018 this cognitive scale was extended to provide better discrimination at the lower levels.

The four Regional Education Consortia in Wales work with schools to raise standards in literacy and numeracy, providing a range of support which includes professional development and intervention programmes. The consortia and the broad geographic area for each is comprised as follows:

- GwE - Towards Excellence (North Wales): Conwy, Denbighshire, Flintshire, Gwynedd, Wrexham, Ynys Mon
- ERW - Educational through Regional Working (South West and Mid Wales): Carmarthenshire, Ceredigion, Neath Port Talbot, Pembrokeshire, Powys, Swansea
- CSC - Central South Consortium - (Central South Wales): Bridgend, Cardiff, Merthyr Tydfil, Rhondda Cynon Taff, Vale of Glamorgan
- EAS - Education Achievement Service (South East Wales): Blaenau Gwent, Caerphilly, Monmouthshire, Newport, Torfaen.

This regional analysis looks at PISA outcomes in reading, mathematics and science of pupils from the sampled schools in Wales that took part in PISA 2018. PISA sampling is designed to produce representative and accurate estimates for the country as a whole. It is possible, however, that mean scores of the sampled schools within different regions in Wales may not be fully representative of the whole region.

2.1 PISA results by Regional Consortia for reading, mathematics and science

Tables 2.1 to 2.3 show the mean scores for each of the consortia in each of the three core PISA subjects.

Pupils in ERW/West Wales gained the highest mean scores in all three subjects and GwE/North Wales pupils had the lowest mean scores, however, no statistically significant[3] differences between regions were found for any of the PISA subjects.

The widest range of scores was seen in reading. This is likely to be because, in 2018, the PISA reading assessment scale was extended to provide better information at the lower levels, which would have the effect of widening the range. The new reading scale is also likely to give a more precise estimate of the means.

In the tables that follow, we provide mean scores, standard errors and 95 per cent confidence intervals alongside the number of pupils.

Table 2.1 Average PISA 2018 reading scores by Regional Consortia

PISA 2018 National mean reading score: 483 (OECD Average: 487)					
Welsh Educational Regional Consortia	**Average (mean)**	**Standard error**	**95% Confidence interval**		**No. of pupils**
			Lower bound	**Upper bound**	
GwE/ North Wales	465	11	442	487	561
ERW/ South West and Mid Wales	493	6	482	505	791
CSC/ Central South Wales	484	6	472	497	943
EAS/ South East Wales	479	8	465	494	585
Missing	510	13	484	535	257

Source: PISA 2018 school census matched database

[3] When statistical significance is reported, it indicates that the compared means are significantly different at the 5% level.

Table 2.2 Average PISA 2018 mathematics scores by Regional Consortia

PISA 2018 National mean mathematics score: 487 (OECD Average: 489)					
Welsh Educational Regional Consortia	**Average (mean)**	**Standard error**	**95% Confidence interval**		**No. of pupils**
			Lower bound	**Upper bound**	
GwE/ North Wales	479	8	464	494	561
ERW/ South West and Mid Wales	495	6	484	506	791
CSC/ Central South Wales	484	6	473	496	943
EAS/ South East Wales	483	7	469	497	585
Missing	504	14	477	531	257

Source: PISA 2018 school census matched database

Table 2.3 Average PISA 2018 science scores by Regional Consortia

PISA 2018 National mean science score: 488 (OECD Average: 489)					
Welsh Educational Regional Consortia	**Average (mean)**	**Standard error**	**95% Confidence interval**		**No. of pupils**
			Lower bound	**Upper bound**	
GwE/ North Wales	481	9	463	498	561
ERW/ South West and Mid Wales	497	6	486	509	791
CSC/ Central South Wales	484	6	472	495	943
EAS/ South East Wales	485	6	473	497	585
Missing	508	15	478	538	257

Source: PISA 2018 school census matched database

3. PISA 2018 scores by GCSE results

Key findings

- The findings show a clear relationship between GCSE grades and PISA mean scores in all subjects, in that higher-achieving GCSE pupils also gained higher PISA mean scores and those with lower GCSE grades gained correspondingly lower mean scores in PISA.
- Similar corresponding patterns were seen when comparing GCSE grades with proficiency levels and with percentiles across all subjects.

In this section, we will look at pupils in Wales' best GCSE grades, from pupil-level annual school census (PLASC) data, mapped against the average (mean) scores, proficiency levels and percentiles in the corresponding PISA subjects.

We will present:

- PISA mean scores by GCSE grades (Section 3.1)
- PISA proficiency levels by GCSE grades (Section 3.2)
- PISA percentile scores by GCSE grades (Section 3.3).

3.1 PISA 2018 mean scores by GCSE grades

Tables 3.1 to 3.4 show PISA mean scores, standard errors and confidence intervals[4] against best:

- Overall GCSE literacy grade (from English/Welsh Language and English/Welsh literature)
- GCSE English/Welsh literature grade (for comparison to 2015[5])
- GCSE mathematics grade
- GCSE science grade.

Table 3.1 PISA reading mean scores by best overall GCSE (English/Welsh) literacy grade

GCSE GRADE	Average (mean)	Standard error of the mean	95% confidence interval lower bound	95% confidence interval upper bound	Number of pupils
A*	589	7	576	602	202
A	550	5	540	560	537
B	495	5	486	504	805
C	453	5	443	463	720
D	419	5	410	428	388
E	380	7	366	393	142
F	347	11	325	370	*
G	346	21	305	387	*
U	379	36	309	450	*
Missing[6]	510	13	484	535	257

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

PISA mean reading scores broadly corresponded with best GCSE literacy grades. The biggest gap in PISA mean scores was between GCSE Grades A and B (55 PISA score points).

[4]These are provided alongside the number of pupils as an indication of the robustness of the estimates and the levels of uncertainty around the means.
[5] The 2015 data tables are provided in Appendix A.
[6] In all tables in this section, 'missing' refers to pupils for whom it was not possible to match their PISA score to a GCSE grade. These include pupils in independent schools and those who had moved school.

To allow direct comparison with the analyses of the 2015 data, we then compared PISA means against GCSE literature grades (only).

Table 3.2 PISA reading mean scores by best GCSE (English/Welsh) literature grade

GCSE GRADE	Average (mean)	Standard error of the mean	95% confidence interval lower bound	95% confidence interval upper bound	Number of pupils
A*	592	8	577	608	155
A	554	5	544	563	474
B	502	5	492	513	753
C	465	5	456	474	671
D	425	6	414	436	315
E	412	8	396	428	116
F	379	16	347	410	*
G	360	18	324	396	*
U	295	29	237	352	*
Missing	449	8	433	464	580

*Indicates that N is suppressed to preserve pupil anonymity.

The number of 'missing' pupils (580) is higher as fewer pupils sit GCSE English and/or Welsh literature than GCSE English and/or Welsh Language.

Source: PISA 2018 school census matched database

Again, the analysis showed that PISA mean scores decreased with the corresponding GCSE grades.

A similar correspondence was recorded in the 2015 analysis, but in 2018 the PISA reading scores were higher at each GCSE grade, except for Grade B, which had a similar PISA score.

Table 3.3 PISA mathematics mean scores by best GCSE mathematics grade

GCSE GRADE	Average (mean)	Standard error of the mean	95% confidence interval lower bound	95% confidence interval upper bound	Number of pupils
A*	577	5	568	587	356
A	541	5	532	550	374
B	504	5	495	513	645
C	474	4	466	482	702
D	437	4	429	445	349
E	413	5	403	422	227
F	387	6	375	400	119
G	382	10	363	402	74
U	359	15	330	388	30
Missing	503	13	477	530	261

‘Missing’ includes one X grade award.

Source: PISA 2018 school census matched database

PISA mean scores in mathematics reflected achievement levels in the same GCSE subject. Pupils had 30 - 37 point differences between each grade in the GCSE A* to D range, with smaller point differences below grade D.

Table 3.4 PISA science mean scores by best GCSE science grade

GCSE GRADE	Average (mean)	Standard error of the mean	95% confidence interval lower bound	95% confidence interval upper bound	Number of pupils
A*	588	5	578	598	326
A	550	5	540	560	430
B	502	4	493	510	795
C	462	4	455	470	637
D	423	5	413	433	254
E	410	5	400	421	177
F	385	7	371	398	131
G	369	9	350	387	53
U	363	13	338	388	52
Missing	496	13	470	522	282

'Missing' includes three X grade awards.

Source: PISA 2018 school census matched database

PISA science performance was also found to correspond to best GCSE science grade, with mean score differences larger at Grades A* to D and smaller in the lowest grades.

Overview of PISA 2018 mean scores by GCSE grades

The findings largely show a clear relationship between GCSE grades and mean scores, in that higher-achieving GCSE pupils show higher PISA mean scores and the lower the GCSE grade, the lower the PISA performance.

3.2 PISA 2018 proficiency levels by GCSE grades

For each subject, we will map the proportions of pupils reaching the different PISA 2018 proficiency levels against their best GCSE grade.

The PISA proficiency levels are devised by the PISA Consortium. They can be used as a guide for interpreting mean scores against the range of difficulty of subject tasks[7]. In 2018, and since reading was the major domain, an additional reading proficiency level was added, Level 1c, to distinguish pupils at the lowest levels of literacy achievement, who would previously have been classified as below Level 1b (OECD, 2019b). Accordingly, for PISA 2018, there are eight proficiency levels for reading (Levels 1c – 6), six for mathematics (Levels 1 – 6) and seven for science (Levels 1b – 6). The level descriptions for each proficiency level are provided in Appendix B for each subject.

Tables 3.5 to 3.8 show the percentage of pupils at each GCSE grade (literacy, English or Welsh literature, mathematics and science) against each of the PISA 2018 proficiency levels for the corresponding subject.

[7]For further detail, see Chapter 2, OECD, 2019a.

Table 3.5 Percentage of pupils at PISA reading proficiency levels by best GCSE literacy grade (English/Welsh language or literature)

PISA 2018 proficiency level	Below Level 1c	Level 1c	Level 1b	Level 1a	Level 2	Level 3	Level 4	Level 5	Level 6
GCSE GRADE	**% at PISA proficiency level**								
A*	0	0	0	*	6	23	40	26	*
A	0	*	*	3	12	32	38	13	*
B	0	*	2	10	29	36	20	*	0
C	0	*	5	18	39	29	7	*	0
D	0	*	10	32	40	13	3	*	0
E	0	*	18	46	23	8	*	0	0
F	2	9	26	46	14	3	0	0	0
G	0	0	44	38	6	12	0	0	0
U	0	0	20	40	40	0	0	0	0
Missing[8]	0	*	*	10	18	30	29	9	*

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

[8] In all tables in this section, 'missing' refers to pupils for whom it was not possible to match their PISA score to a GCSE grade.

Table 3.5 shows a broad relationship between pupils' best GCSE literacy grade and the corresponding PISA reading proficiency level. The majority of pupils at GCSE grades A and A* achieved PISA proficiency Level 4 or above and over half of GCSE grade B pupils were above Level 3[9] in PISA. Most Grade C and D pupils were around Level 2 in PISA reading, while the majority of pupils at the lower grades E to G were below Level 2. Pupils whose best literacy grade was U also tended to be at Level 2 or below.

[9] Level 3 is considered to be a level that prepares pupils for success later in life. Level 2 is considered a baseline level. See: https://www.oecd-ilibrary.org/education/OECD paper on Academic resilience_e22490ac-en

Table 3.6 Percentage of pupils at PISA reading proficiency levels by best GCSE literature grade (English/Welsh literature only)

PISA 2018 proficiency level	Below Level 1c	Level 1c	Level 1b	Level 1a	Level 2	Level 3	Level 4	Level 5	Level 6
GCSE GRADE	**% at PISA proficiency level**								
A*	0	0	0	*	*	25	35	28	*
A	0	*	*	*	12	31	41	13	*
B	0	*	2	9	26	36	22	5	*
C	0	*	3	16	39	31	9	*	0
D	0	*	10	31	37	15	6	*	0
E	0	*	9	34	35	16	*	0	0
F	2	2	22	43	18	10	2	0	0
G	0	6	24	53	12	6	0	0	0
U	0	29	43	14	14	0	0	0	0
Missing	0	2	9	21	25	23	15	4	*

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

Table 3.6 again shows a clear relationship between pupils' best GCSE grade in either English or Welsh literature and their PISA reading proficiency level, with more than half of pupils at GCSE grades A* or A achieving PISA Level 4 or above, and GCSE grade B performing at PISA level 3 and above. This finding is consistent with 2015, although there were some notable differences, with the highest- and lowest-achieving pupils at GCSE showing the greatest differences compared to 2015. The proportions of A* pupils that performed at Level 5 increased from 17 per cent in 2015 to 28 per cent in 2018.

Table 3.7 Percentage of pupils at PISA mathematics proficiency levels by best GCSE mathematics grade

PISA 2018 proficiency level	Below Level 1	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
GCSE GRADE	**% at PISA proficiency level**						
A*	*	*	3	20	48	27	*
A	*	*	10	41	41	7	*
B	*	5	25	45	22	*	0
C	*	12	41	37	7	*	0
D	*	28	50	15	*	0	0
E	10	46	40	*	*	0	0
F	25	50	*	*	*	0	0
G	30	50	15	*	*	0	0
U	43	*	*	0	0	0	0
Missing	*	12	19	26	24	14	*

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

Table 3.8 Percentage of pupils at PISA science proficiency levels by best GCSE science grade

PISA 2018 proficiency level	Below Level 1b	Level 1b	Level 1a	Level 2	Level 3	Level 4	Level 5	Level 6
GCSE GRADE	**% at PISA proficiency level**							
A*	0	*	*	*	27	51	20	*
A	0	0	*	10	44	39	*	0
B	*	*	5	31	46	16	*	0
C	0	2	16	46	33	4	0	0
D	*	*	34	49	12	0	0	0
E	*	7	45	40	7	*	0	0
F	*	14	52	30	*	*	0	0
G	*	*	57	23	0	0	0	0
U	*	37	40	*	0	0	0	0
Missing	0	*	14	21	26	27	8	*

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

Tables 3.7 and 3.8 show that the clear relationship between GCSE grade and PISA proficiency levels continued in GCSE mathematics and science, with over half of pupils at GCSE grades A* - B in these subjects performing at PISA Level 3 and above in the corresponding PISA subjects.

Overview of PISA 2018 proficiency levels by GCSE grades

The majority of pupils who achieve PISA Levels 5 and 6 tend to also get GCSE A*. At PISA Levels 3 and 4 most pupils achieved at GCSE grades A and B.

3.3 PISA 2018 percentile scores by GCSE grades

In this section, we examine the distribution of PISA scores for each GCSE grade by percentile scores, along with the standard deviation (SD) to further explore the spread of performance within the sample.

The 10th percentile is the score below which the lowest performing ten per cent of pupils lie, while the 90th percentile is the score above which the highest performing ten per cent lie. The score at the 50th percentile can also be considered as the 'median' score, as it is the score below which 50 per cent of pupils lie.

For each subject, we will compare 10th, 25th, 50th, 75th and 90th percentile scores for PISA 2018 scores by best GCSE grades, as shown in Tables 3.9 to 3.12.

Table 3.9 PISA reading percentile scores by best GCSE literacy grade

	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	SD	No. of pupils
GCSE GRADE							
A*	497	544	591	637	680	74	202
A	446	505	556	602	641	79	537
B	399	444	498	547	590	78	805
C	353	408	456	503	541	77	720
D	331	374	419	463	501	73	388
E	293	339	377	418	478	72	142
F	247	296	345	392	435	75	*
G	267	279	337	366	473	76	*
U	314	320	357	420	431	54	*
Missing	389	451	515	575	619	94	257

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

The scores at each percentile corresponded to achievement at each GCSE grade. For example, at the 10th percentile, the scores progressively decreased from grade A* through the grades below. Pupils at GCSE grades C and D who performed at the 90th percentile had scores which were higher than the scores of pupils at GCSE grade A* at the 10th percentile.

Table 3.10 PISA reading percentile scores by best GCSE literature grade (English/Welsh)

	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	SD	No. of pupils
GCSE GRADE							
A*	497	539	592	646	688	77	155
A	453	507	561	602	641	76	474
B	402	450	506	559	601	81	753
C	377	419	465	513	555	75	671
D	332	373	418	470	524	80	315
E	326	362	416	463	499	75	116
F	275	332	367	433	501	86	*
G	259	297	349	394	429	76	*
U	215	230	272	313	377	68	*
Missing[10]	314	378	449	516	584	102	580

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

As in GCSE literacy, pupils' performance at each percentile decreased in line with their GCSE grade achievement at that percentile. Pupils at GCSE grades C - D and E - F at the 90th percentile performed better or similarly, respectively, than their grade A* peers at the 10th percentile.

Pupils' GCSE English or Welsh literature achievement and PISA 2018 reading performance was similar to that of 2015, but the standard deviations were higher in 2018, showing a greater variation in scores than in 2015. As mentioned before, this is likely to be a result of the more precise measurement available in 2018, which allows greater discrimination among lower-achieving pupils. The scores for pupils at GCSE A* and F grades at the 90th percentile were around 60 and 40 points higher in 2018, respectively.

[10] The number of 'missing' pupils (580) is higher as fewer pupils sit GCSE English and/or Welsh literature than GCSE English and/or Welsh Language.

Table 3.11 PISA mathematics percentile scores by best GCSE mathematics grade

	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	SD	No. of pupils
GCSE GRADE							
A*	517	547	578	613	638	59	356
A	479	507	541	573	595	57	374
B	437	471	507	540	568	62	645
C	407	439	477	508	537	60	702
D	372	400	441	471	496	58	349
E	358	389	414	437	467	55	227
F	323	357	388	418	444	58	119
G	322	346	374	408	440	61	74
U	287	313	364	398	429	61	30
Missing	392	440	510	564	613	87	261

'Missing' includes one X grade award.

Source: PISA 2018 school census matched database

In mathematics, the scores at each percentile decreased in line with lower GCSE grades. However, the finding that pupils at some lower GCSE grades and at the 90th percentile performed comparably to their A* counterparts at the 10th percentile was not present.

Table 3.12 PISA science percentile scores by best GCSE science grade

	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	SD	No. of pupils
GCSE GRADE							
A*	519	547	589	626	662	64	326
A	478	516	552	590	619	63	430
B	424	462	503	543	578	66	795
C	388	429	466	500	530	64	637
D	352	390	425	457	488	59	254
E	340	372	405	453	477	64	177
F	321	350	382	417	442	62	131
G	310	338	357	402	430	58	53
U	303	318	355	398	442	63	52
Missing	367	416	500	571	617	100	282

'Missing' includes three X grade awards.

Source: PISA 2018 school census matched database

A similar pattern of decreasing performance at each percentile and GCSE grade was found in science. Further, there was no overlap between pupils in the lower GCSE grades' scores at the highest percentile and GCSE A* pupils at the lowest percentile.

Overview of PISA 2018 percentiles by GCSE grades

In general, and for each subject, the mean PISA scores are consistent with expectations, in that the scores increase according to both GCSE grade and PISA percentiles. Higher-achieving GCSE pupils had higher scores at each PISA percentile than their lower-achieving GCSE counterparts.

3.4 Summary of PISA 2018 and GCSE comparisons

When examining PISA mean scores by GCSE grades, the findings show a clear relationship across subjects, in that higher-achieving GCSE pupils also generally gained higher PISA mean scores and those with lower GCSE grades gained correspondingly lower mean scores in PISA.

The same was true of PISA 2018 proficiency levels by GCSE grades where there was also a broad correspondence between pupils' GCSE achievement and PISA performance.

Similarly, the findings were consistent with expectations when comparing pupils' best GCSE grades at each of the five PISA percentiles considered for each subject.

4. PISA and GCSE performance in English- and Welsh-medium schools

Key findings

- PISA results: pupils in Welsh-medium schools gained significanty lower PISA scores than pupils in English-medium schools in reading and mathematics.
- In PISA science, the difference between pupils' scores in English- and Welsh medium schools was not significant.
- GCSE results: pupils in the PISA sample in Welsh-medium schools achieved significantly higher GCSE capped 9 scores than their English-medium counterparts (when weighted by FSM and gender).
- When comparing mean scores by subject, the largest difference between English- and Welsh-medium was found in reading.
- In reading, the PISA performance of Welsh-medium pupils was lower than that of English-medium pupils in every SES quintile, from the most disadvantaged 20 per cent through to the most advantaged 20 per cent.
- In mathematics and science, Welsh-medium pupils' PISA performance was lower than that of their English-medium peers only in the highest (most advantaged) SES quintile.
- Pupils who took the PISA assessment in Welsh performed lower in all three PISA subjects than those that took the assessment in English.

The NFER PISA Additional Analyses report *Welsh- and English-Medium School Results (Classick et al.,* 2020*)* outlines differences in PISA performance in reading between pupils in English- and Welsh-medium schools.

In this section, we offer additional insights into differences between English- and Welsh-medium schools in terms of PISA performance by subject, by GCSE capped 9 scores, by socioeconomic status (SES) and by language of assessment in PISA.

4.1 PISA results in English- and Welsh-medium schools

The analysis comparing the 2018 PISA reading scores of pupils in English- and Welsh-medium schools (Classick *et al*., 2020) shows that pupils in Welsh-medium schools gained PISA scores that were significantly lower than those of their peers in English-medium schools.

In this study we also looked at PISA performance in mathematics and science.

4.1.1 Comparison of PISA performance by subject

Pupils in English-medium schools significantly outperformed their Welsh-medium counterparts in reading and mathematics. However, there was no significant difference in PISA science mean scores, as shown in Table 4.1:

Table 4.1 Mean scores for each PISA subject by school medium

PISA subject	English-medium	Welsh-medium	Difference
Reading	495*	452	43
Mathematics	491*	476	15
Science	492	477	15

* Indicates a significant difference from the Welsh-medium pupils.

Source: PISA 2018 school census matched database

When comparing mean scores by subject, the largest difference between English- and Welsh-medium was found in reading.

These results might suggest that pupils in Welsh-medium schools were, on average, lower achievers than those in English-medium schools. Alternatively, they may have been an unusually weak PISA cohort compared to the English-medium group.

To probe further we conducted a comparison of GCSE scores between the two groups

4.2 GCSE results in English- and Welsh-medium schools

The PISA results were significantly different between pupils in English- and Welsh-medium schools. In order to explore whether these findings may be a true indication of weaker performance among pupils in Welsh-medium schools we looked at the GCSE results of the English- and Welsh-medium pupils in the PISA sample.

The aim of the PISA sampling is to obtain a representative sample of pupils at a national level. This includes all pupils in the required age group, regardless of the medium of the school, as this was not one of the stratifying variables in selecting the sample.

If the Welsh-medium PISA contingent also gained lower GCSE results than either the national sample or the English-medium PISA sample, this would go some way towards explaining the PISA results in Section 4.1.1.

We first compared the PISA cohorts in English-and Welsh-medium schools against the corresponding national cohorts to check whether the PISA cohorts reflected the national distribution in terms of GCSE capped 9 scores. We then compared pupils in English- and Welsh-medium schools in terms of their GCSE capped 9 scores.

4.2.1 PISA samples against national cohorts

The comparisons of PISA school cohorts against the national cohort on GCSE capped 9 scores showed that both the English- and Welsh-medium PISA cohorts' GCSE achievements were slightly higher than those of their respective national cohorts, as shown in Tables 4.2 and 4.3

This finding is not surprising, as the majority of special schools and pupils with some specific/additional learning needs are excluded from the PISA sample.

Table 4.2 Mean GCSE capped 9 score by Welsh-medium PISA and national cohorts

Welsh-medium cohort	Mean GCSE capped 9 score	Standard deviation
PISA	389*	72
National	374	84

*Indicates a significant difference from the pupils in the national cohort.

Source: PISA 2018 school census matched database

Table 4.3 Mean GCSE capped 9 score by English-medium PISA and national cohorts

Welsh-medium cohort	Mean GCSE capped 9 score	Standard deviation
PISA	382*	81
National	360	98

*Indicates a significant difference from the pupils in the national cohort.

Source: PISA 2018 school census matched database

4.2.2 GCSE results in English- and Welsh-medium PISA schools

Table 4.4 shows the difference in GCSE achievement, as measured by capped 9 score, by school medium.

Table 4.4 Mean GCSE capped 9 score by PISA cohort and school medium

Welsh-medium cohort	Mean GCSE capped 9 score	Standard deviation
Welsh-medium	389*	81
English-medium	382	72

*Indicates a significant difference from the English-medium pupils.

Source: PISA 2018 school census matched database

Welsh-medium pupils who took part in PISA had significantly higher average capped 9 GCSE scores than that of their counterparts in English-medium schools.

The discrepancy in PISA scores cannot be explained by low overall achievement in the Welsh-medium sample.

Table 4.5 shows the proportions of English- and Welsh-medium school pupils at each of five GCSE capped 9 score percentiles.

Table 4.5 Proportions of English- and Welsh-medium pupils in the PISA sample by GCSE capped 9 score band

	School medium and cohort	English-medium PISA	Welsh-medium PISA	Total
Percentile group at GCSE Capped 9 average				
Lowest	No. of pupils	265	77	342
	% within group	**12.7**	**9.8**	**11.9**
Second Lowest	No. of pupils	424	138	562
	% within group	**20.3**	**17.5**	**19.5**
Middle	No. of pupils	426	167	593
	% within group	**20.4**	**21.2**	**20.6**
Second highest	No. of pupils	472	224	696
	% within group	**22.6**	**28.5**	**24.2**
Highest	No. of pupils	506	181	687
	% within group	**24.2**	**23.0**	**23.9**
Total	No. of pupils	2093	787	2880
	% within group	**100.0**	**100.0**	**100.0**

Note: Total number of pupils (2,880) is lower than the Wales PISA sample (3,165) due to weighting by free school meals (FSM) eligibility and gender.

Source: PISA 2018 school census matched database

The stronger GCSE performance of pupils in Welsh-medium schools compared to those in English-medium schools is reflected in the two highest GCSE percentiles, with 51.5 per cent of Welsh-medium pupils in the highest bands compared to 46.7 per cent of English-medium pupils achieving these levels in GCSE.

These results suggest that the Welsh-medium pupils in the PISA sample were unlikely to have been an unusually weak cohort compared the English-medium group.

4.3 Socioeconomic status (SES) in English- and Welsh-medium schools

Classick *et al*., (2020), found that English-medium pupils outperformed their Welsh-medium counterparts in PISA reading overall and by FSM.

In this section, we examine PISA performance in all three PISA subjects, by school medium and SES quintiles.

The SES measure used when examining PISA performance is the OECD ESCS (economic, social and cultural status) Index. This measure allows for a more precise examination of SES across five levels whereas for FSM, pupils are either eligible or not. This scale is based on pupils' responses to questions about their parents' backgrounds and education, and possessions in their homes.

4.3.1 PISA performance by SES in English- and Welsh-medium schools

We first compared the proportions of PISA pupils by school medium by PISA ESCS quintile.

There were no significant differences in terms of the distribution across ESCS levels in English- and Welsh-medium PISA schools.

The discrepancy in PISA scores therefore cannot be explained by SES differences between the groups.

There were slightly higher proportions of the most disadvantaged pupils in English-medium schools (see Appendix C) but the differences were not statistically significant.

When comparing school medium and each PISA subject by ESCS, we split the PISA cohort into five SES quintiles, from the lowest 20 per cent (most disadvantaged) to the highest 20 per cent (least disadvantaged).

Table 4.6 shows that, for reading, pupils in English-medium schools significantly outperformed those in Welsh-medium schools at every ESCS band. For mathematics and science, however, the stronger performance of pupils in English-medium schools was seen only at the highest, least disadvantaged, ESCS quintile.

Table 4.6 Mean scores for each PISA subject by ESCS quintile and school medium

PISA subject and school medium	Reading		Mathematics		Science	
	Eng-med	Welsh-med	Eng-med	Welsh-med	Eng-med	Welsh-med
Lowest	472*	441	462	453	465	461
Second lowest	486*	436	480	465	479	469
Middle	494*	457	491	479	494	482
Second highest	503*	476	503	493	502	495
Highest	540*	474	534*	497	536*	499

* Indicates a significant difference from the Welsh-medium pupils.

Missing group not included.

Source: PISA 2018 school census matched database

4.3.2 Comparison of GCSE capped 9 results by SES in English- and Welsh-medium schools

We compared the average GCSE capped 9 scores of pupils in the same five ESCS quintiles in English- and Welsh-medium schools. No significant differences were found at any SES level.

Overall, these results indicate that the Welsh-medium pupils in the PISA sample did not have lower SES levels compared to the English-medium group, and that they had comparable GCSE results.

4.4 PISA scores by language of assessment

Our *Welsh- and English-Medium School Results* report (Classick *et al.,* 2020) found that the language of the PISA assessment had a sizeable negative impact on the PISA reading performance of those who had taken the assessment in Welsh.

Analyses for this report confirmed these findings for reading[11], and showed that the language of assessment had a significant effect on the results for maths and science as well. Significant differences in PISA scores by language of assessment were found, with pupils that took the assessment in English outperforming those who took it in Welsh in all three PISA subjects.

In Welsh-medium schools, over half of pupils took PISA in Welsh (see Appendix D for a breakdown of numbers of pupils who took the assessment in English and Welsh by school medium).

Table 4.7 shows the differences in PISA mean scores by language of assessment:

Table 4.7 Mean scores for each PISA subject by language of assessment

PISA subject	English	Welsh	Difference
Reading	492*	421	71
Mathematics	492*	464	28
Science	500*	460	40

* Indicates a significant difference from pupils that took the test in the Welsh language.

Source: PISA 2018 school census matched database

Although pupils who take the PISA assessment in English have significantly higher scores than those who take the test in Welsh in all three subjects, the difference between reading score is much greater than in mathematics or science.

When we compared the GCSE results for pupils who took the PISA test in Welsh, they were not significantly different from those who took the PISA test in English.

[11] The reading mean scores are slightly lower than those presented in PISA 2018 Additional Analyses: Welsh- and English-Medium School Results (Classick *et al*., 2020) due to the PISA weighted sample of 3,021 vs 3,165. In that report, the mean reading score by language of assessment was 494 for English and 422 for Welsh.

4.5 Summary of PISA and GCSE performance in English- and Welsh-medium schools

English-medium pupils in the PISA sample outperformed their Welsh-medium counterparts in PISA reading and mathematics, but not science.

Welsh-medium pupils in the PISA sample gained significantly higher GCSE capped 9 scores than their English-medium peers, when weighted by FSM and gender.

No differences were found in the proportions of English- and Welsh-medium pupils at each of the five PISA FSM bands. Our analysis by SES found that English-medium pupils' PISA reading performance was stronger than that of Welsh-medium pupils in every SES quintile, from the most disadvantaged 20 per cent through to the least disadvantaged 20 per cent. However, the better performance of English-medium pupils was only seen in the highest (least disadvantaged) SES quintile in PISA mathematics and science. No significant differences between GCSE capped 9 scores in English- and Welsh medium schools were found at any SES level.

Pupils who took the PISA assessment in Welsh performed lower in all three PISA subjects than those who took the assessment in English. However, there was no significant difference in the GCSE achievement of pupils in these groups.

5. Conclusions

Our analysis of regional PISA mean scores for the three core PISA subjects (reading, mathematics and science) found no significant differences between Wales' four Regional Consortia in any subject.

When comparing PISA mean scores and proficiency levels with best GCSE grade, findings were consistent with expectations, showing largely clear relationships between the strength of achievement at GCSE and performance in PISA. Similarly, pupils' best GCSE grades were consistent with each of the five PISA score percentiles.

English-medium pupils significantly outperformed their Welsh-medium counterparts in PISA reading and mathematics, but not science. Welsh-medium pupils who took PISA, however, achieved more highly in their GCSEs than their English-medium PISA peers when weighted by FSM and gender. These findings suggest that the Welsh-medium pupils in the PISA sample were, therefore, unlikely to have been an unusually weak cohort compared the English-medium group.

Socioeconomic status does not appear to be a factor in the stronger PISA performance of pupils in English-medium schools compared to their Welsh-medium peers. There were no significant differences in the proportions of pupils at each of the five PISA ESCS bands when compared across the medium of the schools, suggesting that differences in SES levels cannot explain the differences in PISA performance by medium of school. Pupils in English-medium schools performed better than those in Welsh-medium schools in every SES band for reading. For PISA mathematics and science the stronger performance of English-medium pupils was only found in the *least disadvantaged* 20 per cent of pupils.

There were no significant differences in GCSE capped 9 scores between English- and Welsh-medium schools at any SES level, again suggesting that SES is not the reason for the discrepancy in PISA scores.

What does appear to be a factor in the weaker PISA performance of Welsh-medium pupils is the language of assessment. Pupils who took the PISA assessment in Welsh (all of whom attend Welsh-medium schools) performed lower than those who took the assessment in English in all three PISA subjects. The greatest difference in mean scores was for PISA reading.

These findings suggest that the language of assessment is a salient focus for further research. The definition of "Welsh-medium schools" is broad and includes those that offer fully Welsh-medium education, where the teaching is wholly in Welsh (these comprise around 16 per cent of Welsh-medium schools); bilingual schools; dual-medium schools and schools that teach English with significant Welsh provision (around ten per cent of Welsh-medium schools). Within this group of schools, there are varying proportions of subjects taught in English and/or Welsh; for example, some schools teach 80 plus per

cent of subjects (excluding English and Welsh) solely in Welsh, whereas other schools teach a lower proportion of subjects in both languages. Further analysis would require comparisons of pupil groups in all these different types of Welsh-medium school.

Other factors to consider would be pupils' fluency in Welsh language and the extent to which pupils are considered bilingual or may be classed as second language pupils. Some insights may be drawn from the data on the language spoken at home and/or with friends.

Translation of the PISA questions into Welsh could be a relevant factor, although a thorough analysis of the translations was conducted during the 2018 PISA cycle to ensure that no assessment questions included language that was unfamiliar in Welsh classrooms. The length of the reading texts are often longer in Welsh than in English and, indeed, the word count in some questions can be up to 25 per cent higher in Welsh compared to English. It is possible that the reading time is, therefore, correspondingly longer and could put those taking the test in Welsh at a disadvantage in a timed test such as PISA. This may also account for the greater discrepancies in reading compared with mathematics and science as the reading load is much greater in reading assessments. A comparison of the time taken to read the assessments in each language could be an area for further investigation, particularly in an adaptive testing situation.

Finally, further explorations of curricular differences in English- and Welsh-medium schools, GCSE specifications and moderation of English and Welsh GCSEs may also shed some light on the reasons why PISA results do not match the pattern of GCSE results across school medium.

References

Classick, R., Galvis, M. and Sizmur J. (2020). *PISA 2018 Additional Analyses: Welsh- and English-Medium School Results*. Slough: NFER.

Galvis, M., Hope, C., Sizmur J. and Wheater, R. (2020). *PISA 2018 Additional Analyses: Maths and Science Item Analysis in Wales*. Slough: NFER.

Organisation for Economic Co-operation and Development (2019a). *PISA 2018 Results (Volume I): What Students Know and Can Do, Chapter 2, How PISA Results are Reported: What is a score*? [online]. Available: https://doi.org/OECD PISA 2018 Results (Volume I) What Students Know and Can Do\10.1787/5f07c754-en [28 October, 2020].

Organisation for Economic Co-operation and Development (2019b). *PISA 2018 Assessment and Analytical Framework* [online]. Available: https://doi.org/OECD PISA 2018 Assessment and Analytical Framework\10.1787/b25efab8-en [28 October, 2020].

Sizmur, J., Ager, R., Bradshaw, J., Classick, R., Galvis, M., Packer, J., Thomas D. and Wheater R. (2019). *Achievement of 15-year-olds in Wales: PISA 2018 National Report.* [online]. Available: https://www.nfer.ac.uk/achievement-of-15-year-olds-in-wales-pisa-2018-national-report/PISA 2018 National Report for Wales [28 October, 2020].

Appendix A

Appendices A.1 to A.3 show the 2015 analysis for PISA reading and GCSE English or Welsh literature, by mean scores, percentages and percentiles by GCSE grade.

Appendix A.1 PISA 2015 reading mean scores by best GCSE literature grade (English/Welsh)

GCSE GRADE	Average (mean)	Standard error of the mean	95% confidence interval lower bound	95% confidence interval upper bound	Number of pupils
A*	576	7	563	589	184
A	544	4	535	552	532
B	502	4	494	510	813
C	457	3	450	463	1134
D	420	5	411	430	421
E	388	7	375	401	175
F	369	8	353	385	87
G	346	16	315	377	*
U	351	23	305	395	*
Missing	501	19	463	539	*

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

Appendix A.2 Percentage of pupils at PISA 2015 reading proficiency levels by best GCSE literature grade (English/Welsh literature only)

PISA 2015 proficiency level	Below Level 1b	Level 1b	Level 1a	Level 2	Level 3	Level 4	Level 5	Level 6
GCSE GRADE	**% at PISA proficiency level**							
A*	0	0	0	*	27	48	17	*
A	0	0	*	15	37	36	10	*
B	0	*	6	30	41	19	*	0
C	0	3	20	41	28	7	1	0
D	*	8	32	42	15	*	0	0
E	*	17	45	26	8	*	0	0
F	*	26	48	20	*	0	0	0
G	4	37	47	11	1	0	0	0
U	0	29	71	0	0	0	0	0
Missing	3	6	8	23	27	19	12	2

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

Appendix A.3 PISA 2015 reading percentile scores by best GCSE literature grade (English/Welsh)

GCSE GRADE	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	SD	No. of pupils
A*	500	540	579	614	646	60	184
A	461	499	546	589	626	65	532
B	420	459	501	545	585	65	813
C	371	411	456	502	545	68	1134
D	337	377	422	464	502	65	421
E	304	346	387	427	474	66	175
F	297	328	370	406	441	57	87
G	282	316	344	369	417	48	*
U	332	332	359	368	368	18	*
Missing	342	441	501	587	642	110	*

*Indicates that N is suppressed to preserve pupil anonymity.

Source: PISA 2018 school census matched database

Appendix B

Appendices B.4 to B.6 below describe what pupils can typically do at each proficiency level for the three core subjects: reading, mathematics and science.

Appendix B.4 Reading proficiency levels

Level	Percentage of pupils at this level	What pupils can typically do at each level
6	**OECD**: 1% perform tasks at Level 6 **Wales**: 1%	Readers at Level 6 can comprehend lengthy and abstract texts in which the information of interest is deeply embedded and only indirectly related to the task. They can compare, contrast and integrate information representing multiple and potentially conflicting perspectives, using multiple criteria and generating inferences across distant pieces of information to determine how the information may be used. Readers at Level 6 can reflect deeply on the text's source in relation to its content, using criteria external to the text. They can compare and contrast information across texts, identifying and resolving inter-textual discrepancies and conflicts through inferences about the sources of information, their explicit or vested interests, and other cues as to the validity of the information. Tasks at Level 6 typically require the reader to set up elaborate plans, combining multiple criteria and generating inferences to relate the task and the text(s). Materials at this level include one or several complex and abstract text(s), involving multiple and possibly discrepant perspectives. Target information may take the form of details that are deeply embedded within or across texts and potentially obscured by competing information.
5	**OECD**: 9% perform tasks at least at Level 5 **Wales**:	Readers at Level 5 can comprehend lengthy texts, inferring which information in the text is relevant even though the information of interest may be easily overlooked. They can perform causal or other forms of reasoning based on a deep understanding of extended

Level	Percentage of pupils at this level	What pupils can typically do at each level
	7%	pieces of text. They can also answer indirect questions by inferring the relationship between the question and one or several pieces of information distributed within or across multiple texts and sources. Reflective tasks require the production or critical evaluation of hypotheses, drawing on specific information. Readers can establish distinctions between content and purpose, and between fact and opinion as applied to complex or abstract statements. They can assess neutrality and bias based on explicit or implicit cues pertaining to both the content and/or source of the information. They can also draw conclusions regarding the reliability of the claims or conclusions offered in a piece of text. For all aspects of reading, tasks at Level 5 typically involve dealing with concepts that are abstract or counterintuitive, and going through several steps until the goal is reached. In addition, tasks at this level may require the reader to handle several long texts, switching back and forth across texts in order to compare and contrast information.
4	**OECD**: 28% perform tasks at least at Level 4 **Wales**: 25%	At Level 4, readers can comprehend extended passages in single or multiple-text settings. They interpret the meaning of nuances of language in a section of text by taking into account the text as a whole. In other interpretative tasks, pupils demonstrate understanding and application of ad hoc categories. They can compare perspectives and draw inferences based on multiple sources. Readers can search, locate and integrate several pieces of embedded information in the presence of plausible distractors. They can generate inferences based on the task statement in order to assess the relevance of target information. They can handle tasks that require them to memorise prior task context. In addition, pupils at this level can evaluate the relationship between specific statements and a person's overall stance or conclusion about a topic.

Level	Percentage of pupils at this level	What pupils can typically do at each level
		They can reflect on the strategies that authors use to convey their points, based on salient features of texts (e.g. titles and illustrations). They can compare and contrast claims explicitly made in several texts and assess the reliability of a source based on salient criteria. Texts at Level 4 are often long or complex, and their content or form may not be standard. Many of the tasks are situated in multiple-text settings. The texts and the tasks contain indirect or implicit cues.
3	**OECD**: 54% perform tasks at least at Level 3 **Wales**: 51%	Readers at Level 3 can represent the literal meaning of single or multiple texts in the absence of explicit content or organisational clues. Readers can integrate content and generate both basic and more advanced inferences. They can also integrate several parts of a piece of text in order to identify the main idea, understand a relationship or construe the meaning of a word or phrase when the required information is featured on a single page. They can search for information based on indirect prompts, and locate target information that is not in a prominent position and/or is in the presence of distractors. In some cases, readers at this level recognise the relationship between several pieces of information based on multiple criteria. Level 3 readers can reflect on a piece of text or a small set of texts, and compare and contrast several authors' viewpoints based on explicit information. Reflective tasks at this level may require the reader to perform comparisons, generate explanations or evaluate a feature of the text. Some reflective tasks require readers to demonstrate a detailed understanding of a piece of text dealing with a familiar topic, whereas others require a basic understanding of less familiar content. Tasks at Level 3 require the reader to take many features into account when comparing, contrasting or categorising information. The required information is

Level	Percentage of pupils at this level	What pupils can typically do at each level
		often not prominent or there may be a considerable amount of competing information. Texts typical of this level may include other obstacles, such as ideas that are contrary to expectation or negatively worded.
2	**OECD**: 77% perform tasks at least at Level 2 **Wales**: 78%	Readers at Level 2 can identify the main idea in a piece of text of moderate length. They can understand relationships or construe meaning within a limited part of the text when the information is not prominent by producing basic inferences, and/or when the text(s) include some distracting information. They can select and access a page in a set based on explicit though sometimes complex prompts, and locate one or more pieces of information based on multiple, partly implicit criteria. Readers at Level 2 can, when explicitly cued, reflect on the overall purpose, or on the purpose of specific details, in texts of moderate length. They can reflect on simple visual or typographical features. They can compare claims and evaluate the reasons supporting them based on short, explicit statements. Tasks at Level 2 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.
1a	**OECD**: 92% perform tasks at least at Level 1a **Wales**: 94%	Readers at Level 1a can understand the literal meaning of sentences or short passages. Readers at this level can also recognise the main theme or the author's purpose in a piece of text about a familiar topic, and make a simple connection between several adjacent pieces of information, or between the given information and their own prior knowledge. They can select a relevant page from a small set based on simple prompts, and locate one or more independent pieces of information within short texts.

Level	Percentage of pupils at this level	What pupils can typically do at each level
		Level 1a readers can reflect on the overall purpose and on the relative importance of information (e.g. the main idea vs. non-essential detail) in simple texts containing explicit cues. Most tasks at this level contain explicit cues regarding what needs to be done, how to do it, and where in the text(s) readers should focus their attention.
1b	**OECD**: 99% perform tasks at least at Level 1b **Wales**: 99%	Readers at Level 1b can evaluate the literal meaning of simple sentences. They can also interpret the literal meaning of texts by making simple connections between adjacent pieces of information in the question and/or the text. Readers at this level can scan for and locate a single piece of prominently placed, explicitly stated information in a single sentence, a short text or a simple list. They can access a relevant page from a small set based on simple prompts when explicit cues are present. Tasks at Level 1b explicitly direct readers to consider relevant factors in the task and in the text. Texts at this level are short and typically provide support to the reader, such as through repetition of information, pictures or familiar symbols. There is minimal competing information.
1c	**OECD**: 100% perform tasks at least at Level 1c **Wales**: 100%	Readers at Level 1c can understand and affirm the meaning of short, syntactically simple sentences on a literal level, and read for a clear and simple purpose within a limited amount of time. Tasks at this level involve simple vocabulary and syntactic structures.

Appendix B.5 Mathematics proficiency levels

Level	Percentage of pupils at this level	What pupils can typically do at each level
6	**OECD**: 2% perform tasks at Level 6 **Wales**: 1%	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 together and flexibly translate amongst 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.
5	**OECD**: 11% perform tasks at least at Level 5 **Wales**: 7%	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. Pupils at this level have begun to develop the ability to reflect on their work and to communicate conclusions and interpretations in written form.
4	**OECD**: 29% perform tasks at least at Level 4 **Wales**: 25%	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 representations, 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.
3	**OECD**: 54% perform tasks at least at Level 3 **Wales**: 53%	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.
2	**OECD**: 76% perform tasks at least at Level 2 **Wales**: 79%	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 results.
1	**OECD**: 91% perform tasks at least at Level 1 **Wales**: 94%	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 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 B.6 Science proficiency levels

Level	Percentage of pupils at this level	What pupils can typically do at each level
6	**OECD**: 1% perform tasks at Level 6 **Wales**: 0.4%	At Level 6, pupils can draw on a range of interrelated scientific ideas and concepts from the physical, life, and earth and space sciences and use content, procedural and epistemic knowledge in order to offer explanatory hypotheses of novel scientific phenomena, events and processes or to make predictions. In interpreting data and evidence, they are able to discriminate between relevant and irrelevant information and can draw on knowledge external to the normal school curriculum. They can distinguish between arguments that are based on scientific evidence and theory and those based on other considerations. Level 6 pupils can evaluate competing designs of complex experiments, field studies or simulations and justify their choices.
5	**OECD**: 7% perform tasks at least at Level 5 **Wales**: 5%	At Level 5, pupils can use abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, events and processes involving multiple causal links. They are able to apply more sophisticated epistemic knowledge to evaluate alternative experimental designs and justify their choices and use theoretical knowledge to interpret information or make predictions. Level 5 pupils can evaluate ways of exploring a given question scientifically and identify limitations in interpretations of data sets including sources and the effects of uncertainty in scientific data.
4	**OECD**: 25% perform tasks at least at Level 4 **Wales**: 23%	At Level 4, pupils can use more complex or more abstract content knowledge, which is either provided or recalled, to construct explanations of more complex or less familiar events and processes. They can conduct experiments involving two or more independent variables in a constrained context. They are able to justify an experimental design, drawing on elements of procedural and epistemic knowledge. Level 4 pupils can interpret data drawn from a moderately complex data set or less familiar context, draw appropriate conclusions that go beyond the data and provide justifications for their choices.

3	**OECD**: 52% perform tasks at least at Level 3 **Wales**: 52%	At Level 3, pupils can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context. Level 3 pupils are able to distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim.
2	**OECD**: 78% perform tasks at least at Level 2 **Wales**: 80%	At Level 2, pupils are able to draw on everyday content knowledge and basic procedural knowledge to identify an appropriate scientific explanation, interpret data, and identify the question being addressed in a simple experimental design. They can use basic or everyday scientific knowledge to identify a valid conclusion from a simple data set. Level 2 pupils demonstrate basic epistemic knowledge by being able to identify questions that can be investigated scientifically.
1a	**OECD**: 94% perform tasks at least at Level 1a **Wales**: 96%	At Level 1a, pupils are able to use basic or everyday content and procedural knowledge to recognise or identify explanations of simple scientific phenomena. With support, they can undertake structured scientific enquiries with no more than two variables. They are able to identify simple causal or correlational relationships and interpret graphical and visual data that require a low level of cognitive demand. Level 1a pupils can select the best scientific explanation for given data in familiar personal, local and global contexts.
1b	**OECD**: 99% perform tasks at least at Level 1b **Wales**: 100%	At Level 1b, pupils can use basic or everyday scientific knowledge to recognise aspects of familiar or simple phenomena. They are able to identify simple patterns in data, recognise basic scientific terms and follow explicit instructions to carry out a scientific procedure.

Appendix C

Appendix C.7 Proportions of pupils at each SES quintile by school medium

School medium		English-medium	Welsh-medium
SES quintile band			
Lowest	No. of pupils	470	105
	% within group	21.4	15.4
Second Lowest	No. of pupils	441	134
	% within group	20.1	19.7
Middle	No. of pupils	437	137
	% within group	19.9	20.1
Second highest	No. of pupils	425	150
	% within group	19.4	22.1
Highest	No. of pupils	420	154
	% within group	19.2	22.6
Total	No. of pupils	**2193**	**680**
	% within group	**100**	**100**

Missing (264, i.e.156 English- and 108 Welsh-medium) not included.

Source: PISA 2018 school census matched database

Appendix D

Appendix D.8 School medium sample breakdown by language of assessment

Category	English-medium school	Welsh-medium school
Language of assessment	**No. of pupils**	**No. of pupils**
English	2349	338
Welsh	0	450
Total	**2349**	**788**

Total N = 3,137

Source: PISA 2018 school census matched database

Evidence for excellence in education

The Mere, Upton Park, Slough, Berks SL1 2DQ

T: +44 (0)1753 574123 • F: +44 (0)1753 691632 • enquiries@nfer.ac.uk

www.nfer.ac.uk

NFER ref. WPAB

ISBN. 978-1-912596-22-5