

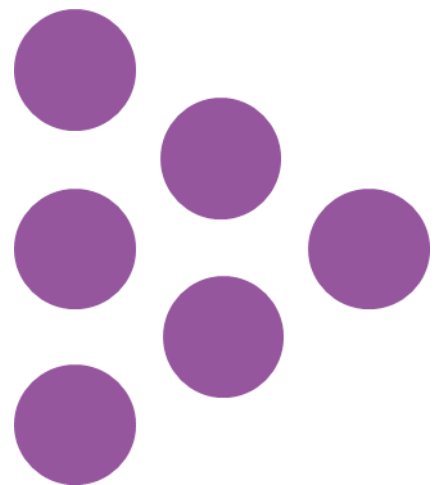


Llywodraeth Cymru
Welsh Government

PISA 2018 Additional Analyses:

Mathematics and Science Item Analysis in Wales

National Foundation for Educational Research (NFER)



PISA 2018 Additional Analyses: Mathematics and Science Item Analysis in Wales

Maria Galvis, Chris Hope, Juliet Sizmur and Rebecca Wheeler

Published in November 2020

By the National Foundation for Educational Research,

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

www.nfer.ac.uk

© 2020 National Foundation for Educational Research
Registered Charity No. 313392

ISBN: 978-1-912596-23-2

How to cite this publication:

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

Contents

List of figures	1
List of tables	3
Executive Summary	4
1 Introduction	6
2 Items analysed	7
2.1 Criteria for item selection	7
2.2 Item analysis context	8
2.2.1 Item domains	8
2.2.2 Trend analysis	9
2.2.3 International context	9
2.2.4 Subgroup analysis	10
3 Mathematics	12
3.1 'Omitted' mathematics items	13
3.1.1 Process categories	13
3.1.2 Content categories	14
3.1.3 Context categories	15
3.1.4 Response format categories	16
3.2 'No credit' mathematics items	17
3.2.1 Process categories	18
3.2.2 Content categories	18
3.2.3 Context categories	19
3.2.4 Response format categories	20
3.3 'Partial credit' mathematics items	21
4 Science	25
4.1 'Omitted' science items	25
4.1.1 Competency categories	26
4.1.2 Knowledge categories	27
4.1.3 Content categories	28
4.1.4 Context categories	29
4.1.5 Response format categories	30
4.2 'No-credit' science items	31

4.2.1 Competency categories	31
4.2.2 Knowledge categories	32
4.2.3 Content categories	33
4.2.4 Context categories	34
4.2.5 Response format categories	35
4.3 'Partial credit' science items	36
5 Subgroup analysis	39
5.1 Analysis background	39
5.2 Gender	40
5.3 Language of Assessment	41
5.4 Medium of schooling	42
5.5 Socioeconomic status	43
5.6 Lowest and highest national attainment quartile	44
6 Conclusions	46
6.1 Subject knowledge and skills	46
6.2 Assessment skills	47
References	48
Appendix A: Item composition of the PISA 2018 assessment.	49

List of figures

Figure 2. 1 Trends in average mathematics PISA scores for Wales and comparator countries	9
Figure 2. 2 Trends in average science PISA scores for Wales and comparator countries	10
Figure 3. 1 Number of omitted mathematics analysis items in each process category	14
Figure 3. 2 Number of omitted mathematics analysis items in each content category	15
Figure 3. 3 Number of omitted mathematics analysis items in each context category	16
Figure 3. 4 Number of 'omitted' mathematics analysis items in each response format category..	17
Figure 3. 5 Number of no-credit mathematics analysis items in each process category	18
Figure 3. 6 Number of no-credit mathematics analysis items in each content category	19
Figure 3. 7 Number of no-credit mathematics analysis items in each context category	20
Figure 3. 8 Number of no-credit mathematics analysis items in each response format category .	20
Figure 3. 9 Scoring pattern of the multi-mark mathematics items in Wales.....	24
Figure 4. 1 Number of omitted science analysis items in each competency category.....	27
Figure 4. 2 Number of omitted science analysis items in each knowledge category	28
Figure 4. 3 Number of omitted science analysis items in each content category	29
Figure 4. 4 Number of omitted science analysis items in each context category	30
Figure 4. 5 Number of omitted science analysis items in each response format category	31
Figure 4. 6 Number of no-credit science analysis items in each competency category	32
Figure 4. 7 Number of no-credit science items analysed in each knowledge category	33
Figure 4. 8 Number of no-credit science analysis items in each content category	34
Figure 4. 9 Number of no-credit science analysis items in each context category	34
Figure 4. 10 Number of no-credit science analysis items in each response format category	35
Figure 4. 11 Scoring pattern for the ten multi-mark science items in Wales.....	38
Figures 5. 1 Average omission and no-credit rates of the analysis items by gender.....	40
Figures 5. 2 Average omission and no-credit rates of the analysis items by language of assessment.....	41

Figures 5. 3 Average omission and no-credit rates of the analysis items by medium of schooling	42
Figures 5. 4 Average omission and no credit rates of the analysis items by socioeconomic status	44

List of tables

Table 2. 1 Number of items omitted or where learners received no credit in Wales (mathematics and science aggregated).....	8
Table 3. 1 Distribution of 'partial credit' mathematics items analysed in each category	22
Table 3. 2 Classification and facility of the seven multi-mark mathematics items in Wales.....	23
Table 4. 1 Distribution of 'partial credit' science items analysed in each category	36
Table 4. 2 Classification and facility of the ten multi-mark science items in Wales	37
Table 5. 1 Item subgroup analysis by lowest and highest attainment quartile.....	45

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). The assessment takes place every three years and examines how 15-year-olds can apply what they have learned in school to real-life situations, rather than their ability to remember facts. Schools and learners from 79 participating countries and economies participated in PISA 2018. Learners in Wales have participated in PISA since 2006.

The purpose of this report is to explore the strengths and weaknesses of learners in Wales in science and mathematics in PISA 2018. We did this by identifying what type of items (questions) learners tended to skip or fail to gain credit on. The items we analysed in this report followed the same selection criteria as in previous studies to allow for comparability across years: items omitted by more than 15 per cent of learners, items in which 55 per cent of learners gained no credit, and multi-mark items. In addition, we examined whether performance of learners in Wales was different from a group of comparator countries and the OECD in general.

Further analysis was undertaken to identify if there were noticeable differences in the pattern of responses by gender, language of assessment, medium of schooling, socioeconomic status and attainment quartile.

Some of the key findings of this report are:

- Learners in Wales tended to omit and fail to gain credit on more mathematics items than science items (proportional to the pool of items in each subject).
- The number of no-credit items in both mathematics and science was higher in Wales than the OECD average.
- The number of omitted items in Wales was similar to the OECD average for mathematics and lower for science. This represents an improvement from 2015, and suggests that in PISA 2018 learners in Wales were more confident in attempting science questions, but continued to failed to gain credit in more items than their international peers do.
- Compared to the OECD, learners in Wales had fewer omitted science items related to living systems. Across comparator countries, Wales was the only country where no items set in personal contexts met the omission criteria (as opposed to items set in local/national and global contexts).
- Compared to high performing countries, learners in Wales were less confident in attempting mathematics items related to quantification, items that required them to perform mathematical procedures needed to derive results, and problems set in the natural world and related to science and technology.

- In science, learners in Wales struggled more than their international peers to gain credit on items that asked them to explain phenomena scientifically (this competency includes generating hypotheses and predicting possible changes, recalling and using theories and facts).
- Compared to high performing countries, learners in Wales struggled most to gain credit in mathematics items that asked them to formulate situations mathematically (identify the mathematical aspects of a problem situated in a real-world context), as well as items related to interpretation and presentation data, probability and statistics.
- Wales' performance in multi-mark items relative to the OECD average improved for both subjects in PISA 2018 compared to 2015, indicating that learners in Wales are giving better or more complete responses to these types of question.
- In Wales and comparator countries, the highest incidence of omitted and no-credit items corresponds to open constructed-response items, following the same pattern as PISA 2015.
- In both mathematics and science, more learners in Wales tended to gain no credit in the complex multiple-choice items than in the OECD, on average.
- The difference between learners with high and low socio-economic background, in terms of omission rates, was more pronounced than the differences found to exist between genders, medium of schooling, and language of assessment.
- The gap in performance between learners in the highest and lowest attainment quartiles was larger for science than for mathematics in terms of both omission and no-credit rates.

1 Introduction

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 learners from 79 participating countries and economies participated in PISA 2018. Wales has participated in PISA since 2006.

PISA is designed to examine how 15-year-olds can apply what they have learned in school to real-life situations. Learners are asked to use their skills of reasoning, interpretation and problem solving, rather than simply remembering facts.

In Wales, 3165 15-year-old learners in 107 schools completed a two-hour computer based assessment of reading, mathematics and science, and a pupil questionnaire. Details of the item composition of the PISA 2018 assessment are in Appendix A.

Further details on the development of the PISA survey, what PISA measures, the PISA scales, frameworks and proficiently levels can be found in the PISA 2018 Technical Report (OECD), and the PISA 2018 Assessment and Analytical Framework (OECD, 2019a).

The purpose of this report is to explore the strengths and weaknesses in maths and science skills among learners in Wales in PISA 2018. In order to do this, in-depth item analysis was conducted of items (questions) that met certain criteria in terms of omission rates and zero credit as outlined in Section 2.1.

The features of the identified items were considered to examine patterns of performance, and whether performance of learners in Wales was different from five comparator countries and performance in the OECD generally. This item analysis builds on similar analyses of PISA 2015 (Andrade *et al.*, 2017) and 2012 (NFER, 2015). Where possible, we compare the 2018 and 2015 analyses to identify changes in performance.

In PISA 2018, the reading assessment adopted an adaptive approach, whereby learners were assigned items based on their performance in earlier units (OECD, 2019b). Because of this, facility levels would not be comparable to those of 2015 and, as a result, the item analysis in this report will focus on maths and science only.

Chapter 2 provides further details of the analysis. Chapters 3 and 4 present the analysis of PISA performance of learners in Wales, five comparator countries and performance in the OECD generally, in mathematics and science. Chapter 5 presents analysis of the performance of different groups of learners in Wales (gender, language of assessment, medium of schooling, socioeconomic status, and lowest and highest attainment quartile). Finally, conclusions are presented in Chapter 6.

2 Items analysed

By identifying the features of the items omitted by a large proportion of learners in Wales and the items that were attempted by a large proportion of learners but no credit was gained, we can better understand learners' PISA assessment skills. These insights could then be used to inform teaching and learning strategies aimed at addressing the weaknesses of learners in Wales and, ultimately, lead to an improvement in Wales's performance in PISA science and mathematics domains.

Learners in Wales were presented with 429 items in PISA 2018; 57 per cent were reading items (major domain), 27 per cent were science and 16 per cent were mathematics. All 429 items had at least three per cent of learners who attempted the item and failed to receive credit for it, as was the case in 2015. Across all subjects, seven per cent of the items had zero omission rate (i.e. all learners attempted them). Responses to these 32 items with no omissions shows that, on average, 40 per cent of learners failed to gain credit on them. In PISA 2015, the average percentage of learners receiving no credit on items with zero omission rate was 50 per cent.

This report analyses only mathematics and science items due to challenges in comparability of reading items across cycles as a result of the introduction of adaptive testing in PISA 2018.

2.1 Criteria for item selection

Learners in Wales were presented with 70 mathematics items and 115 science items (185 in all). Table 2.1 below shows the number of items and percentage of learners that omitted and/or gained no credit for mathematics and science items at each percentage threshold.

Items that met the following criteria were selected for analysis, in line with analyses of PISA 2015 and 2012 (Andrade et al., 2017; NFER, 2015):

- items which were 'omitted' (or skipped) by 15 per cent or more of learners in Wales
- items which were attempted but 55 per cent or more of learners in Wales failed to gain credit (no credit items)
- multi-mark (partial credit) items.

As a result, 13 items were identified where over 15 per cent of learners omitted the item and there were 62 items where over 55 per cent of learners gained no credit. In addition, all 17 science and mathematics multi-mark items were analysed (7 mathematics and 10 science), some of which overlapped with those identified using the criteria described above.

Table 2.1 Number of items omitted or where learners received no credit in Wales (mathematics and science aggregated)

Percentage of learners omitting the item	Number of items (out of 185)	Percentage of learners gaining no credit	Number of items (out of 185)
Over 5%	49	Over 20%	167
Over 10%	25	Over 30%	143
Over 15%	13	Over 40%	110
Over 20%	8	Over 50%	83
Over 25%	4	Over 55%	62
Over 30%	1	Over 60%	48
Over 35%	0	Over 70%	20
-	-	Over 80%	7

Note: Percentages and number of items in bold indicates the agreed analysis threshold

Source: PISA 2018 database

In total, we analysed 78 items in this report; 65 of them fell into one of the three item-selection criteria (i.e. omission rate above 15 per cent, no credit rate above 55 per cent, or multi-mark item), 12 items fell into two of the three possible criteria, and one mathematics item fell into all three categories.

2.2 Item analysis context

2.2.1 Item domains

Of the 78 items selected for analysis:

- 34 were mathematics items (44 per cent of the total number of analysed items)
- 44 were science items (56 per cent of the total number of analysed items)

From the total pool of 185 mathematics and science items administered to learners in Wales, 38 per cent correspond to mathematics items and 62 per cent to science items. In line with this, the mathematics items analysed in this report are slightly over-represented and the science items analysed are slightly under-represented.

2.2.2 Trend analysis

Where possible, we provide a descriptive commentary on the most notable variations in performance between PISA 2018 and PISA 2015. GCSE specifications have changed since PISA 2015. The first set of reformed GCSEs were first taught from September 2015 (maths and maths numeracy qualifications), science qualifications in September 2016, and Welsh second language from September 2017¹.

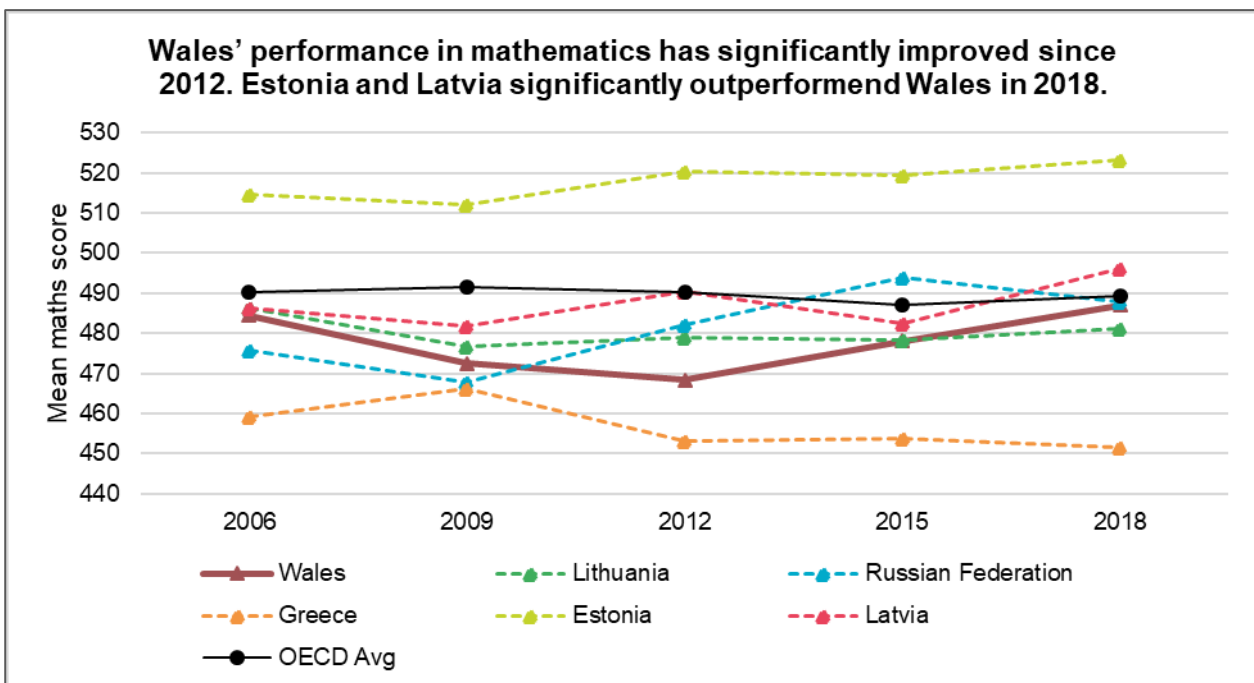
2.2.3 International context

The results of learners in Wales were compared with the OECD average² and five comparator countries for each of mathematics and science. For consistency in reporting, the comparator countries selected were the same countries selected for comparison in trend analysis in the PISA 2018 Wales national report (Sizmur et al., 2019), with the addition of Estonia due to its rapid improvement over successive PISA cycles. These were:

- Mathematics: Estonia, Greece, Latvia, Lithuania, Russia
- Science: Estonia, Iceland, Latvia, Spain, Sweden.

The performances of Wales and the comparator countries are presented for mathematics and science in Figures 2.1 and 2.2 below.

Figure 2. 1 Trends in average mathematics PISA scores for Wales and comparator countries



*OECD Average is AV37 for 2006, 2012, 2015, 2018, and AV36 for 2009 Source: PISA 2018 database

¹ More information in <https://www.qualificationswales.org/english/qualifications/gcses-and-a-levels/gcses/>

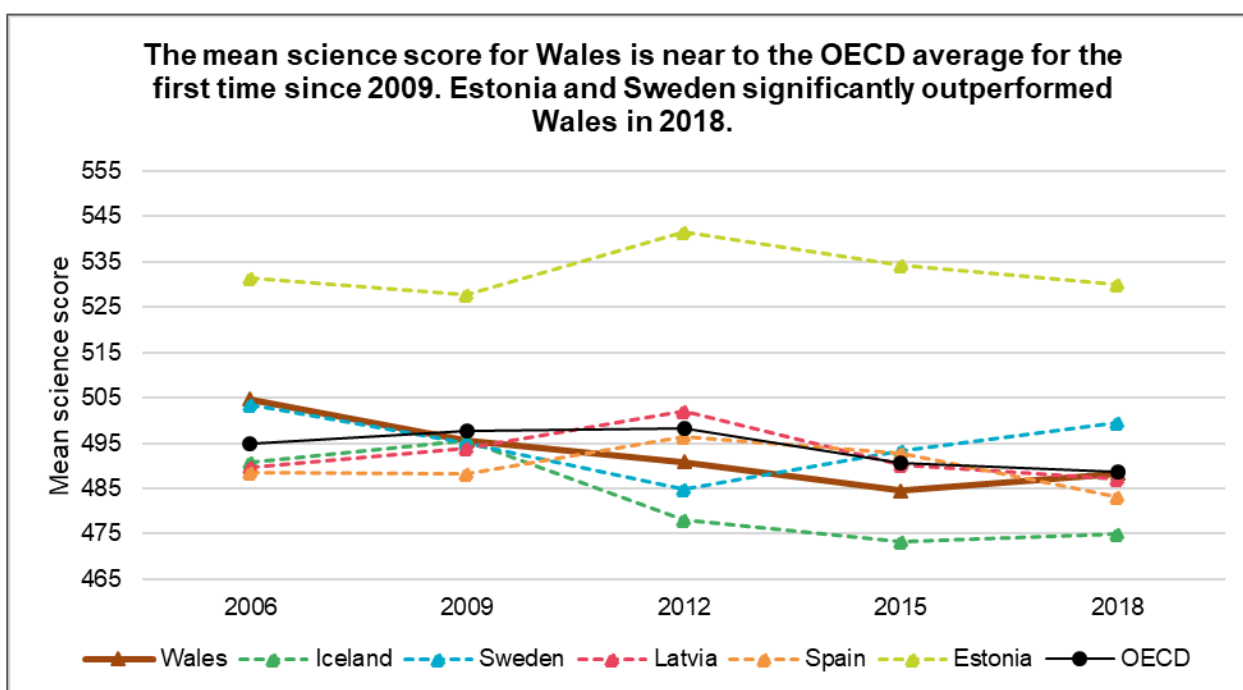
² The OECD average used excludes Chile, Mexico and Colombia as learners in these three countries received an easier version of the assessment.

In PISA 2018, the mean score for Wales in mathematics was similar to the OECD average for the first time and was significantly better than in 2009 and 2012.

In PISA 2018, the mean score for Wales in science was not significantly different from the OECD average, which represents a comparative improvement from the last two PISA cycles when science scores in Wales were significantly lower than the OECD average.

Estonia ranked highest among European countries in science, mathematics and reading in PISA 2018, with all scores significantly higher than the OECD average. Mathematics performance has improved steadily since Estonia joined PISA in 2006. Science performance in the same period has remained stable and high. Figures 2.1 and 2.2 show that the other comparator countries had similar performances to Wales in 2009 but have followed different trajectories through successive cycles.

Figure 2. 2 Trends in average science PISA scores for Wales and comparator countries



*OECD Average is AV37 for 2006, 2012, 2015, 2018, and AV36 for 2009 Source: PISA 2018 database

2.2.4 Subgroup analysis

A subgroup analysis was conducted to identify if there were noticeable differences in the pattern of responses by:

- gender
- medium of schooling (Welsh or English)
- language of assessment (Welsh or English)

- socioeconomic status (i.e. the 30 per cent most deprived learners in Wales according to the PISA index of economic, social and cultural status (ESCS) compared with the 70 per cent least deprived)
- lowest and highest national attainment quartile.

3 Mathematics

Key findings

- In PISA 2018, learners in Wales achieved a mean score of 487 in mathematics which, for the first time, was not significantly below the OECD average (489). Compared with the comparator countries, Wales' score was similar to Lithuania and Russian Federation, higher than Greece, and lower than Estonia and Latvia. Estonia achieved the highest mathematics score amongst European countries in 2018.
- Fourteen per cent of mathematics items were not attempted by more than 15 per cent of the learners in Wales. More than 55 per cent of learners in Wales failed to gain credit in 40 per cent of mathematics items. The amount and distribution of omitted and no-credit items in the mathematics domain was similar to PISA 2015.
- The incidence of skipped items in higher performing comparator countries Estonia and Latvia was similar than in Wales. In terms of items for which more than 55 percent of learners gained no credit, the performance of Wales was lower to that of Estonia.
- Compared with higher performing comparator countries, learners in Wales tended to skip more items in the *employ* category (which require learners to apply mathematical concepts, facts, procedures and reasoning to reach mathematical conclusions), in the *quantity* category (which involve understanding measurements, counts, magnitudes, units, numerical trends and patterns), and in the *scientific* category (which relate to the application of mathematics to the natural world and issues and topics related to science and technology).
- The no credit analysis showed that, compared with the OECD on average and high performing Estonia, learners in Wales had problems gaining credit in items that required them to *formulate situations mathematically*, items in the *uncertainty and data* category, and items set in *occupational* and *societal* contexts.
- In 2018, Wales had a similar incidence of no-credit mathematics items in the *simple multiple-choice* category as compared with the OECD average. This represents an improvement from 2015, when this rate was higher for Wales than for the OECD.
- There was an improvement in Welsh facility levels of multi-mark mathematics items as compared with the OECD average from PISA 2015. This suggests that in 2018 learners in Wales found fewer maths items more difficult than their international peers.

3.1 ‘Omitted’ mathematics items

Learners in Wales were presented with 70 mathematics items in PISA 2018. Ten items were skipped by more than 15 per cent of the learners, similar to the OECD average and similar to the number³ of items omitted in comparator countries. In order to compare changes since 2015, we focused on the proportion of items omitted in each category. In 2018, 14 per cent of mathematics items met the omission criteria, compared with 16 per cent in 2015.

PISA classifies mathematical literacy in terms of following three interrelated aspects (OECD, 2019a), which are described and analysed in more detail in the sections that follow:

- the mathematical **processes** that describe what individuals do to connect the problem with the solution,
- the mathematical **content**
- the **contexts** in which the assessment items are located.

Items are also classified into three question formats: open constructed-response, closed constructed-response and selected-response (simple and complex multiple-choice) items.

3.1.1 Process categories

PISA mathematics items are assigned to one of three mathematical processes: i) *formulating* situations mathematically; ii) *employing* mathematical concepts, facts, procedures and reasoning; and iii) *interpreting*, applying and evaluating mathematical outcomes.

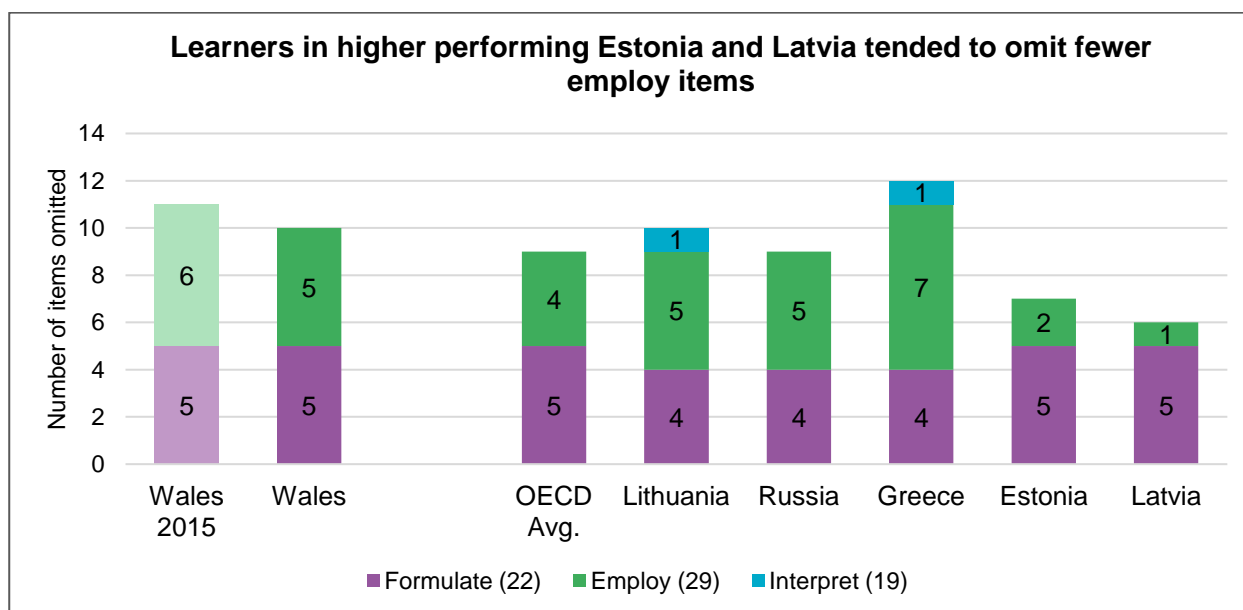
Figure 3.1 shows the distribution of the 10 omitted items⁴ across process categories. In Wales, around one quarter of the *formulate* items and around one fifth of the *employ* items had omission rates of over 15 per cent. This was in line with the OECD average and similar to those in the comparator countries; except for learners from Estonia and Latvia (the highest performing comparator countries) who omitted fewer *employ* items, and learners from Greece, who omitted more *employ* items.

An exploration into the reasons why learners omitted items is beyond the scope of this report, but could be explained by low motivation in general, or, specifically for low-stakes assessments, inadequate time or item difficulty.

³ We have considered the number of total items different when they differed by five or more. Since the item classifications can have three or four subcategories, this criterion corresponds to an average difference at least two item per subcategory.

⁴ Note that in this section we will sometimes refer to items that have been skipped by more than 15 per cent of learners as ‘omitted items’. This is for the sake of brevity, but it should be interpreted as ‘items that have not been attempted by more than 15 per cent of learners’.

Figure 3. 1 Number of omitted mathematics analysis items in each process category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items

Source: PISA 2018 database

Learners across comparator countries tended to omit more items that required to *formulate situations mathematically* compared to the other mathematics process categories; this was in line with PISA 2015. The *formulate* items involve learners recognising and identifying opportunities to use their mathematics skills in a problem solving context and then providing a structure to analyse and set up and solve a problem. The *employ* items require learners to apply mathematical concepts, facts, procedures and reasoning to reach mathematical conclusions.

As in 2015, no items in the *interpreting* process were omitted by more than 15 per cent of learners in Wales. This process involves reflecting on mathematical solutions, results, or conclusions and interpreting them in the context of real-life problems and can require learners to evaluate mathematical outcomes.

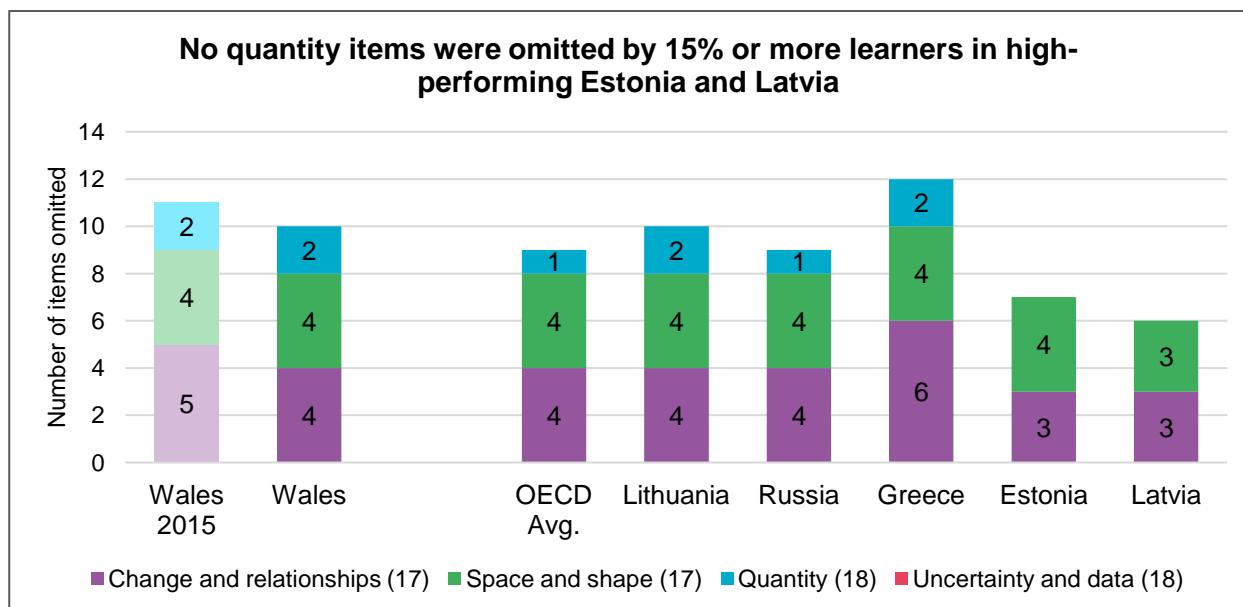
3.1.2 Content categories

PISA classifies mathematical knowledge (and the ability to apply that knowledge to the solution of problems) into four content categories, which reflect what is typically found in national mathematics curricula. Each mathematics item falls into one of these categories: i) *change and relationships*, ii) *space and shape*, iii) *quantity*, and iv) *uncertainty and data*.

The distribution of the omitted mathematics items by content category is presented in Figure 3.2. The *change and relationships* and *space and shape* categories had in proportion more omitted items than the *quantity* and *uncertainty and data*. This finding

was consistent across comparator countries. None of the *uncertainty and data* items met the criteria for analysis in Wales or comparator countries.

Figure 3. 2 Number of omitted mathematics analysis items in each content category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

Source: PISA 2018 database

In Wales, around one quarter of items in the *change and relationships* and *space and shape* categories had omission rates of over 15 per cent, similar to 2015. A smaller proportion of items in the *quantity* category met the omission criteria (11 per cent).

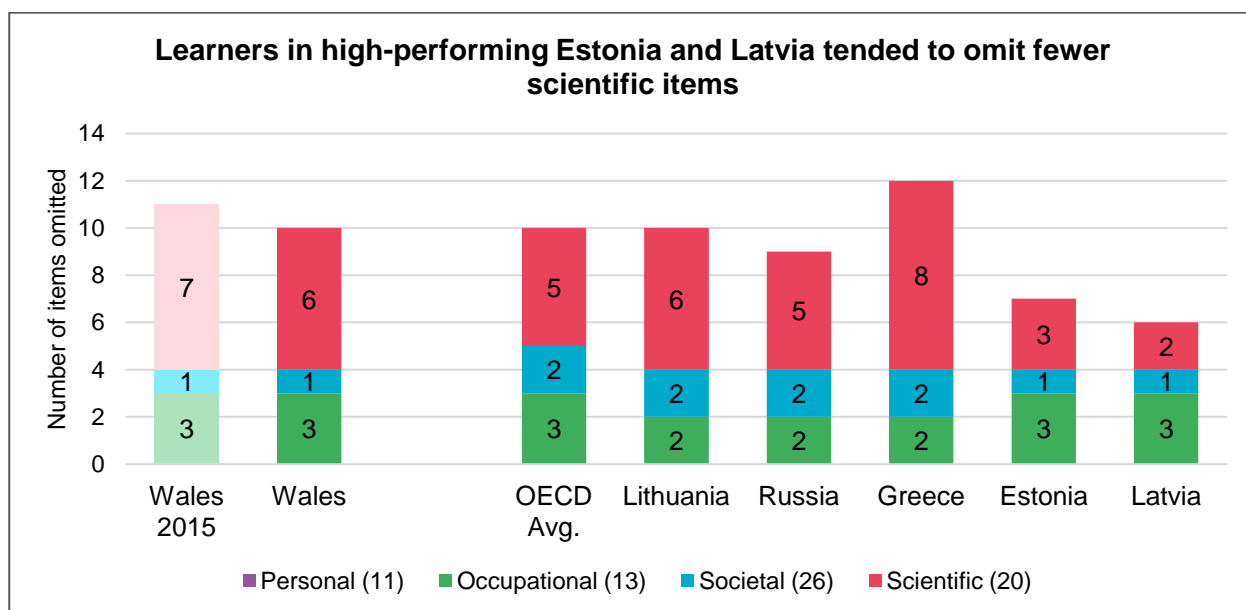
Wales had a similar number of omitted items across all content categories compared with the OECD average and most of the comparator countries. Latvia and Estonia (the highest performing comparator countries) did not have any *quantity* items omitted by more than 15 per cent of learners. *Quantity* items assess number sense, multiple representations of numbers, mental calculation, estimation and assessment of reasonableness of results (OECD, 2019a).

3.1.3 Context categories

There are four context categories into which PISA mathematics assessment items are classified: i) *personal*, ii) *occupational*, iii) *societal* and iv) *scientific*.

The distribution of omitted mathematics items by context category is presented in Figure 3.3, which shows that the majority of mathematics items that were not attempted by over 15 per cent of learners in Wales correspond to *scientific* context (30 per cent), followed by *occupational* context (23 per cent) and a small proportion of items in the *societal* context (8 per cent). This was in line with the OECD average.

Figure 3. 3 Number of omitted mathematics analysis items in each context category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

Source: PISA 2018 database

With the exception of Latvia and Estonia, the highest proportion of omitted items across comparator countries fall into the *scientific* context category. Problems classified in the *scientific* category relate to the application of mathematics to the natural world and issues and topics related to science and technology. These may include areas such as weather or climate, ecology, medicine, space science, genetics, measurement and the world of mathematics itself (OECD, 2019a).

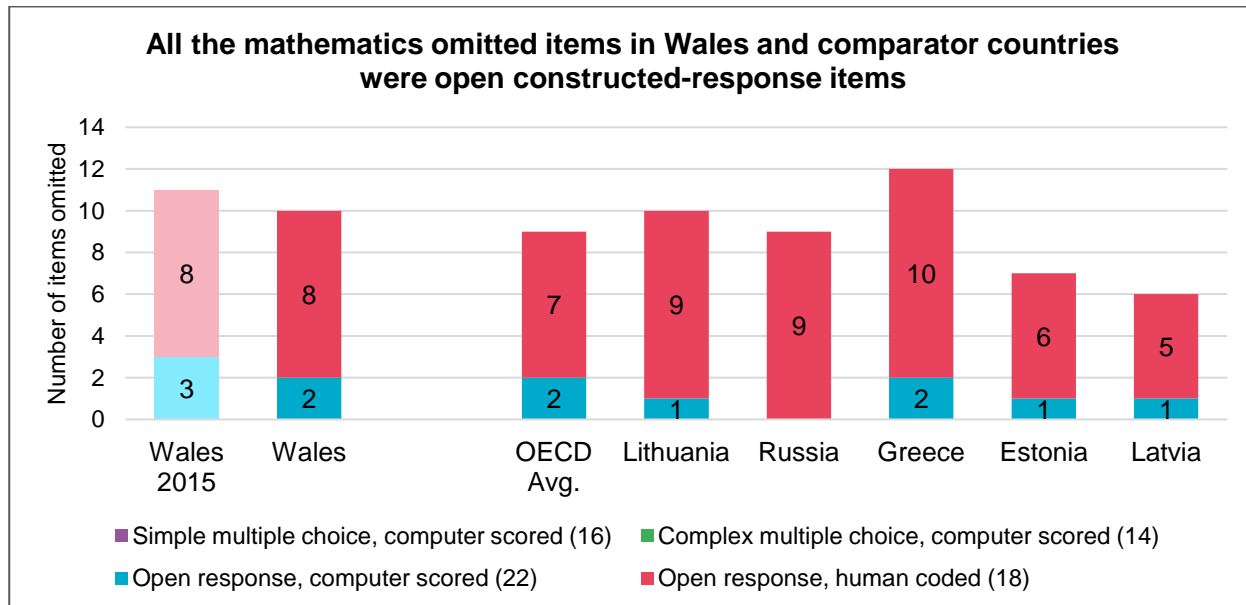
Learners in Greece (the lowest performer comparator country) had more omitted items than Wales in the *scientific* category, whilst learners from Estonia and Latvia (the highest performing comparator countries) had fewer omitted items than Wales in the same category.

3.1.4 Response format categories

The mathematics items are classified into three types of response format: open constructed-response items, closed constructed-response, and selected-response items (simple or complex multiple-choice).

Open response items require an extended written response. Learners may be asked to show the steps taken or to explain how the answer was reached. Selected response items require learners to choose one or more responses from a number of response options. In this report, we analyse four subtypes of formats: i) *open response - computer scored*, ii) *open response – human coded*, iii) *simple multiple-choice – computer scored*, and iv) *complex multiple-choice – computer scored*.

Figure 3. 4 Number of ‘omitted’ mathematics analysis items in each response format category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

Source: PISA 2018 database

Figure 3.4 shows the number of omitted items in each response format. All the items omitted by more than 15 per cent of learners in Wales were *open constructed-response* items, as in PISA 2015. The highest proportion of omitted items corresponded to *open response - human coded* (44 per cent of the pool of this type of items), in line with the OECD average.

Learners in higher-performing Estonia and Latvia were more likely to attempt the *open response* items than learners in other comparator countries and the OECD on average, whilst learners in Greece (the lowest performing comparator country) tended to skip more of these type of items. In Russia, none of the *open response - computer scored* were omitted by more than 15 per cent of learners.

3.2 ‘No credit’ mathematics items

In PISA 2018, there were 28 mathematics items where more than 55 per cent of learners in Wales failed to gain credit. This represents 40 per cent of all mathematics items, which is similar to the 43 per cent in PISA 2015.

Wales had a higher proportion of no-credit items than the OECD average (33 per cent) and the highest-achieving comparator country Estonia (24 per cent). Russia and Latvia had a similar number of no-credit mathematics items to Wales.

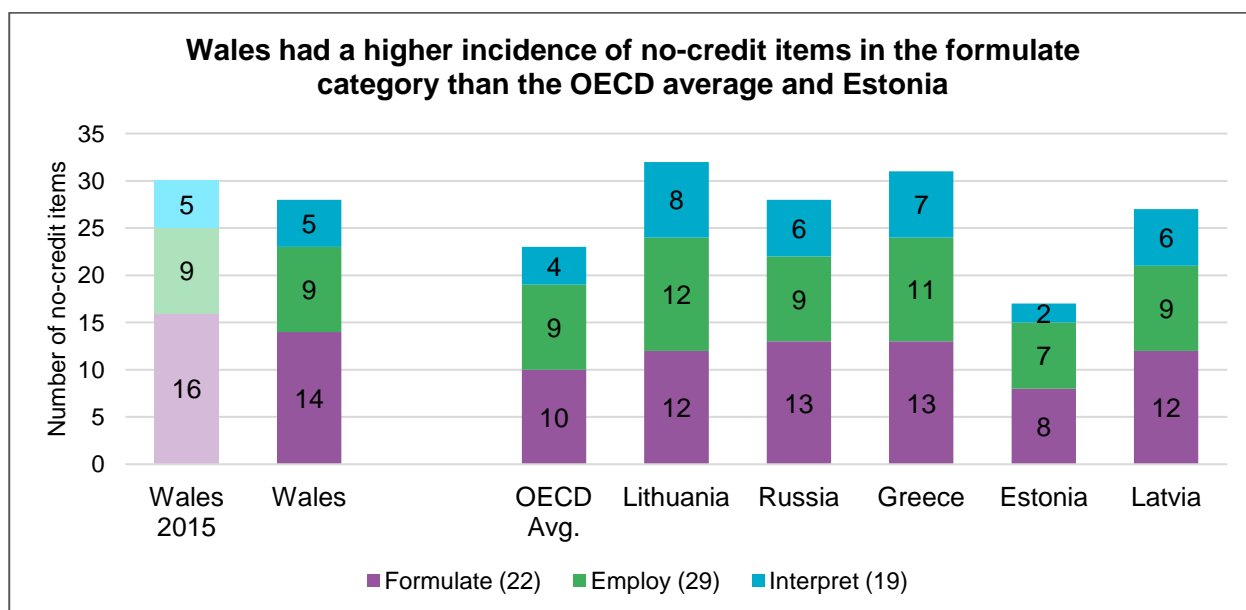
Figures 3.7 to 3.10 show the distribution of these items according to the three aspects of mathematical literacy (process, content and context), and according to the items' response format.

3.2.1 Process categories

As in PISA 2015, process items that assessed the *formulating situations mathematically* process had the highest incidence of no-credit across all comparator countries. More than 55 per cent of learners in Wales failed to gain credit in almost two thirds of items in this category. This was followed by *employing mathematical concepts* items (31 per cent) and *interpreting, applying and evaluating mathematical outcomes* items (26 per cent).

Learners in Wales had a higher incidence of no-credit items⁵ in the *formulate* category than the OECD average, Lithuania, and the higher-performing comparator countries of Estonia and Latvia. Lithuania and Greece, the low-achieving comparator countries, had higher levels of no-credit items than Wales in the *employ* and *interpret* categories.

Figure 3. 5 Number of no-credit mathematics analysis items in each process category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

Source: PISA 2018 database

3.2.2 Content categories

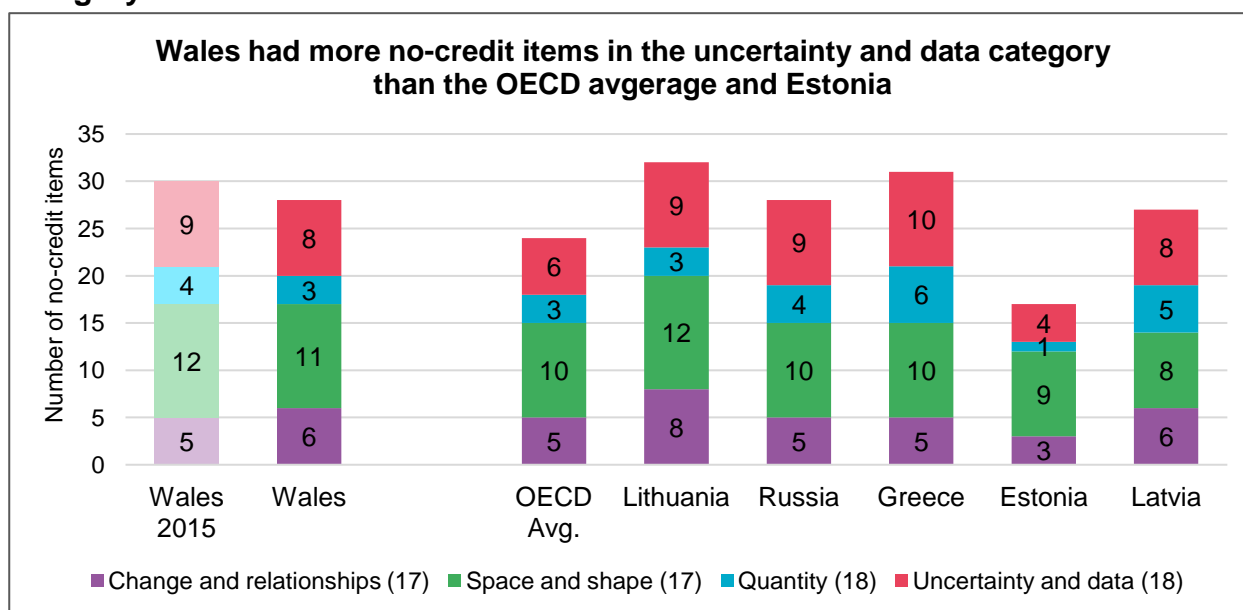
Figure 3.6 shows the number of no-credit items in each content category. The highest proportion of no-credit items in Wales assessed *space and shape* (65 per cent of the total

⁵ Note that in this section, for the sake of brevity, we will sometimes refer to items where more than 55 per cent of learners gained no credit as 'no-credit items'.

of items in this category), followed by *uncertainty and data* (44 per cent), *change and relationships* (35 per cent) and *quantity* (17 per cent). This distribution was in line with the OECD average, except for items in the *uncertainty and data* category, in which the incidence of no-credit items was lower than in Wales.

Space and shape items had the highest incidence of no-credit analysis items in all the comparison countries, as in PISA 2015. Although no *uncertainty and data* items were omitted by more than 15 per cent of pupils in Wales and comparator countries, the amount of no-credit items in this category ranged from 22 to 56 per cent across the countries included in this analysis.

Figure 3. 6 Number of no-credit mathematics analysis items in each content category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

Source: PISA 2018 database

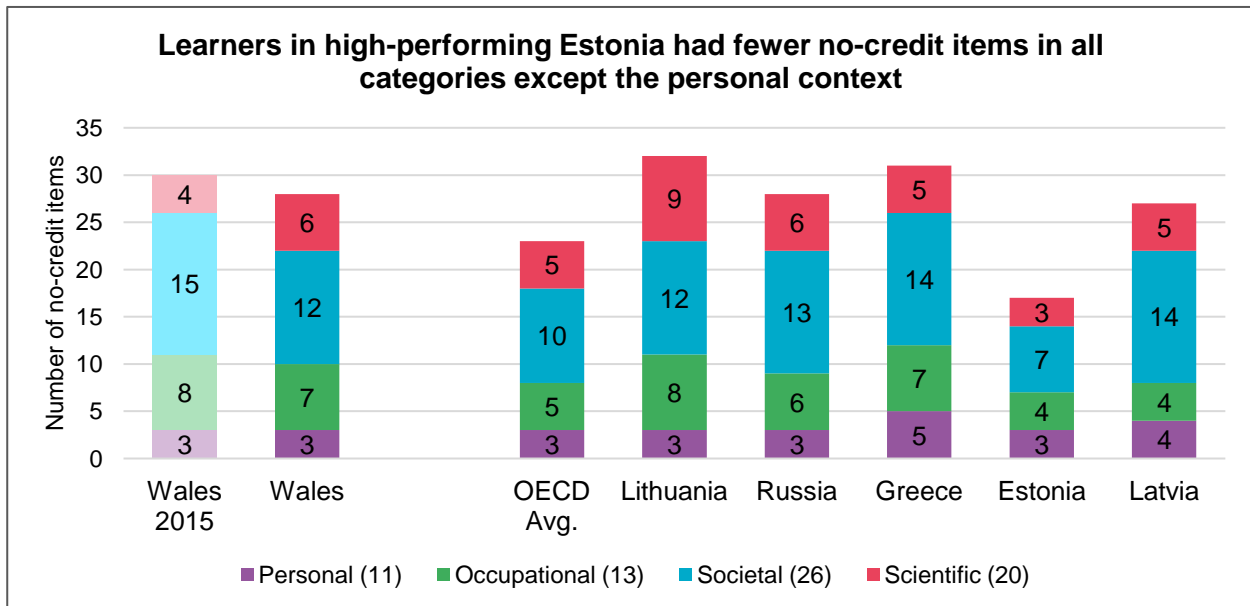
3.2.3 Context categories

The highest proportion of no-credit items in Wales corresponded to the *societal* context (46 per cent of this type of item), similar to the OECD average and comparator countries. This was a lower proportion than in PISA 2015, when learners in Wales gained no credit on 58 per cent of *societal* items. Problems classified in this category may involve contexts such as voting systems, public transport, government, public policies, demographics, advertising, national statistics and economics (OECD, 2019a).

There were no *personal* items omitted by more than 15 per cent of learners in Wales or any comparator country. The proportion of *personal* items in which learners in Wales gained no credit (27 per cent) was lower than low-performing Greece. High-achieving

Estonia had fewer no-credit items than Wales in the *scientific, societal* and *occupational* categories.

Figure 3. 7 Number of no-credit mathematics analysis items in each context category

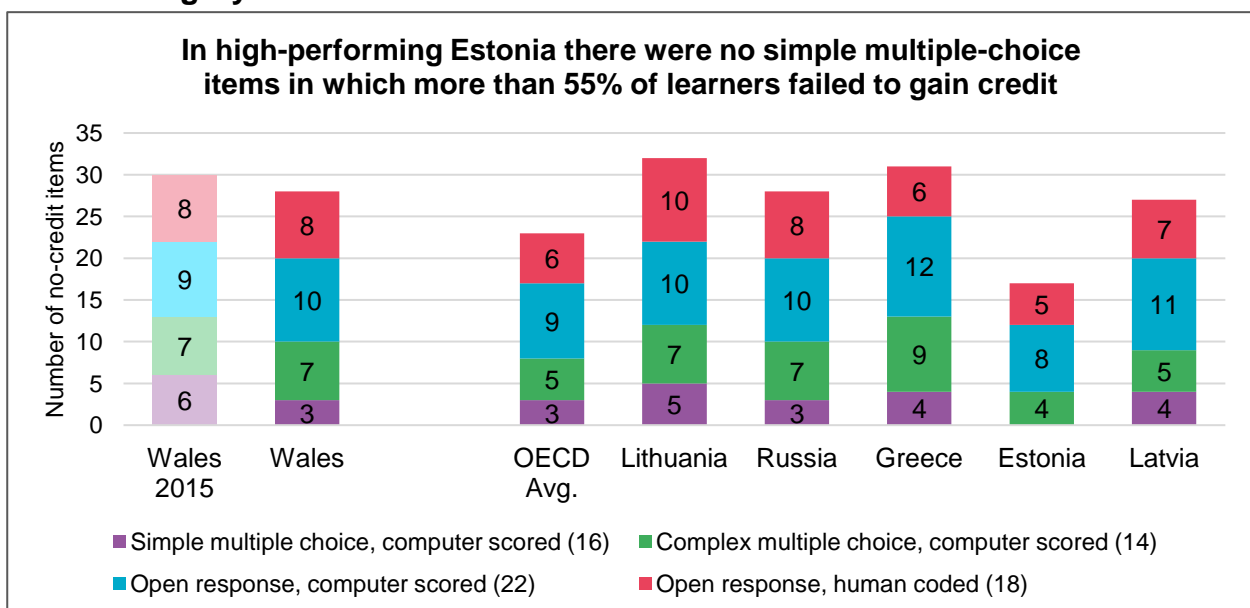


Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

3.2.4 Response format categories

The number of no-credit mathematics analysis items in each response format category is shown in Figure 3.8.

Figure 3. 8 Number of no-credit mathematics analysis items in each response format category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 69 mathematics items.

In Wales, the highest incidence of no-credit items analysed corresponded to *complex multiple-choice*, with half of items in this category being items for which over 55 per cent of learners gained no credit. This was followed by *open constructed-response* format items (around 45 per cent of both computer human scored items). The incidence of no-credit items (19 per cent) for the *simple multiple-choice* category was lower than those of the other types of response formats.

Learners in Wales had a higher incidence of no-credit items in the *open response human coded* and the *complex multiple-choice* categories than the OECD average and Estonia. In 2015, Wales had a higher incidence of no-credit items in the *simple multiple-choice* category as compared with the OECD average, but this difference disappeared in 2018.

Although there were no mathematics *single* or *multiple-choice* items omitted by more than 15 per cent of learners in Wales and comparator countries, the incidence of no-credit items in the latter category ranged from 29 to 64 per cent across comparator countries.

3.3 'Partial credit' mathematics items

Most PISA items are awarded full credit or no credit. However, there are some *open response* and *complex multiple-choice* items which have partial credit scoring, which allows more nuanced scoring to take into account that some answers, even though incomplete, are better than others (OECD, 2019a).

Of the 17 multi-mark items analysed, seven were mathematics items. All of them were *open response* items, six of which were human coded and one was computer scored. Table 3.1 shows the distribution of the seven mathematics partial credit items across to the process, content, context and item format categories. In PISA 2015, there were also seven partial credit items in mathematics, and the distribution across categories was similar.

The majority of mathematics multi-mark items fell into the *scientific* context category and the *employ* process category. The items were more evenly distributed across content categories.

Table 3. 1 Distribution of ‘partial credit’ mathematics items analysed in each category

	Number of partial credit items analysed	Total items of this type in PISA 2018 ⁶
Mathematical process		
Formulate	2	22
Employ	4	29
Interpret	1	19
Mathematical content		
Change and relationships	2	17
Space and shape	2	17
Quantity	1	18
Uncertainty and data	2	18
Mathematical context		
Personal	0	11
Occupational	1	13
Societal	1	26
Scientific	5	20
Response format		
Simple multiple-choice – computer scored	0	16
Complex multiple-choice – computer scored	0	14
Open response – computer scored	1	22
Open response – human coded	6	18

Source: PISA 2018 database

Table 3.2 shows the classification and facility (relative to the OECD) of the seven multi-mark mathematics items in Wales and Figure 3.11 shows the scoring pattern for these items. The facility is the average percentage of marks attained out of the maximum available for that item in the country.

⁶ Not including items in the cluster M6B

Table 3. 2 Classification and facility of the seven multi-mark mathematics items in Wales

Item ID	Facility Wales/ facility OECD Average. ⁷	Classification
DM906Q02C*	42%/42% Wales similar to OECD avg.	Quantity, Employing Mathematical Concepts, Facts and Procedures, Scientific
DM953Q04C**	13%/15% Wales similar to OECD avg.	Uncertainty and Data, Formulating Situations Mathematically, Scientific
CM955Q03S**	11%/11% Wales similar to OECD avg.	Uncertainty and Data, Employing Mathematical Concepts, Facts and Procedures, Societal
DM155Q02C	61%/63% Wales similar to OECD avg.	Change and Relationships, Interpreting, Applying and Evaluating Mathematical Outcomes, Scientific
DM155Q03C*	21%/21% Wales similar to OECD avg.	Change and Relationships, Employing Mathematical Concepts, Facts and Procedures, Scientific
DM949Q03C*	25%/30% Wales similar OECD avg.	Space and Shape, Formulating Situations Mathematically, Occupational
DM462Q01C***	1%/9% Wales lower than OECD avg.	Space and Shape, Employing Mathematical Concepts, Facts and Procedures, Scientific

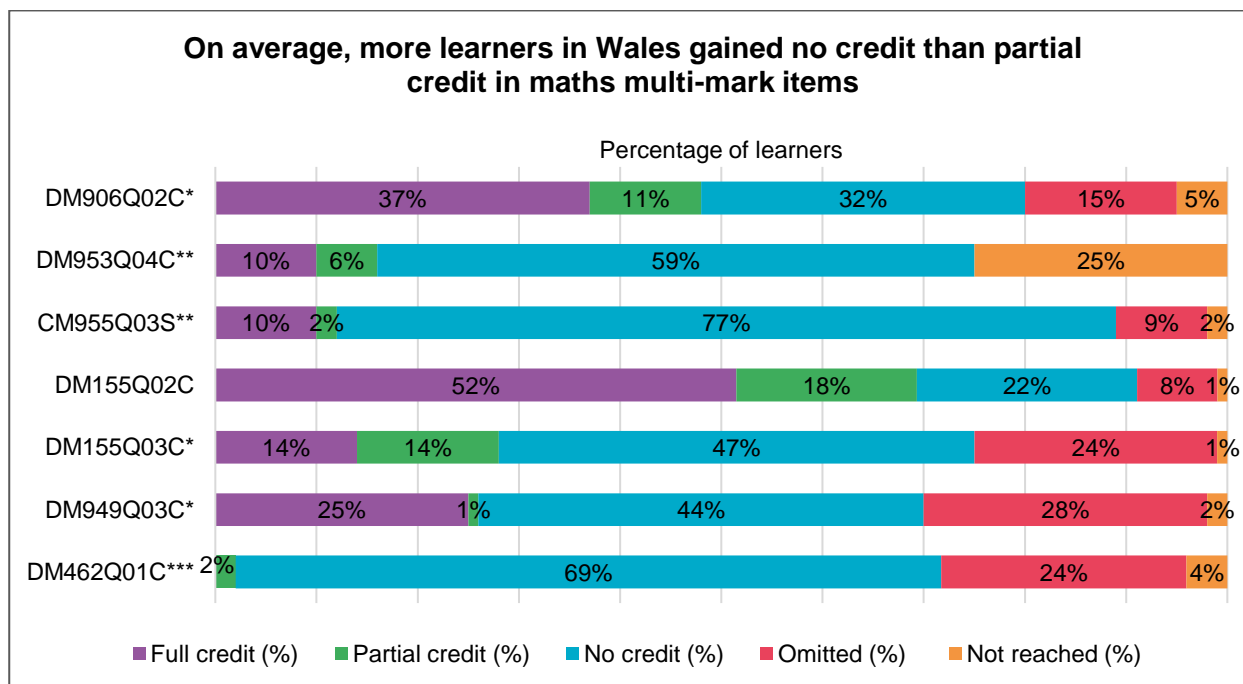
Note: *Item fell into the analysis category where more than 15% of learners omitted the item **Item fell into the analysis category of where more than 55% of learners in Wales attempted but gained no credit in the item ***Item fell into both analysis categories

Source: PISA 2018 database

Four items fell into the omission rate analysis criteria (>15 per cent) and three items fell into the no-credit analysis criteria (>55 per cent). One of the multi-mark items fell into both analysis categories i.e. it was omitted by more than 15 per cent of learners in Wales and attempted but failed by more than 55 per cent of learners in Wales. The latter was the only item in which the average facility of Wales was lower than the OECD average, meaning that learners in Wales achieved one per cent of the all the marks available on that item, as compared to nine per cent of learners in the OECD average. This represents an improvement from 2015, where Wales had a lower facility than the OECD average in three multi-mark mathematics items.

⁷ First percentage correspond to Wales and second percentage correspond to the OECD average. If Wales facility was within one standard deviation above or below of the OECD facility then it was classified as having a similar facility

Figure 3. 9 Scoring pattern of the multi-mark mathematics items in Wales



Notes: *Item fell into the analysis category where more than 15% of learners omitted the item **Item fell into the analysis category of where more than 55% of learners in Wales attempted but gained no credit in the item ***Item fell into both analysis categories

Source: PISA 2018 database

The percentage of learners gaining partial credit in the seven multi-mark items ranged from one to 18 per cent. It was generally more common for learners in Wales to earn either no credit or full credit on multi-mark mathematics items, which followed the pattern seen in PISA 2015 and 2012.

4 Science

Key findings

- In PISA 2018, learners in Wales achieved a mean score of 488 in science, which was the closest it has been to the OECD average (489) since 2009. Wales' science score was similar to Latvia and Spain, higher than Iceland and lower than Estonia and Sweden. Estonia achieved the highest science score amongst European countries in 2018.
- Three per cent of science items were not attempted by more than 15 per cent of the learners. More than 55 per cent of learners in Wales failed to gain credit in 30 per cent of science items. Both omission and no-credit rates were lower than those of mathematics.
- Wales had the same number of science items omitted by 15 per cent of learners as higher performing Estonia, and less than Sweden and the OECD average. In terms of items in which more than 55 percent of learners gained no credit, Wales had a higher proportion than the OECD average and higher performing Estonia and Sweden.
- The omission analysis showed that, compared with the OECD average, learners in Wales tended to skip fewer items related to evaluate and design scientific enquiry and interpret data and evidence scientifically. Learners in Wales also omitted fewer items than the OECD average assessing procedural and epistemic knowledge, physical and living systems, and items related to personal issues.
- The no-credit analysis showed that, compared with higher performing comparator countries, learners in Wales tended to have higher incidence of no-credit items from the explain phenomena scientifically category, the living systems category, and items related to local/national issues. The latter category refers to items related to the community and issues such as food security, energy supply, disposal of waste, and environmental impact.
- Compared with PISA 2015, there was an improvement in omission rates and in the facility of science items in Wales with respect to the OECD average. In 2018, three multi-mark science items proved to be easier for learners in Wales than for their peers internationally.

4.1 'Omitted' science items

Learners in Wales were presented with 115 science items in PISA 2018. Only three items were not attempted by more than 15 per cent of the learners – considerably lower than

mathematics. Only three per cent of science items met the omitted item criteria, which is lower than the OECD average (six per cent). Latvia and high performing Estonia had the same number of omitted science items as Wales, while the other three comparator countries had a higher number of items skipped by more than 15 per cent of learners.

In PISA 2015, learners in Wales omitted seven per cent of science items. Science was the major domain in 2015 and, therefore, learners received a higher number of science items in that cycle (184 items); hence, comparisons with 2015 are focused on the proportion of items per category, rather than the number.

PISA assesses learners' performance in science literacy on four interrelated aspects:

- The **competencies** to understand and engage in critical discussion about issues that involve science and technology
- The scientific **knowledge** to understand the major facts, concepts and explanatory theories
- The **content** knowledge of science
- The **context** in which the items is set

Items are further classified into three response format types: simple multiple-choice, complex multiple-choice, and constructed response.

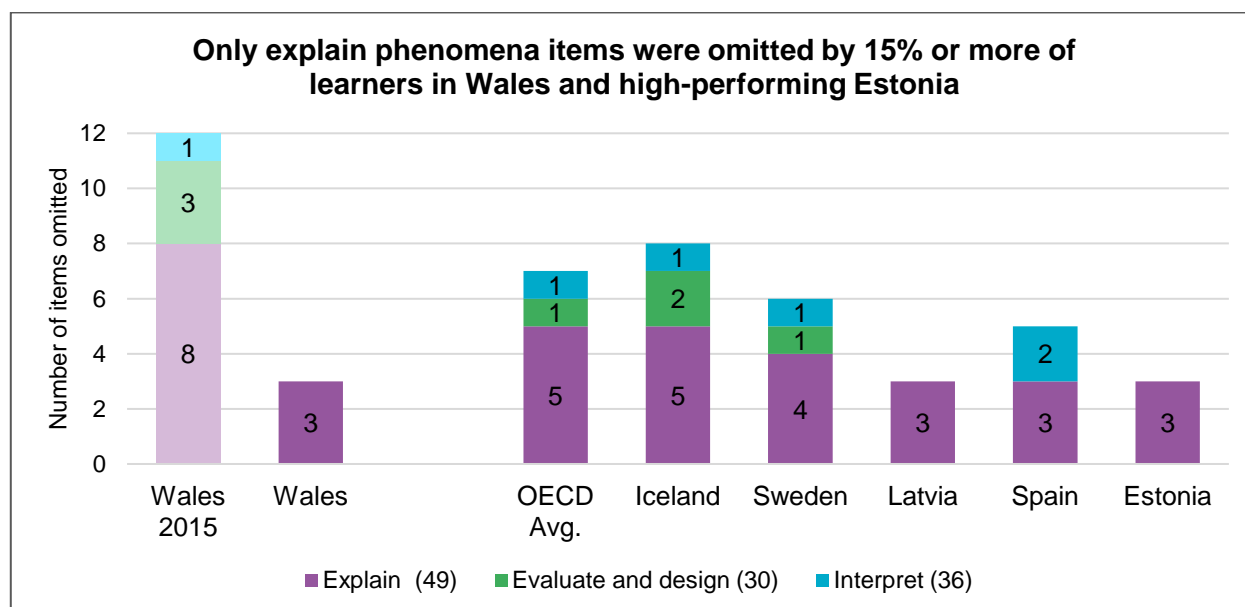
4.1.1 Competency categories

PISA science items are assigned to one of three science competencies: i) *explaining* phenomena scientifically, ii) *evaluating and designing* scientific enquiry, and iii) *interpreting* data and evidence scientifically.

Figure 4.1 shows the distribution of the analysed omitted items across competency categories. In all comparator countries and the OECD on average, items that asked learners to *explain phenomena scientifically* accounted for most of the items omitted by 15 per cent of learners. *Explain* items involve learners recognising, offering and evaluating explanations for natural and technological phenomena, and using these representations to predict possible changes (OECD, 2019a). In Wales, all the analysed omitted items corresponded to the *explain* category, six per cent of all *explain* items, as was also the case in Latvia and high performing Estonia.

In Wales, the proportion of omitted *explain* items was lower than the OECD average and Iceland. In 2015, learners in Wales, omitted items from all three categories, with most items falling predominantly into the *explain phenomena* category. The proportion of items omitted in the *explain phenomena* category in Wales fell from nine to six per cent from PISA 2015.

Figure 4. 1 Number of omitted science analysis items in each competency category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

4.1.2 Knowledge categories

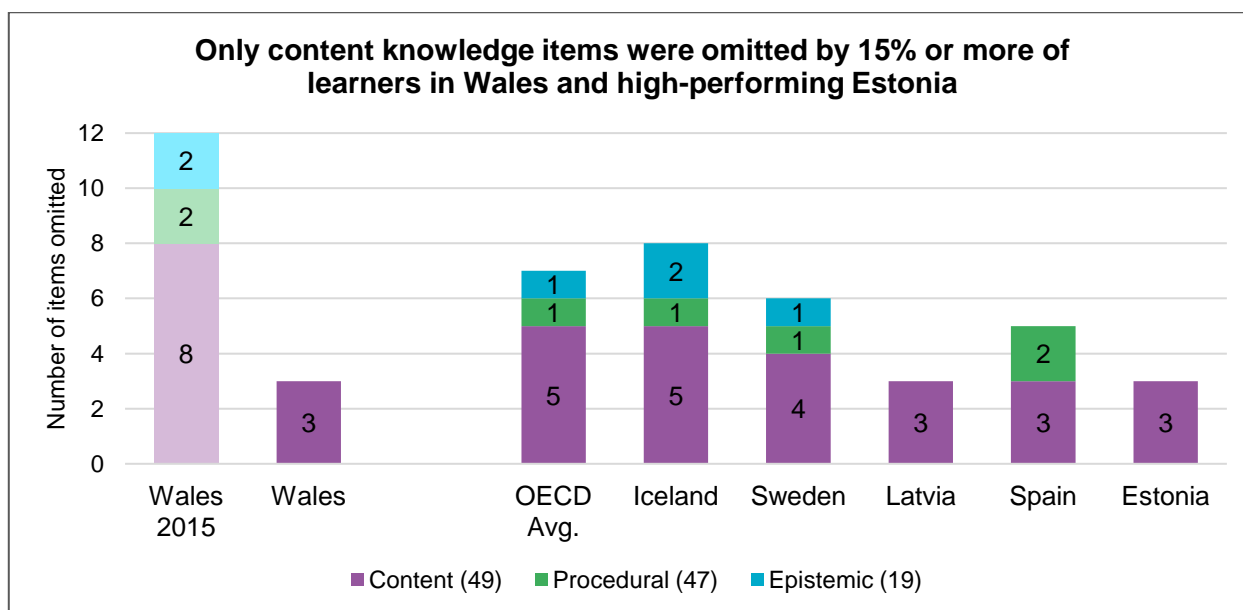
PISA classifies scientific knowledge into three distinguishable and related elements (OECD, 2019a):

- Content knowledge: The knowledge of facts, concepts, ideas and theories about the natural world that science has established
- Procedural knowledge: The procedures that scientist use to establish scientific knowledge
- Epistemic knowledge: The understanding of the role of specific constructs and defining features in building scientific knowledge, such as questions, hypotheses, and peer review.

Figure 4.2 shows the analysed omitted items across knowledge categories. In all comparator countries and the OECD on average, items that assessed *content knowledge* accounted for the highest number of items omitted by 15 per cent of learners. All omitted items in Wales fell into the *content knowledge* category, as was also the case for Latvia and Estonia. The number of omitted items was lower than the OECD average.

In 2015, the analysed items for Wales were from all three categories, with most items falling into the *content knowledge* category; eight per cent of omitted items were *content* in 2015, similar to six per cent in 2018.

Figure 4. 2 Number of omitted science analysis items in each knowledge category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

4.1.3 Content categories

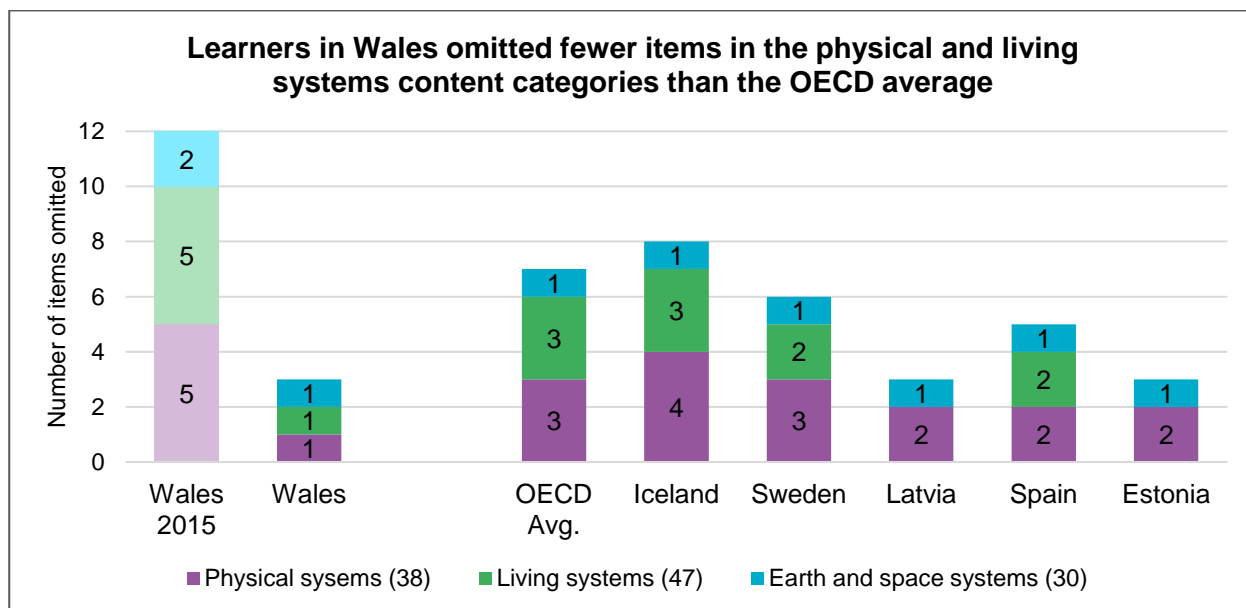
All PISA science items require content knowledge. PISA science has three categories of content knowledge, drawn from the major fields of physics, chemistry, biology, and earth and space sciences. These are:

- Physical systems (related to matter, energy, motion, etc.)
- Living systems (related to cells, organisms, humans, ecosystems, etc.)
- Earth and space systems.

In Wales, the three omitted science analysis items were distributed across the three content categories, representing between two and three per cent of all items in each category. In PISA 2015, the incidence of analysed omitted items in Wales was eight per cent for *physical systems*, seven per cent for *living systems*, and four per cent for *earth and space systems*.

Learners in Wales had fewer items omitted in the *physical systems* and the *living systems* category than the OECD average, whilst in 2015, Wales had a similar number of omitted items to the OECD average in all three competencies.

Figure 4.3 Number of omitted science analysis items in each content category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

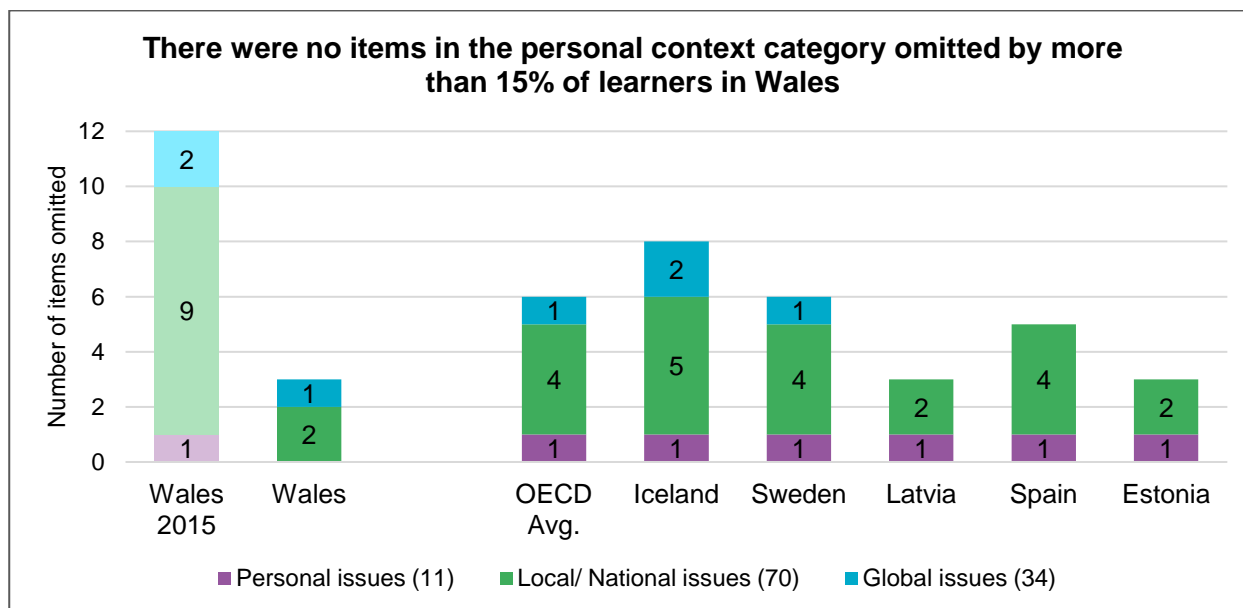
4.1.4 Context categories

PISA categorises science items into three contexts: i) *personal issues* (situations relating to self, family and peer groups), ii) *local and national issues* (situations relating to community), and iii) *global issues* (situations relating to life across the world).

Figure 4.4 shows the distribution of science omitted items across context categories. In Wales, learners omitted three per cent of items related to *local/national* issues and *global* issues. Compared with the OECD average and the comparator countries, Wales was the only country with no items omitted by more than 15 per cent of learners in the *personal issues* category.

In 2015, learners in Wales omitted items from all three categories. Most omitted items had a *local/national* context and there has been a decrease in the omission of these items, from 8 per cent in 2015 to 3 per cent in 2018. In 2015 learners in Wales had a similar number of items omitted in the *local/national* category than the OECD average, whilst in 2018 Wales had fewer items that met the mission criteria in that category.

Figure 4. 4 Number of omitted science analysis items in each context category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

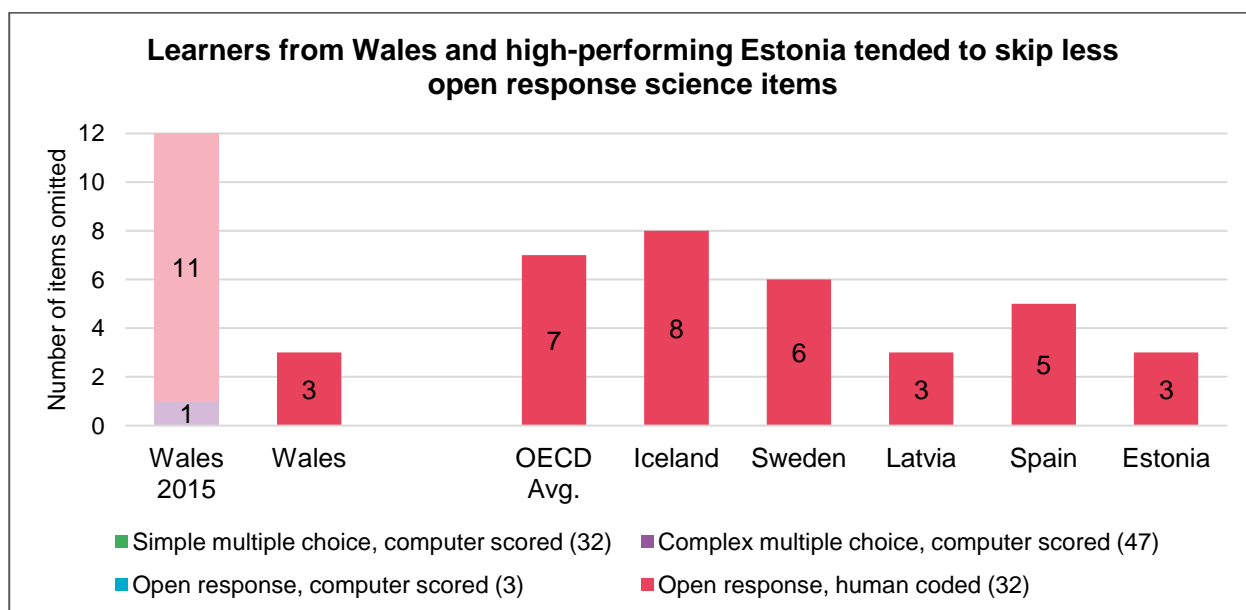
4.1.5 Response format categories

The science items are classified into three types of response format: *simple multiple-choice*, *complex multiple-choice*, and *constructed response*. *Constructed response* items in the scientific literacy requires students to write responses ranging from a phrase to a short paragraph, and sometimes ask students to draw a graph or diagram (OECD, 2019).

In PISA 2018, all science items omitted by more than 15 per cent of learners in Wales and comparator countries were *open response (human coded)* items. Wales had the same number of items that met the omission criteria than high performing Estonia, and fewer items than the OECD average.

There was a decrease in proportion of items omitted in the pool of *open response (human coded)* category in Wales, from 19 per cent in 2015 to nine per cent in 2018. However, it is important to note that the amount of open response (human coded) items fell from 58 in 2015 to 32 in 2018.

Figure 4. 5 Number of omitted science analysis items in each response format category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain

Source: PISA 2018 database

4.2 'No-credit' science items

There were 34 science items where more than 55 per cent of learners in Wales failed to gain credit. This represents 30 per cent of all science items in PISA 2018, which is similar to 33 per cent of no-credit science items analysed in PISA 2015.

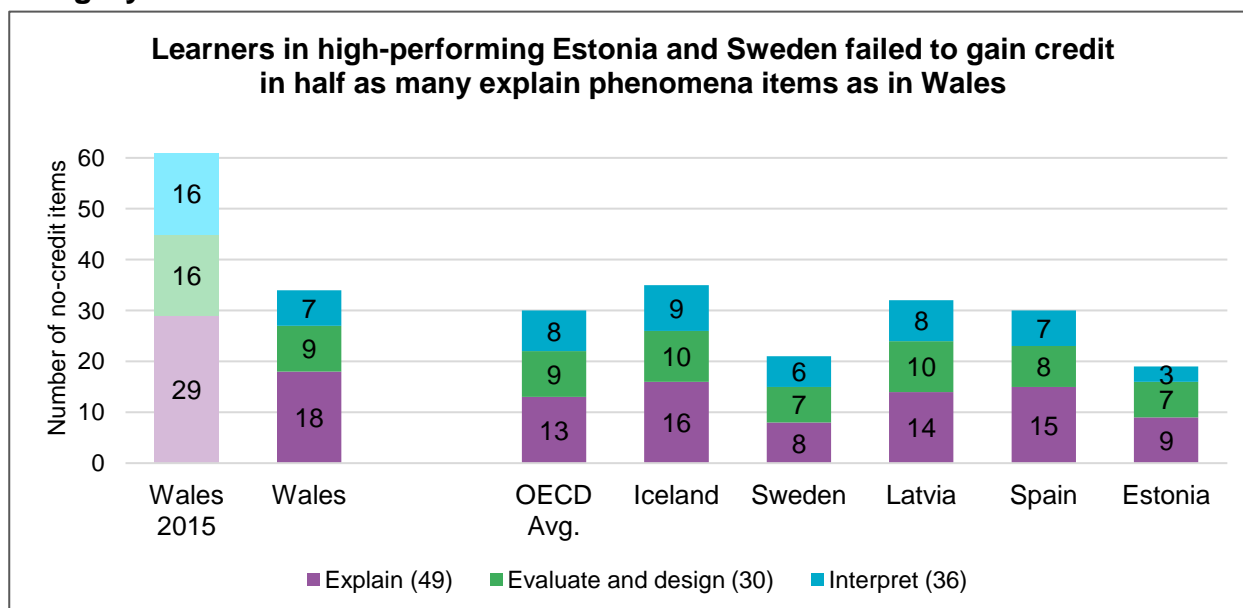
Wales had a similar proportion of no-credit items to Iceland, Latvia and Spain, and a higher proportion than the OECD average (25 per cent) and the two top performer countries Sweden and Estonia. These two countries had lowest incidence of no-credit items, with more than 55 per cent of learners gaining no credit in between 17 and 18 per cent of science items.

Figures 4.6 to 4.10 show the distribution of these items according to the four aspects of science literacy (competency, knowledge, content and context), and according to the items' response format.

4.2.1 Competency categories

Figure 4.6 shows the distribution of the three no-credit items across the competency categories. Learners in Wales failed to gain credit in 37 per cent of *explain* items, 30 per cent of *evaluate and design* items and 19 per cent of *interpret* items. The amount of no-credit items in the *interpret* and the *evaluate and design* categories was consistent with the OECD average.

Figure 4. 6 Number of no-credit science analysis items in each competency category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

Learners in Wales had a higher incidence of no-credit items in the *explain phenomena* category than all comparator countries. The amount of no-credit items in the *evaluate* and *interpret* competencies were similar to the OECD average but higher than high-performing Estonia.

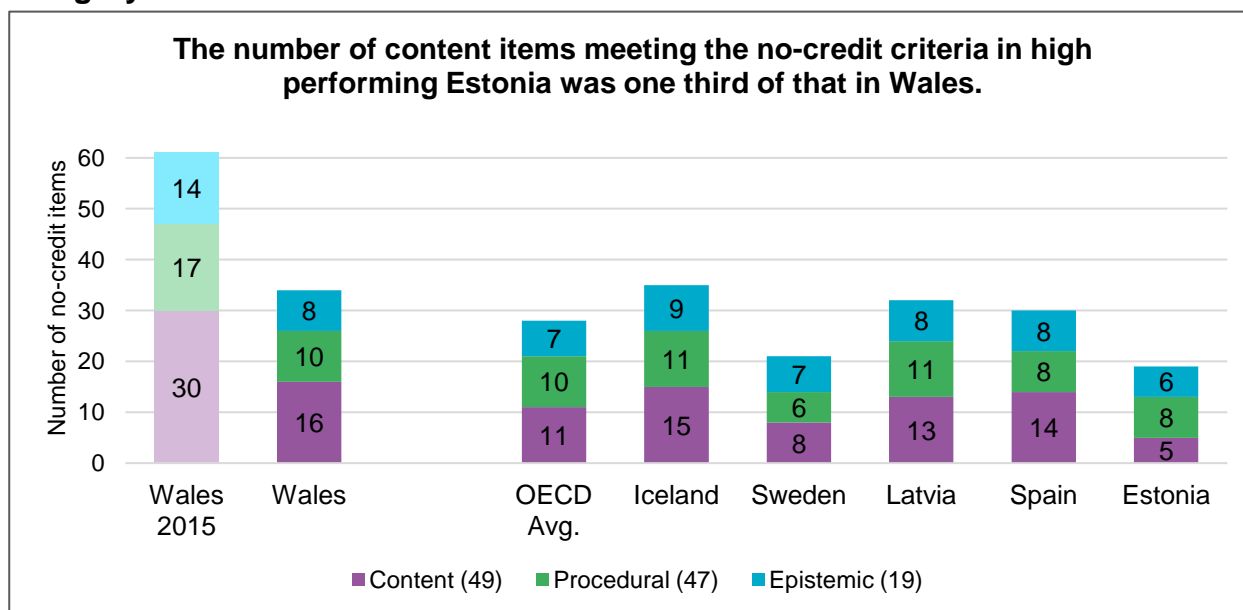
Compared with 2015, the percentage of no-credit items in the *explain phenomena* category remained similar in Wales, whereas the percentage of *evaluate* and *interpret* analysed items reduced by around ten percentage points.

4.2.2 Knowledge categories

Figure 4.7 shows the distribution of no-credit science analysis items across knowledge categories. In Wales, the *epistemic* knowledge category had the highest proportion of no-credit items (42 per cent), followed by *content* knowledge (33 per cent), and items relating to *procedural* knowledge (21 per cent). The OECD average had a similar number of no-credit items in the *epistemic* and *procedural* categories, but fewer than Wales on the *content* category.

Across comparator countries, the highest incidence of no-credit items belonged to the *epistemic* category. This type of knowledge is related to the understanding of the underlying practices of scientific enquiry and the meaning of foundational terms such as theory, hypothesis, observations and models; and it is most likely to be tested in a pragmatic way rather than asking for specific theory (OECD, 2019a). Higher-performing Sweden and Estonia had fewer no-credit items in this category than Wales.

Figure 4. 7 Number of no-credit science items analysed in each knowledge category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

All items omitted by more than 15 per cent of learners in Wales corresponded to *content* items, and more than 55 per cent of learners did not gain credit in a third of the *content* items attempted. This proportion was greater than higher-performing Sweden and Estonia.

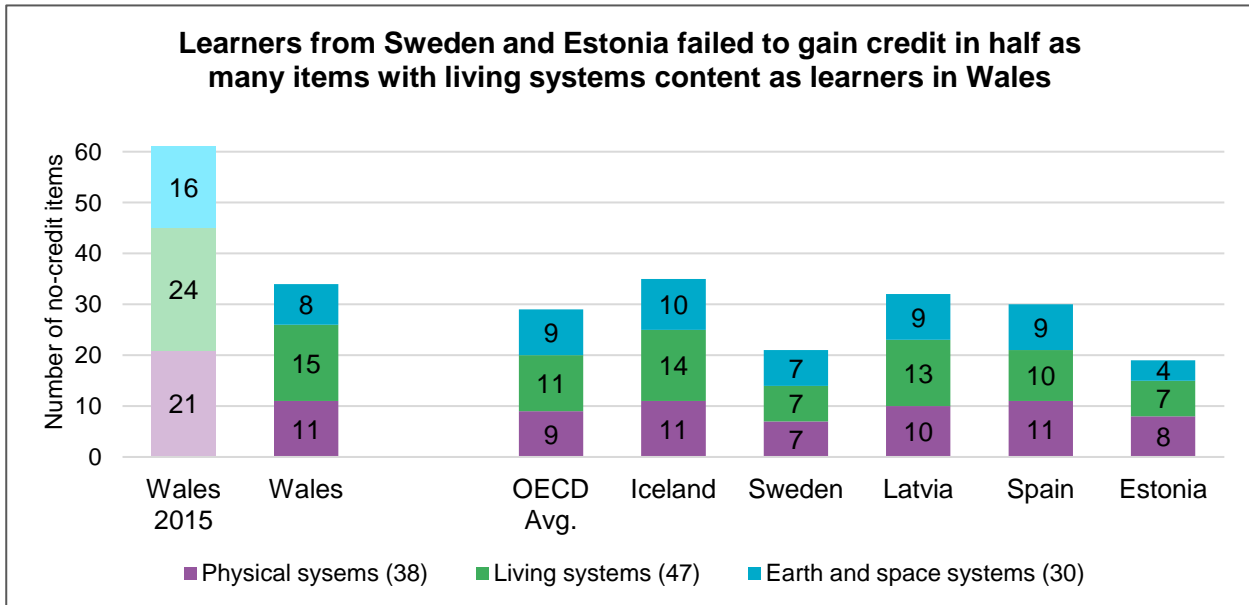
The distribution of items across knowledge categories was similar to PISA 2015, when the highest incidence of no-credit items in Wales also fell in the *epistemic* category (54 per cent), followed by *content* (30 per cent) and *procedural* knowledge (28 per cent).

4.2.3 Content categories

The numbers of no-credit analysis items in each science content knowledge category are shown in Figure 4.8. The three categories had very similar proportions of items on which more than 55 per cent of the learners in Wales failed to gain credit: 29 per cent for the *physical systems* content, 32 per cent for *the living systems* content and 27 per cent for the *earth and space* content. The proportion of no-credit items in the *physical systems* and *earth and space* categories were in line with the OECD average, but greater than higher-performing Sweden and Estonia.

The distribution was similar to PISA 2015, when more than 55 per cent of learners in Wales gained no marks in 34 per cent of *physical systems*, 32 per cent of *living systems* and 33 per cent of *earth and space* items. Wales had more no-credit items in the *living systems* category than the OECD and all the comparator countries except from Iceland. This followed the same pattern as PISA 2015.

Figure 4. 8 Number of no-credit science analysis items in each content category

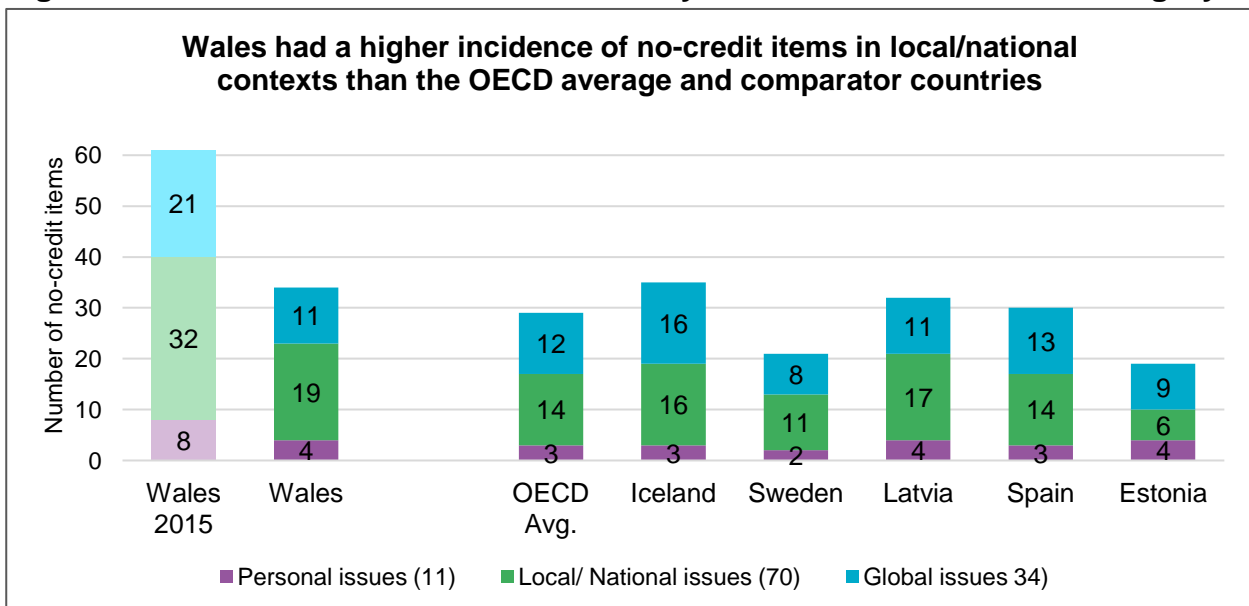


Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

4.2.4 Context categories

Figure 4.9 shows the number of no-credit science items in each context category. More than 55 per cent of learners in Wales failed to obtain credit in around a third of items in each category: 36 per cent for the *personal* context, 32 per cent for the *global* context and 27 per cent for the *local/national* context. Wales had more no-credit items in the *local/national issues* category than the OECD average and all comparator countries, with Estonia showing a notably lower proportion of no-credit items in the category (9 per cent).

Figure 4. 9 Number of no-credit science analysis items in each context category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

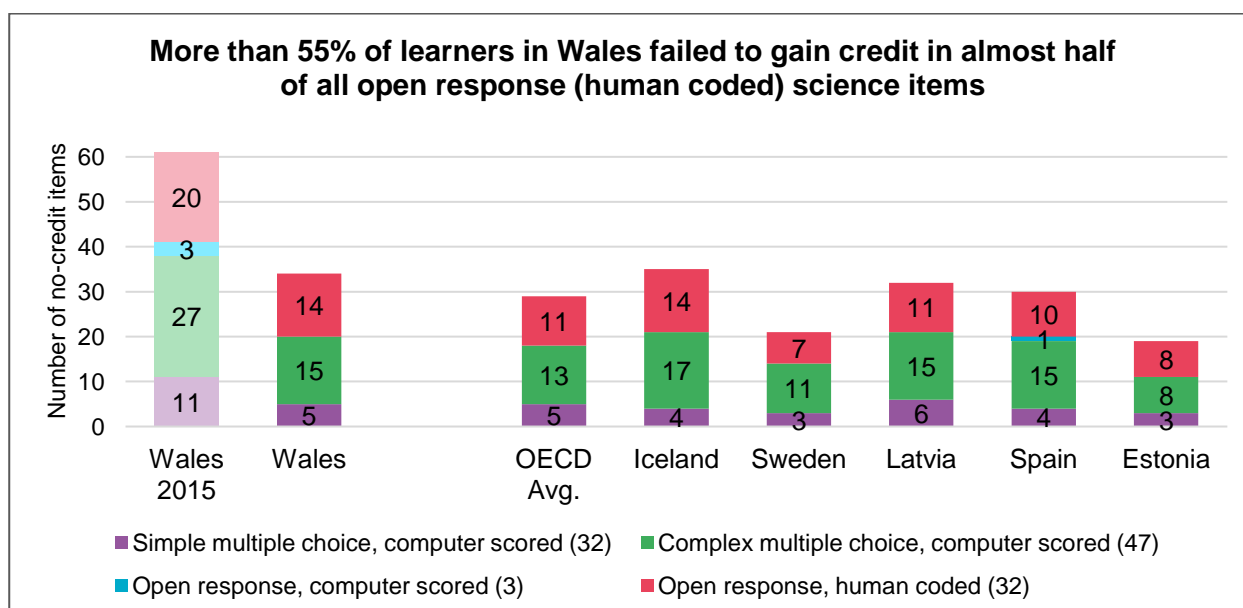
4.2.5 Response format categories

Figure 4.10 shows the number of no-credit science items in each response format category. In Wales, 44 per cent of *open response (human coded)* items and 32 per cent of *complex multiple-choice* items are no-credit analysis items, that is attempted but failed to gain credit by more than 55 per cent of learners. The *simple multiple-choice* category had the lowest incidence of no-credit items, and similar to the OECD average. Higher-performing Sweden and Estonia had a lower proportion of no-credit items in all categories compared to Wales and the OECD average.

All science items in Wales omitted by more than 15 per cent of learners fell into the *open response (human coded)* category. Learners that attempt this type of item failed to gain credit in almost half of them, indicating that *open response* science items are challenging for learners in Wales. This pattern is similar for the OECD average and all comparator countries.

Compared with 2015, learners in Wales had a higher proportion of no-credit *open response (human coded)* items, and a lower proportion of no-credit *complex multiple-choice* items. This refers to the proportion of items that met the no-credit analysis criteria within each category.

Figure 4. 10 Number of no-credit science analysis items in each response format category



Note: The numbers in brackets correspond to the total number of items in the category in PISA 2018. In 2015, learners in Wales were presented with 184 science items, as science was the major domain.

Source: PISA 2018 database

4.3 'Partial credit' science items

Of the 17 multi-mark items in PISA 2018, ten were science items. Seven of the multi-mark science items were *open response* items, six human coded and one computer scored. Three were *complex multiple-choice - computer scored* items. The distribution of 'partial credit' analysis items across the different scientific literacy categories is shown in Table 4.1.

Table 4. 1 Distribution of 'partial credit' science items analysed in each category

	Number of partial credit items analysed	Total items of this type in PISA 2018
Science competency		
Explain phenomena scientifically	7	49
Evaluate and design scientific enquiry	1	30
Interpret data and evidence scientifically	2	36
Science content		
Physical systems	1	38
Living systems	7	47
Earth and space systems	2	30
Science knowledge		
Content	6	49
Procedural	4	47
Epistemic	0	19
Science context		
Personal issues	0	11
Local/ National issues	8	70
Global issues	2	34
Response format		
Simple multiple-choice – computer scored	0	32
Complex multiple-choice – computer scored	3	47
Open response – computer scored	1	3
Open response – human coded	6	32

Source: PISA 2018 database

Seven items fell into the *explain phenomena* competency, eight items were related to *local/national issues*, six items fell into the *content* type of knowledge, and seven items were related to *living systems*.

Table 4.2 shows the classification and facility (relative to the OECD) of the ten multi-mark science items in Wales and Figure 4.11 shows the scoring pattern for each these items. Facility stands for the average percentage of marks attained out of the maximum available for that item in the country.

Table 4. 2 Classification and facility of the ten multi-mark science items in Wales

Item ID	Facility Wales/ facility OECD Average ⁸	Classification
DS607Q03C**	34%/40% Wales lower than OECD avg.	Explain phenomena scientifically, Content, Local/ National, Living
CS645Q01S	51%/45% Wales higher than OECD avg.	Explain phenomena scientifically, Content, Global, Earth and Space
DS498Q04C	66%/58% Wales similar to OECD avg.	Explain phenomena scientifically, Content, Global, Earth and Space
DS605Q04C*	56%/43% Wales higher than OECD avg.	Explain phenomena scientifically, Content, Local/National, Earth and Space
DS657Q04C	30%/19% Wales higher than OECD avg.	Explain phenomena scientifically, Content, Local/ National, Living
CS635Q01S	56%/48% Wales similar to OECD avg.	Explain phenomena scientifically, Content, Local/ National, Living
CS635Q04S	45%/44% Wales similar to OECD avg.	Evaluate and design scientific enquiry, Procedural, Local/ National, Living
DS635Q05C**	13%/13% Wales similar to OECD avg.	Explain phenomena scientifically, Content, Global, Physical
DS629Q01C	46%/49% Wales similar to OECD avg.	Explain phenomena scientifically, Content, Local/National, Physical
CS634Q02S	31%/32% Wales similar to OECD avg.	Interpret data and evidence scientifically, Procedural, Global, Living

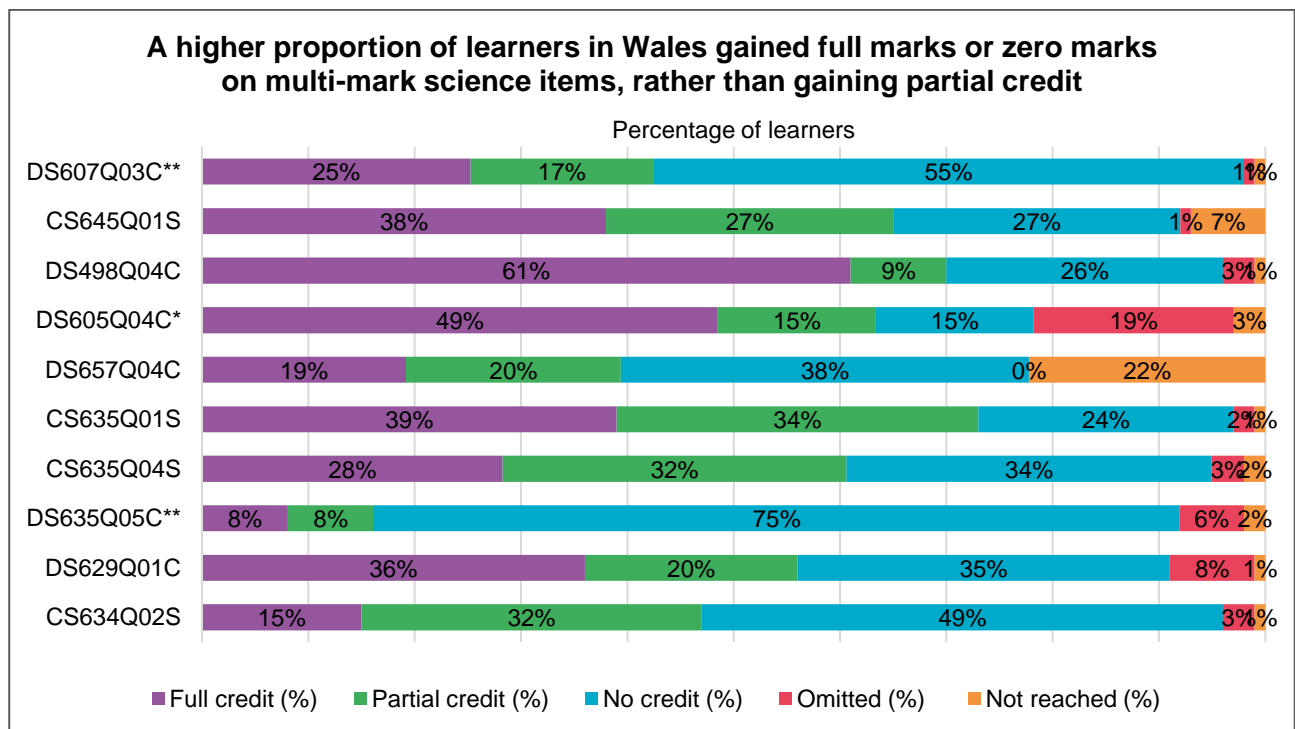
Note: *Item fell into the analysis category where more than 15% of learners omitted the item **Item fell into the analysis category of where more than 55% of learners in Wales attempted but gained no credit in the item ***Item fell into both analysis categories

Source: PISA 2018 database

⁸ First percentage correspond to Wales and second percentage correspond to the OECD average. If Wales facility was within 1 standard deviation above or below of the OECD facility then it was classified as having a similar facility

One of the ten multi-mark science items fell into the omission rate analysis criteria (>15 per cent of learners) and two items fell into the no-credit analysis criteria (>55 per cent learners). These three items belonged to the *explain phenomena* category and to the *content* type of knowledge category. Six items had a similar average facility than the OECD average, one item was more difficult in Wales than internationally, and three items proved to be easier for learners in Wales than internationally. This represents an improvement from 2015, where no multi-mark science items had a higher facility in Wales than the OECD average.

Figure 4. 11 Scoring pattern for the ten multi-mark science items in Wales



Note: *Item fell into the analysis category where more than 15% of learners omitted the item **Item fell into the analysis category of where more than 55% of learners in Wales attempted but gained no credit in the item, ***Item fell into both analysis categories

Source: PISA 2018 database

The percentage of learners gaining partial credit in the ten multi-mark science items ranged from eight to 34 per cent, which is higher than those rates of multi-mark mathematics items and similar to PISA 2015. On average around 38 per cent of learners gained no credit in these items, and 32 per cent gained full credit.

5 Subgroup analysis

Key findings

- Compared with PISA 2015, the gender gap in omission rates widened in science, due to boys skipping fewer science items than girls. In terms of items in which more than 55 per cent of learners failed to gain credit, the gender gap decreased for maths due to a smaller proportion of girls gaining no credit in maths items.
- The changes in omission and no-credit rates in mathematics between 2015 and 2018 suggests that learners who took the test in Welsh were more confident in attempting maths items in 2018, but on average failed to gain credit in more of the attempted maths items than their peers who took the test in English. This pattern was similar for learners in Welsh medium schools, as compared with their peers in English medium schools.
- The difference in omission rates between learners with high and low socio-economic background was more pronounced than the differences found to exist between genders, medium of schooling, and language of assessment.
- The gap between learners in the highest and lowest attainment quartiles for both omission and no-credit rates was larger for science than for mathematics.

5.1 Analysis background

This section comprises a high-level investigation into whether there were differences in the proportions of omitted and no-credit items in five subgroups of learners in Wales. The subgroups analysed were gender, language of assessment (Welsh or English), medium of schooling (Welsh or English), socioeconomic status, and lowest and highest attainment quartile. A comparison of 2015 and 2018 is shown for each subgroup.

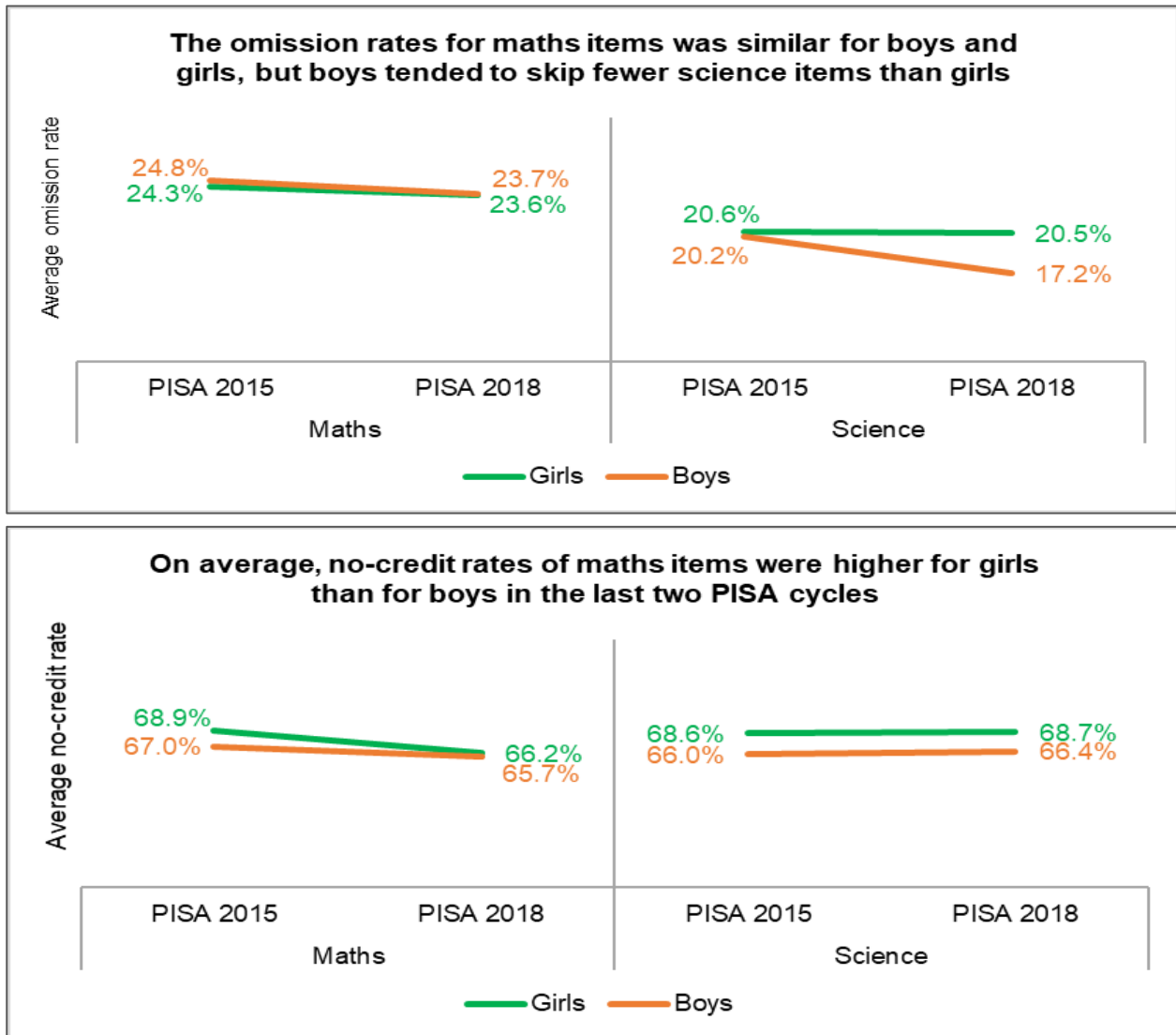
The analysis was conducted separately by maths and science and focussed on the subset of items identified in the main analysis (items omitted by more than 15 per cent of learners and items with a percentage of no-credit higher than 55 per cent). Using the gender analysis as an example, first we pooled together the items with an omission rate higher than 15 per cent in each subject, and then we computed an average omission rate on each pool of items by gender.

It should be noted that the results reported are for the subsample of items identified in the previous part of the analysis and cannot be generalised to the overall assessment. Differences between sub-groups may not be statistically significant, and do not constitute evidence of causal relationships between the characteristics and outcomes considered.

5.2 Gender

Boys and girls had a similar omission rates in the mathematics items analysed in this report (around 24 per cent), following the same pattern as in PISA 2015. In terms of learners receiving no credit, a lower proportion of girls and boys failing to gain credit in mathematics items in 2018 contributed to closing the gender gap by 1.5 percentage points.

Figures 5. 1 Average omission and no-credit rates of the analysis items by gender



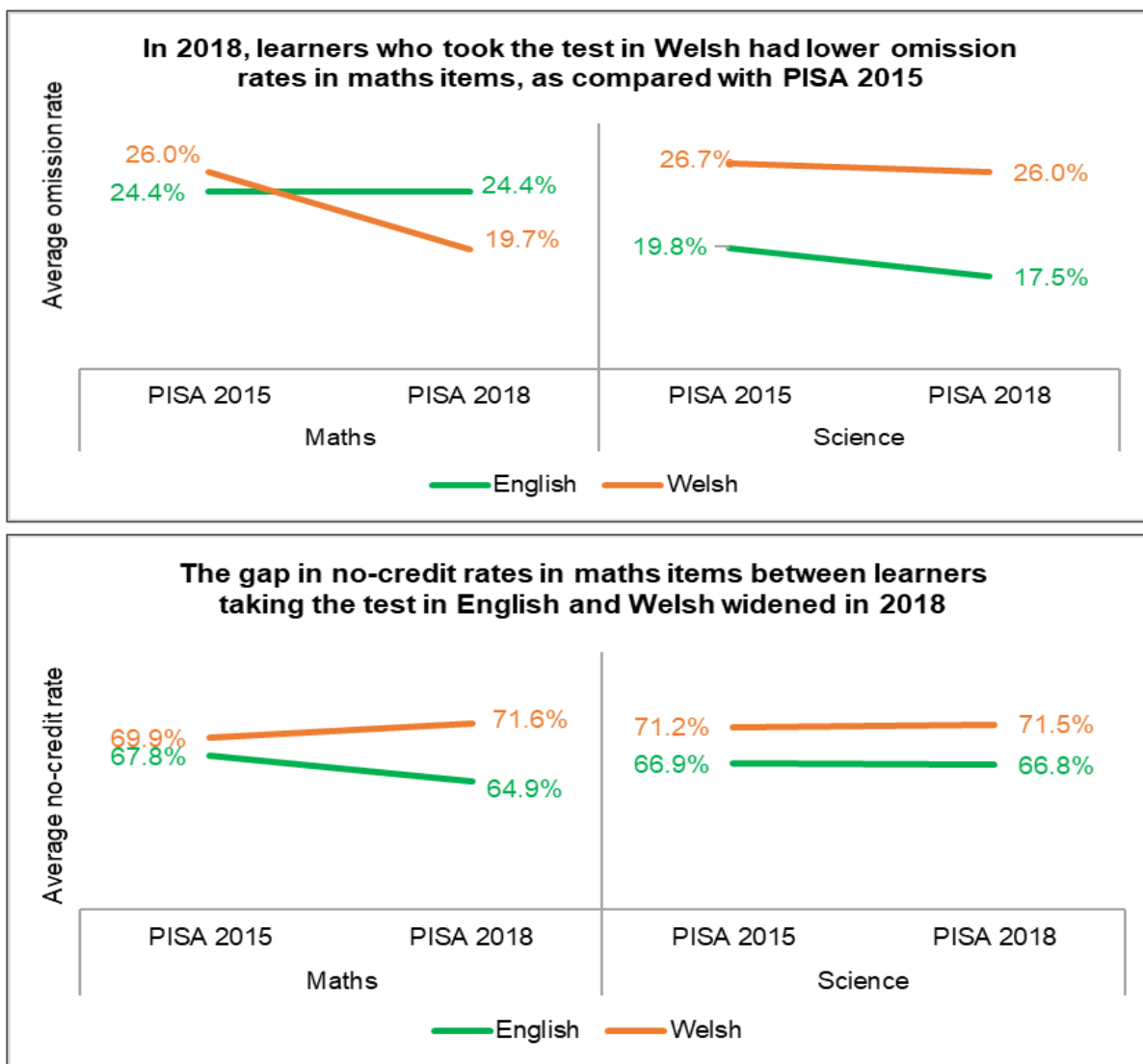
Source: PISA 2018 database

The gender gap in omission rates in science items increased by three percentage points in PISA 2018, due to boys omitting on average fewer science items than in PISA 2015, while the proportion for girls remained similar. The proportion of boys receiving no-credit in science items was two percentage points lower than girls, following the pattern of PISA 2015.

5.3 Language of Assessment

Learners who completed the assessment in Welsh had a higher tendency to skip science items, as was the case in PISA 2015. However, a lower proportion of learners taking the test in Welsh omitted maths items as compared with learners taking the test in English.

Figures 5. 2 Average omission and no-credit rates of the analysis items by language of assessment



Source: PISA 2018 database

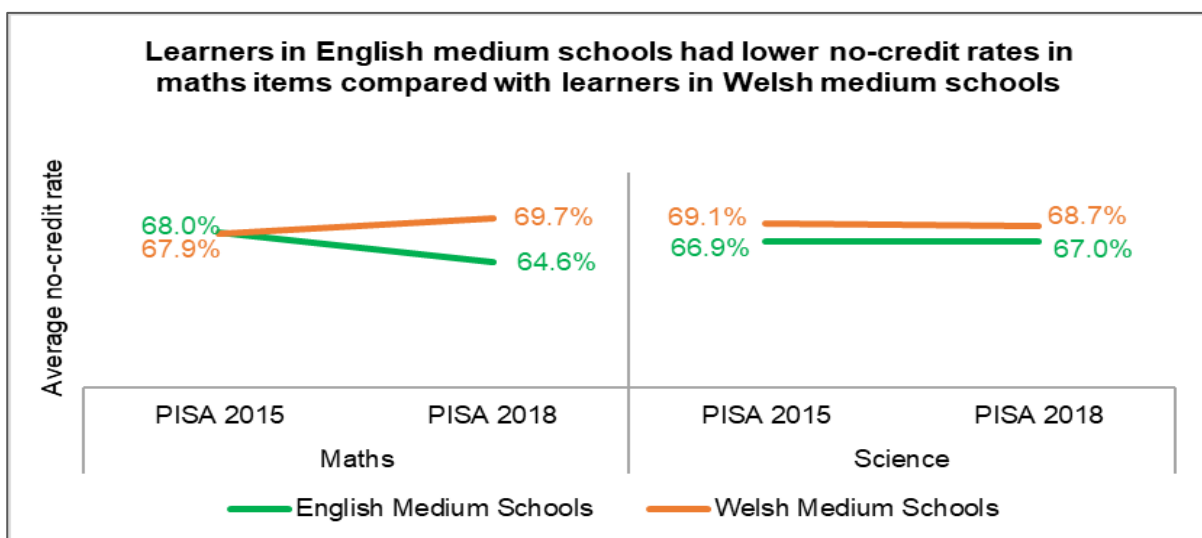
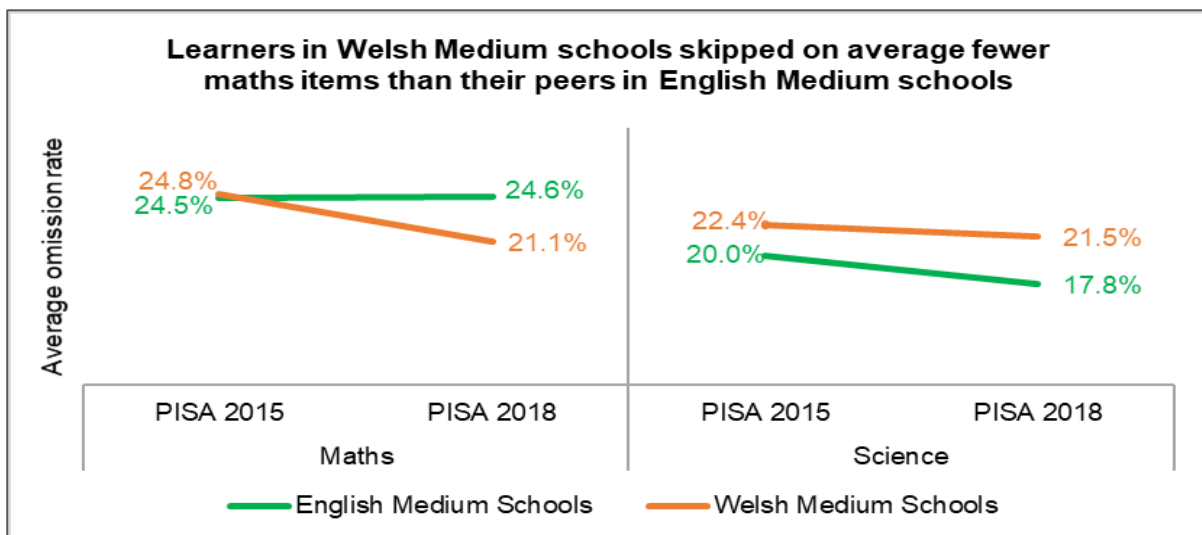
A higher proportion of learners who took the test in Welsh obtained no credit in the set of mathematics and science items that were analysed. Compared with PISA 2015, the gap remained constant for science but widened in mathematics (from 2.1 to 6.7 percentage points). The increased gap in mathematics was due to both more learners taking the assessment in Welsh failing to gain credit and less learners taking the test in English failing to gain credit in maths items.

This suggests that learners who took the test in Welsh were more confident in attempting mathematics items in 2018, but on average failed to gain credit in more of the attempted items than their peers who took the test in English. In science, the gap in the percentage of learners omitting and gaining no credit in the subset of items analysed remains in favour of learners taking the test in English.

5.4 Medium of schooling

Figures 5.3 (a and b) compare the results of learners enrolled in Welsh medium schools with those of learners enrolled in English medium schools. In 2018, the gap between learners in Welsh and English medium schools widened in mathematics as compared with 2015, in terms of the omitted and the no-credit omission rates.

Figures 5.3 Average omission and no-credit rates of the analysis items by medium of schooling



Source: PISA 2018 school census matched database

Omission rates in mathematics items was lower for learners in Welsh medium schools (21 per cent) than for their peers in English medium schools (25 per cent). On average, 65 per cent of learners from English medium schools gained no credit in the subsample of mathematics items analysed as compared with 70 per cent of learners in Welsh medium schools, increasing the gap from 0.1 percentage points in 2015 to 5.1 percentage points in 2018.

For science items, the omission and no-credit rates were lower for learners in English medium schools compared with their peers in Welsh medium schools. Although a lower proportion of pupils (both from English and Welsh medium schools) omitted science items than in 2015, the size of the gap remained similar across PISA cycles.

See *Welsh- and English-Medium School Results and Regional Performance and PISA/GCSE Matching in Wales* for further analyses of Wales' PISA 2018 performance (Classick *et al.*, 2020 and Gambhir *et al.*, 2020).

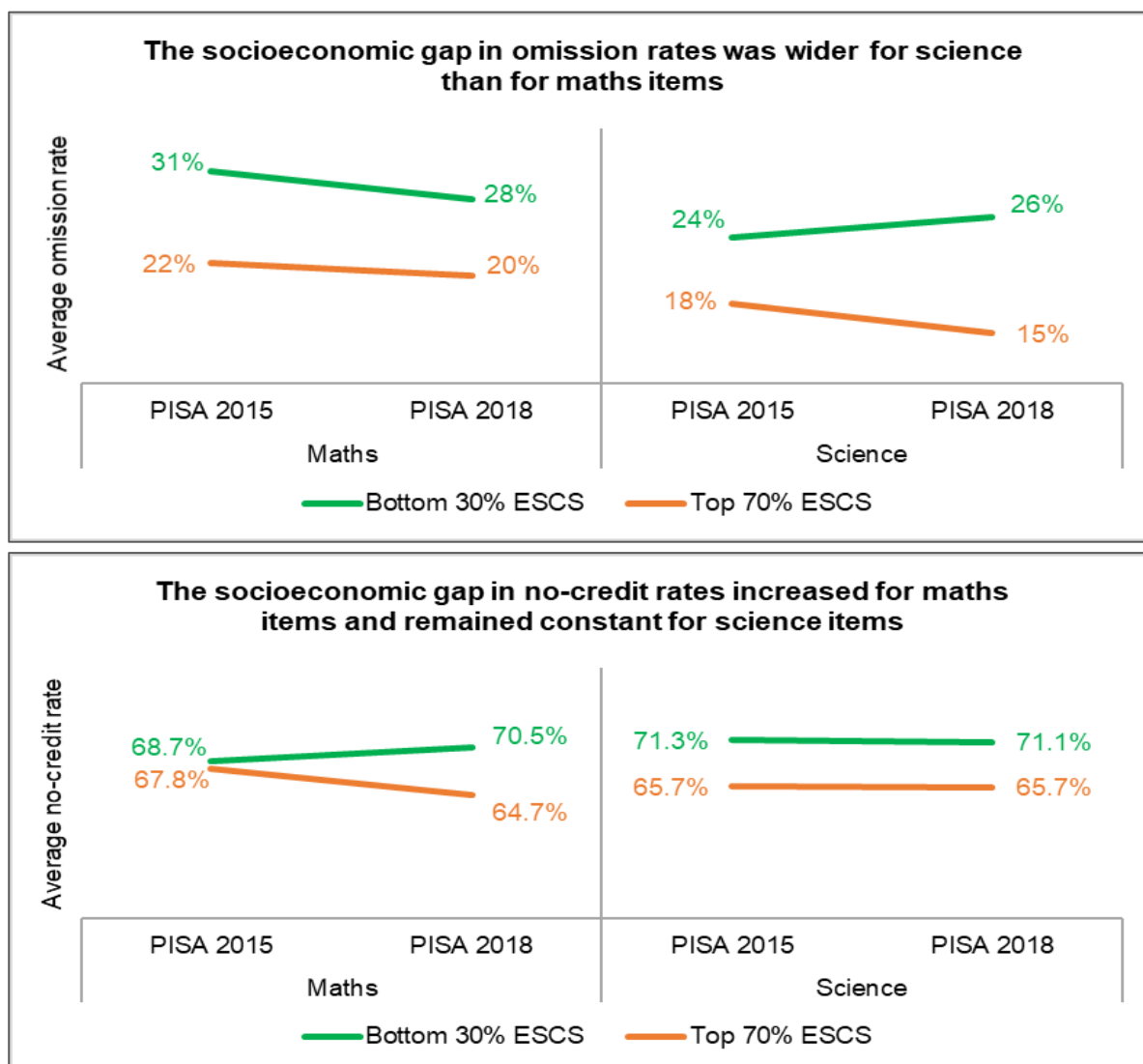
5.5 Socioeconomic status

Figures 5.4 (a and b) report the patterns of responses of learners in the bottom 30 per cent of the socio-economic distribution with those of individuals in the top 70 per cent, using PISA's Economic, Social and Cultural Status (ESCS) measure. The difference in the omission rates between learners with different socio-economic statuses is more pronounced than the differences found to exist between genders, medium of schooling, and language of assessment. Following the pattern of PISA 2015, the socio-economic background seemed to be more associated with the proportion of items that learners left blank than with the proportion of items answered incorrectly.

Compared to PISA 2015, the gap in omission rates between learners from different socio-economic statuses became more pronounced in science, going from 6.6 to 11.4 percentage points. This was a result of both disadvantaged learners skipping more science items and advantaged learners being more confident in attempting science items. Although the size of the gap in mathematics items remained similar, learners at both sides of ESCS distribution had lower omission rates in 2018.

In terms of no-credit rates in the pool of analysed items, the gap in favour of advantaged learners remained similar for science but increased for maths items, going from 0.9 to 5.8 percentage points from 2015 to 2018. This suggest that learners from the bottom 30 per cent of the ESCS distribution attempted more maths items but also failed to gain credit in more of the attempted items as compared with PISA 2015.

Figures 5. 4 Average omission and no credit rates of the analysis items by socioeconomic status



Source: PISA 2018 database

5.6 Lowest and highest national attainment quartile

Table 5.1 compares the performance of learners in the analysis items by national attainment quartiles. The difference in the omission and no-credit rates between learners in the highest and lowest attainment quartiles in PISA 2018 was higher for science than for mathematics.

Whilst 80 per cent of learners in the lowest attainment quartile gained no credit in the analysed science items, this figure was 47 per cent for their peers in the highest quartile. The percentage of learners in the highest quartile gaining no credit was similar for mathematics items, but the gap was narrower since fewer learners in the lowest quartile gained no credit in maths as compared to science (73 and 80 per cent, respectively).

Table 5. 1 Item subgroup analysis by lowest and highest attainment quartile

Subject	Average no-credit rates by attainment quartile			Average omission rates by attainment quartile		
	Lowest quartile	Highest quartile	Difference lowest-highest	Lowest quartile	Highest quartile	Difference lowest-highest
Mathematics	72.6%	47.2%	25.5 pp	37.5%	10.1%	27.4 pp
Science	80.1%	46.7%	33.4 pp	41.1%	1.9%	39.2 pp

Source: PISA 2018 database

In terms of omission rates, the gap between the lowest and highest quartile was more pronounced than for no-credit rates. There was also a noticeable difference in the proportion of learners in the highest quartile skipping items across subjects. While 10 per cent of learners in the highest attainment quartile omitted the analysed maths items, only 2 per cent of learners in the same quartile omitted science items. Learners in the lowest quartile omitted more science than mathematics items, and failed to gain credit in more science than mathematics items. This mirrors the findings of the PISA 2018 national report, which showed a greater attainment gap in science than in maths.

6 Conclusions

6.1 Subject knowledge and skills

Compared to the OECD average, Wales had similar a number of items meeting the omission criteria in mathematics; following the pattern of the previous two PISA cycles. In science, Wales had a lower number of omitted items than the OECD average and a similar number to that of higher-performing Estonia. As in PISA 2015, the number of items meeting the no-credit criteria in both mathematics and science in Wales was higher than the OECD average. This suggests that in PISA 2018, learners in Wales were more confident in attempting science questions, but continued to fail to gain credit in more items than their international peers do.

By looking at the performance of higher performing countries, we can understand the comparative strengths and weaknesses of learners in Wales. For instance in mathematics, learners from higher-performing Latvia and Estonia had on average fewer omitted items than Wales in the *employing mathematical concepts* process category, the *quantity* content category, and the *scientific* context category. This indicates that learners in Wales were less confident in attempting items related to quantification (e.g. measurements, units and sizes), items that required them to perform mathematical procedures needed to derive results (e.g. arithmetic computations, solving equations, extracting information from tables and charts, and analysing data), as well as problems related to the natural world, science and technology.

In terms of science, learners in Wales skipped the same small number of items as their Estonian peers. Compared to the OECD average, Wales had fewer omitted items across all categories, and in particular items related to *personal issues* and *living systems*. Conversely, the no-credit item analysis showed that Wales failed to gain credit in more items of the *living systems* content category than almost all comparator countries. Therefore, although learners in Wales were relatively confident in attempting questions related to cells, organisms, humans and ecosystems, they performed poorly on them.

Regarding items in which more than 55 per cent of learners failed to gain credit, Wales had on average more items meeting the criteria than higher performing comparator countries and the OECD average. In mathematics, items in four categories drove the difference with international peers: *formulate process*, *uncertainty and data* content, and *societal* and *occupational* contexts. This suggests that learners in Wales need to improve skills related to identifying the mathematical aspects and the significant variables of a problem situated in a real-world context, interpretation and presentation of data, probability and statistics.

In science, the categories in which learners in Wales had comparatively more items with no credit were the *explain phenomena* competency, the *content* knowledge, *local/national* context and, as mentioned before, the *living systems* content. This

suggests that learners in Wales struggle with science questions that ask them to describe or interpret phenomena, generate hypothesis and predict possible changes, recall and use theories, explanatory ideas, information, and facts.

6.2 Assessment skills

All the mathematics items omitted by more than 15 per cent of learners in Wales and comparator countries in PISA 2018 corresponded to *open constructed-response* items, as was the case in PISA 2015. This suggests that learners in general are less willing to try to answer questions that require an extended written response or which ask learners to show the steps taken to reach an answer.

The highest incidence of omitted and no-credit science items also fell into the *open constructed-response* category. This pattern was similar for the OECD average and almost all comparator countries. Most science *open constructed-response* items asked for a written responses ranging from a phrase to a short paragraph, and a small number of items asked learners draw a graphs or diagram in a drawing editor (OECD, 2019).

Learners in Wales showed a weakness in performance in *complex multiple-choice* items, failing to gain credit in more items of this type than their international peers in both mathematics and science domains. This item format can comprise a series of interrelated 'Yes/No' questions, the completion of a sentence from a drop-down menu, a 'drag-and-drop' response, or an interactive task (the latter introduced in PISA 2018), and it helps to measure learners' understanding of higher-order concepts which may be difficult to express as an open response (OECD, 2019).

Wales' performance in multi-mark items relative to the OECD average improved for both subjects in PISA 2018. Learners in Wales gained on average fewer marks than the OECD average in one mathematics multi-mark item, as compared with three items in 2015. In science, learners in Wales gained on average more marks than the OECD average in three multi-mark items, whilst in 2015 Wales' facility rates were similar or lower than the OECD average in all multi-mark science items. This indicates that learners in Wales gave more complete and/or better responses to multi-mark items than in 2015.

References

Andrade, J., de Lazzari, G. and Wheeler, R. (2017). *PISA 2015 Welsh Item Analysis*. Slough: NFER.

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

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

National Foundation for Educational Research (2015). *PISA 2012 Wales Item Analysis*. Slough: NFER.

Organisation for Economic Co-operation and Development (2019a). *PISA 2018 Assessment and Analytical Framework* [online]. Available: https://www.oecd-ilibrary.org/education/PISA_2018_Assessment_and_Analytical_Framework_b25efab8-en [28 October, 2020].

Organisation for Economic Co-operation and Development (2019b). *PISA Technical Report* [online]. Available: https://www.oecd.org/pisa/data/PISA_2018_Technical_Report/ [28 October, 2020].

Sizmur, J., Ager, R., Bradshaw, J., Classick, R., Galvis, M., Packer, J., Thomas D. and Wheeler R. (2019). *Achievement of 15-year-olds in Wales: PISA 2018 National Report* [online]. Available: https://gov.wales/sites/default/files/statistics-and-research/2019-12/achievement-15-year-olds-program-international-student-assessment-PISA_2018_National_Report_for_Wales_0.pdf [28 October, 2020].

Appendix A: Item composition of the PISA 2018 assessment.

There were 36 forms of the assessment which were randomly assigned to learners. Each form consisted of four 30-minute clusters (set of reading, mathematics or science items). All learners responded to 60 minutes of reading items. The content of the other 60 minutes of the test varied: 46 per cent of learners were administered two clusters of mathematics items, 46 per cent were administered two clusters of science items, and eight per cent were administered one cluster of mathematics and one cluster of science items. (OECD, 2019b)

Before 2015, there were significantly fewer items in the minor domains than the major domain. Since 2015, and the introduction of computerised assessment, a new approach has been implemented to improve the measurement of trends, by including a greater number of items.

Table A. 1 Numbers of items by domain across cycles

Domain	PISA cycle				
	2006	2009	2012	2015	2018
Reading	28	131*	44	103	245*
Mathematics	48	35	109*	83	82†
Science	103*	53	53	184*	115

Note: Number of items in the major domain marked with an asterisk.

† There were two versions of one mathematics cluster: M6A (standard) and M6B (easier). Learners in Wales and comparator countries in this report were presented with M6B cluster for a total pool of 70 mathematics items.

Source: PISA 2018 school census matched database

Evidence for excellence in education

© National Foundation for Educational Research 2020

All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, or otherwise, without prior written permission of NFER.

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. WPAC

ISBN. 978-1-912596-23-2

