PIRLS and ePIRLS 2016
Test content and Irish pupils' performance

Eemer Eivers
Mary Delaney

2018

ERC Research Series: Report 4

Educational Research Centre
Foras Taighde ar Oideachas

PIRLS and ePIRLS 2016

Test content and Irish pupils' performance

Eemer Eivers
Mary Delaney

2018

ERC Research Series: Report 4

Educational Research Centre
Foras Taighde ar Oideachas

www.erc.ie

Cataloguing-in-publication data:

Eivers, Eemer.

PIRLS and ePIRLS 2016: test content and Irish pupils' performance
Eemer Eivers and Mary Delaney.
Dublin: Educational Research Centre.

iv, 44 pages: illustrations, tables; 30cm
ERC research series; report 4

ISBN 978 0 900440 70 0 – (print)
ISBN 978 0 900440 71 7 – (pdf) ISSN 2565-5604

1. Progress in International Reading Literacy Study.
2. Reading comprehension– Ireland - Evaluation.
3. Reading (Primary) – Ability testing – Ireland.
4. Educational tests and measurements – Ireland – Data processing.
5. Academic achievement – Ireland.
6. Sex differences in education.
7. Gender.

2018
I Title II Eivers, Eemer III Delaney, Mary
372.4809417 dc23

Design: Silverbark Creative.
Printed in the Republic of Ireland by Naas Printing Limited.

Table of Contents

Back to Contents

Acknowledgements

PIRLS is administered in Ireland by the Educational Research Centre with the support of a National Advisory Committee. The members of the Committee advising on the implementation of PIRLS 2016 were as follows (in alphabetical order):

- Aedín Ní Thuathail (Irish Primary Principals' Network).
- Aidan Clerkin (Educational Research Centre).
- Áine Lynch (National Parents Council – Primary).
- Arlene Forster (National Council for Curriculum and Assessment).
- Clare Griffin (Department of Education and Skills) (Chair) (replacing Yvonne Keating).
- Deirbhile Nic Craith (Irish National Teachers' Organisation).
- Eamonn Moran (Department of Education and Skills).
- Eddie Fox (Educate Together)
- Eemer Eivers (Educational Research Centre).
- Máirín Ní Chéileachair (Gaelscoileanna).
- Mia Treacy (Professional Development Service for Teachers).
- Seán Delaney (Marino Institute of Education).
- Suzanne Cobbe (Catholic Primary Schools Management Association).

The authors thank the National Advisory Committee for their contributions, and Lorraine Gilleece, Emer Delaney and Brenda Donohue for their key roles in implementing the study. We also extend our thanks to the pupils, parents, teachers and principals who took part in PIRLS 2016.

Introduction

This report examines performance on reading items included as part of PIRLS 2016 (Progress in International Reading Literacy Study) and its related online reading assessment, ePIRLS. It has a particular focus on the responses of pupils in Ireland, including items on which Irish pupils performed particularly well or poorly, and on which gender differences were markedly different to the general pattern of girls outperforming boys (in Ireland, and internationally).

The rest of the report is divided into eight parts, the first of which summarises the main features of PIRLS and ePIRLS and how they define reading literacy. The second section examines commonalities between the PIRLS and ePIRLS assessment frameworks and the framework underpinning recent National Assessments of English Reading, the Primary School English Curriculum (DES/NCCA, 1999a; 1999b; NCCA, 2005), and the more recent Department of Education and Skills' literacy and numeracy strategy (DES, 2011). Section three summarises Irish performance on PIRLS and ePIRLS, in terms of overall performance, subscale scores, and pupils reaching various International Benchmarks. The fourth section describes how pupils rated their enjoyment of the various texts and projects, while the fifth examines differences in item difficulty by various item characteristics (e.g., multiple choice versus constructed response, reading purpose, cognitive process, and the nature of text content). The sixth section provides examples of texts and items used as part of PIRLS – content that has been released, subsequent to the completion of PIRLS 2016. It also contains information about the percentages of pupils in Ireland and internationally who correctly answered selected items, with a particular focus on atypical items. The seventh section contains similar information about ePIRLS released content. The concluding section suggests some implications for curriculum and instruction.

As noted, a number of texts were released at the end of PIRLS 2016. Some are presented later in this report but, for reasons of space, not in the more visually appealing format presented to pupils. To see content as presented to pupils, see http://www.erc.ie/pirls. Clicking **sample questions** shows two texts, with commentary and explanation of how each item is scored, and information about the percentages of pupils in Ireland and internationally who answered each item correctly. Other texts can be accessed from the international report (Mullis, Martin, Foy, & Hooper, 2017a, Appendix H), while sample ePIRLS interactive items can be seen at http://timssandpirls.bc.edu/pirls2016/international-results/epirls/take-the-epirls-assessment/.

Readers should note that this report examines only one aspect of the PIRLS 2016 data. Those who would like more general information about PIRLS 2016 are referred to Eivers, Gilleece and Delaney (2017) for a focus on Ireland's performance, and to the international reports on PIRLS (Mullis et al., 2017a) and ePIRLS (Mullis, Martin, Foy, & Hooper, 2017b) for a more international perspective.

> **NOTE: Throughout this review, the term "item" is used instead of the more familiar term "question", as not all test items are phrased as questions.**

What are PIRLS and ePIRLS?

PIRLS is the largest international comparative assessment of reading achievement at primary school level. It assesses pupils in their fourth year of formal schooling (Mullis et al., 2017a). First run in 2001, it takes place every five years, with 50 countries and 11 benchmarking regions[1] taking part in the 2016 cycle. The most recent cycle included a new element called ePIRLS, an assessment of online reading skills. Ireland was one of 14 countries and two benchmarking regions that participated in both PIRLS and ePIRLS. As they use different modes of assessment and are designed to assess related, but different, skills, there is no overlap in test content between PIRLS and ePIRLS.

Both assessments are guided by a framework that defines who and what is to be measured, and how measurement will take place. The framework identifies children in their fourth year of formal schooling as the target for the assessment, as that is the point at which pupils have learned how to read and are beginning to read to learn (Mullis, 2015). Reading literacy is defined as follows:

> "… the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment."
>
> *(Mullis, Martin & Sainsbury, 2015, p.12)*

Thus, the definition emphasises reading comprehension rather than reading fluency, recognises that there are many types of text (including digital forms), and sees reading as an active, meaning-making, and social process.

The framework also identifies the main purposes for which children read, and comprehension processes they use when doing so. The reading purposes specified are: *reading for literary experience,* and, *reading to acquire and use information* (hereafter referred to as Literary and Informational). The comprehension processes are: *focus on and retrieve explicitly stated information; make straightforward inferences; interpret and integrate ideas and information;* and, *evaluate and critique content and textual elements*. For reporting purposes these were combined into two subscales (Retrieve/Infer and Interpret/Evaluate). The emphases placed on each purpose and process in PIRLS and ePIRLS is summarised in Table 1.

1 Benchmarking regions are entities (usually cities or states) that follow PIRLS quality standards and procedures. For example, Canada included two additional samples to provide more detail on performance in Ontario and Quebec.

Table 1: Reading purposes and processes in the PIRLS and ePIRLS assessment frameworks, and associated subscales

Framework Element	PIRLS Subscales	ePIRLS Subscales
Purposes		
Literary experience Acquire and use information	Literary (50%) Informational (50%)	Informational (100%)
Processes		
Retrieve explicitly stated information Make straightforward inferences	Retrieve/Infer (50%)	Retrieve/Infer (50%)
Interpret and integrate ideas and information Examine and evaluate content, language and textual elements	Interpret/Evaluate (50%)	Interpret/Evaluate (50%)

PIRLS is designed to assess both Literary and Informational reading skills. The assessment is in paper format, and uses a matrix sampling technique, in which 12 texts are systematically presented across 16 different test booklets (Martin, Mullis & Foy, 2015). Each booklet contains one Literary and one Informational text. Total testing time is a maximum of 80 minutes, with a break between texts. PIRLS provides five main outcome scores (an overall reading score, and scores for two purpose and two process subscales).

In contrast, the ePIRLS assessment is delivered via a computer and is comprised of Informational texts only (Martin et al., 2015). Pupils complete two of five randomly assigned "ePIRLS projects". Projects require navigation through webpages in a closed, fake internet environment, containing features such as tabs, pop-ups and hyperlinks. Pupils are guided through the project by a teacher avatar who prompts with questions about the online content. As with the paper-based assessment, total testing time is a maximum of 80 minutes, with a break between projects. ePIRLS provides three main outcome scores (an overall [Informational] reading score, and two process subscale scores).

PIRLS in an Irish context

In many ways, there is considerable overlap between how reading is defined in PIRLS and how it is defined in Ireland's National Assessments[2] and Primary School English Curriculum (PSEC). For example, the definition of reading that underpinned the 2004, 2009 and 2014 National Assessments of English Reading also notes that reading is a constructive, social, process, and that children read for enjoyment:

> "the process of constructing meaning through the dynamic interaction among the reader's existing knowledge, the information suggested by the written language, and the context of the reading situation. Young readers read to learn, to participate in communities of readers, and for enjoyment.
>
> *(Eivers, Shiel, Perkins & Cosgrove, 2005, p.15)*

In contrast to the PIRLS definition, the definition in the National Assessments also refers to the influence of existing knowledge on how text is interpreted, but does not explicitly reference a variety of text sources, such as online texts. The PSEC (DES/NCCA, 1999a; 1999b) refers to the need for pupils to encounter a variety of genres, and emphasises the importance of fostering positive attitudes towards reading, and discussing texts. Unsurprisingly for documents released in 1999, there is little emphasis on digital or online reading, and relatively little emphasis in the additional support material published in 2005 (NCCA, 2005). However, the Department of Education and Skills' more recent literacy and numeracy strategy, and the new Primary Language Curriculum (for up to Second class) explicitly include digital media in their definitions of literacy and text (DES, 2011; DES/NCCA, 2015).

The assessment framework for reading in PIRLS and in the National Assessments also share similarities in the balance of text types. As with PIRLS, test content in the two most recent National Assessments (Eivers et al., 2010; Shiel, Kavanagh & Millar, 2014), has been evenly divided between Literary and Informational texts. However, the two most recent National Assessments found narrative texts and children's literature were far more likely to be regularly used than were real-life texts or reference materials, while use of digital texts has remained quite limited (Eivers et al., 2010; Kavanagh, Shiel, Gilleece & Kiniry, 2015). In addition, there remained a reliance on reading schemes (many or most of which are largely composed of literary types of texts).

Data from PIRLS 2016 also suggest that Irish pupils were slightly less likely than the PIRLS average to be regularly exposed to a variety of non-fiction resources, but more likely to be exposed to fiction materials. For example, 39% of Irish pupils encountered longer fiction books as part of their reading lessons on a daily or almost daily basis, compared to only 15% of pupils internationally.

Unlike PIRLS, the National Assessments and the PSEC place relatively more emphasis on retrieval skills. This is particularly true at Second class, where the National Assessments contain no Examine & Evaluate items, and relatively few Interpret and Integrate items (on the basis that the

2 The National Assessments are periodic evaluations of English reading, conducted since 1972. Originally, the assessment of reading and mathematics were conducted separately. However, since 2009, the same grades (and pupils) are assessed on both reading and mathematics.

PSEC does not expect these processes to be taught until Third and Fourth class). The Sixth class element of the National Assessments places relatively more emphasis on the two "higher order" comprehension skills (again, in line with the PSEC), but still falls well short of the 50% of items in the PIRLS assessment. Thus, pupils in Ireland taking part in PIRLS may have had limited classroom experience with these processes.

The PSEC encourages teachers of Third and Fourth class to get pupils to argue and discuss ideas and assimilate, deduct, infer, analyse, predict, evaluate and summarise when reading. However, in relation to writing there is a continuing focus on narrative and composition at this level. Writing to present an argument or for persuasion, evaluation and critique is not encouraged until the next stage of the curriculum. In terms of how the PSEC is reflected in practice, data from PIRLS 2016 show that Irish pupils are slightly less likely than the international average to be regularly taught how to summarise main ideas, but more likely to regularly be asked to identify the main ideas of what they have read, and to challenge opinions expressed in texts.

Summary of Ireland's performance

Irish pupils performed very well on PIRLS and ePIRLS (Eivers et al., 2017; Mullis et al., 2017a, 2017b). The Irish mean score for PIRLS was 567, significantly above the study centrepoint of 500. Only two countries – the Russian Federation and Singapore – had significantly higher mean scores than Ireland. Four countries (Hong Kong SAR, Finland, Poland and Northern Ireland) achieved means that did not differ significantly from Ireland's mean, while the Irish mean was higher than the mean in 43 countries. With a mean score of 571, Irish pupils showed a relative strength on the Literary subscale, although the mean of 565 on the Informational subscale was also very high.

Ireland's performance on PIRLS in 2016 represented a significant improvement on performance in PIRLS 2011, when Ireland's mean score was 552. The 2011 performance was well above the PIRLS international centrepoint of 500, but significantly lower than the mean achieved in 2016. In 2011, five countries achieved mean scores that were significantly higher than Ireland's, eight did not differ significantly, and 31 countries had a significantly lower mean. As in 2016, Ireland demonstrated a relative national strength on Literary texts.

On ePIRLS, Singapore's mean score of 588 was significantly higher than the mean scores in all other countries. The next highest-achieving countries were Norway and Ireland. With almost identical means (of 568 and 567, respectively), Norway and Ireland significantly outperformed all remaining countries and benchmarking participants. As 2016 was the first time ePIRLS took place, no trend data are available.

Ireland and Canada were the only two countries where there was no significant national-level difference between performance on PIRLS and ePIRLS. In seven participating countries, the average ePIRLS national mean score was significantly higher than PIRLS, while in five countries, achievement in PIRLS was significantly higher than in ePIRLS.

Gender

On PIRLS, gender differences in Ireland were smaller than the international average: a 12-point gap in favour of girls, compared to 19 points on average internationally (Table 2). In Ireland the gender gap was most pronounced on the Literary subscale (17 points) and smallest on the Informational subscale (8 points). The PIRLS average gender gap was also largest for Literary texts (23 points), but the smallest gender gap internationally was found on the Retrieve/Infer subscale (a 14-point gap).

Examining trends by gender, both boys and girls in Ireland showed significant improvements since 2011. However, the magnitude of improvement was larger for boys than girls (17 points versus 13 points, respectively), meaning that the gap has reduced in Ireland.

Table 2: Summary gender differences in Ireland and internationally, PIRLS						
	Ireland			International		
	Girls	Boys	Gap	Girls	Boys	Gap
Overall	572	561	12	520	501	19
Literary	580	563	17	522	499	23
Informational	569	561	8	519	503	16
Retrieve/Infer	571	561	10	520	503	17
Interpret/Evaluate	576	562	14	520	500	20

On ePIRLS, the gender gap in Ireland (11 points) was about the same as the international average gap (12 points) (Table 3). The gender gap was larger in Ireland and internationally for mean scores on Retrieve/Infer than for Interpret/Evaluate.

Table 3: Summary gender differences in Ireland and internationally, ePIRLS						
	Ireland			International		
	Girls	Boys	Gap	Girls	Boys	Gap
(Informational)	572	561	11	545	533	12
Retrieve/Infer	572	559	13	548	533	15
Interpret/Evaluate	573	563	10	542	533	9

In almost all countries, the results for boys and girls on PIRLS Informational reading and ePIRLS scores were in line with national results.[3] As such, if there was a national strength on PIRLS or ePIRLS, both boys and girls reflected that national pattern. In Ireland, however, girls showed a slight relative strength (5 points) on ePIRLS versus PIRLS Informational reading, while boys' performance was almost identical on both measures.

As Ireland's overall performance was well above average, Irish pupil performance on most test items was also well above average, as will be apparent in the sections examining performance on selected items.

Performance at the International Benchmarks

PIRLS and ePIRLS also report performance at four International Benchmarks:

- Advanced (625 points)
- High (550)
- Intermediate (475)
- Low (400)

The International Benchmarks include descriptions of what pupils scoring at each Benchmark can achieve, and estimates of the percentages of pupils in each country achieving them. For detailed information about how Benchmark descriptions were created, see Mullis and Prendergast (2017), and see Eivers et al. (2017) or Mullis et al. (2017a, 2017b) for a fuller description of the skills demonstrated by pupils at each Benchmark. Summary descriptions are provided in Tables 4 and 5.

3 Comparisons of ePIRLS and PIRLS performance is based *only* on the subset of pupils who completed both assessments.

Table 4: Summary of pupils' reading skills at each PIRLS International Benchmark, by text purpose

Level	**Literary Text:** *Pupils can…*	**Informational Text:** *Pupils can…*
Advanced *(625 points)*	■ interpret events and actions to describe reasons, feelings, and character development. ■ begin to evaluate the effect of the author's language and style choices.	■ integrate complex information from different parts of text, to explain relationships and sequences. ■ begin to evaluate visual and textual elements to consider the author's view point.
High *(550 points)*	■ distinguish significant actions and details embedded across the text. ■ infer relationships between intentions, actions, events, and feelings. ■ integrate events and actions, traits, and feelings across the text. ■ recognise the use of metaphor, tone, imagery.	■ distinguish relevant information in a text or a complex table, infer logical connections to provide explanations. ■ integrate textual and visual information to interpret the relationship between ideas. ■ evaluate and make generalisations about context and textual elements.
Intermediate *(475 points)*	■ locate and reproduce explicitly stated actions, events, and feelings. ■ make simple inferences about attributes, feelings, and motivations. ■ interpret obvious reasons, recognise evidence, and give examples. ■ begin to recognise language choices.	■ locate and reproduce two or three pieces of information from text. ■ make straightforward inferences to provide factual explanations. ■ begin to interpret and integrate information to order events.
Low *(400 points)*	■ retrieve explicitly stated information, actions, or ideas. ■ make simple inferences about events and reasons for actions. ■ begin to interpret story events and central ideas.	■ reproduce explicitly stated information from text and other formats (e.g., chart). ■ begin to make straightforward inferences about explanations, actions, and descriptions.

Content adapted from Exhibits 2.3, 2.4, 2.5, and 2.6 in the main PIRLS report (Mullis et al., 2017a)

Table 5: Summary of the skills that pupils display at each of the ePIRLS International Benchmarks

Level	**When reading Online Informational Texts,** *pupils can…*
Advanced *(625 points)*	■ make inferences from complex information to support an explanation. ■ interpret and integrate information to explain relationships, and show thorough understanding. ■ evaluate the effects of textual, visual, and interactive elements and begin to consider the writer's point of view.
High *(550 points)*	■ infer relevant information and provide comparisons. ■ integrate information to provide examples and make contrasts. ■ evaluate how graphic elements and language choices support content.
Intermediate *(475 points)*	■ reproduce information presented in various forms, including independent use of navigation features. ■ make straightforward inferences to recognise reasons and actions. ■ integrate information to recognise causes, comparisons, and explanations. ■ begin to evaluate the use of interactive features to convey information.
Low *(400 points)*	■ reproduce explicitly stated information from sites that contain text and a variety of dynamic, navigable features (e.g., timelines, pop-up boxes). ■ begin to make straightforward inferences about descriptions.

Content adapted from Exhibits 2.2, 2.3, 2.4, and 2.5 in the main ePIRLS report (Mullis et al., 2017b)

Benchmark skills are cumulative, meaning that pupils achieving the upper Benchmarks are also able to demonstrate the skills associated with the lower level Benchmarks. Conversely, pupils at the Low and Intermediate International Benchmarks are unable to consistently display the skills associated with higher levels. The descriptions of the skills associated with each level are different for PIRLS Literary texts, PIRLS Informational texts, and for ePIRLS – because the skills required to read these types of texts are different. However, all three sets of International Benchmarks are comparable as the same cutpoints are used on each to categorise achievement levels.

Looking first at PIRLS, almost all pupils in Ireland (98%) and internationally (96%)[4] demonstrated the skills associated with the lowest Benchmark (Table 6). This means that virtually all pupils in Ireland had successfully mastered the most basic reading skills. Most pupils in Ireland were also able to demonstrate the skills associated with the Intermediate Benchmark (89% versus an international median of 82%), while 62% reached the High Benchmark, considerably higher than the international median of 47%. At the highest level – the Advanced International Benchmark – 21% of pupils in Ireland reached this level, more than double the international median value of 10%.

A small number (2%) of Irish pupils did not reach the Low International Benchmark, meaning that PIRLS cannot describe their reading skills. This percentage is lower than other assessments of reading literacy (e.g., the National Assessments, PISA), partly due to how Benchmarks were established in PIRLS, but also due to the strong performance of Irish pupils on PIRLS.

The percentages of Irish pupils reaching the higher Benchmarks in ePIRLS was also well above the international median (Table 6). Twenty percent of pupils in Ireland reached the Advanced International Benchmark (international median: 12%), while 63% reached the High International Benchmark (international median 50%). As with PIRLS, almost all Irish pupils reached at least the Low Benchmark, just above the international median of 97%, while 90% of pupils in Ireland reached the Intermediate Benchmark, compared to the ePIRLS median of 84%.

Table 6: Cumulative percentage of pupils, Ireland and the international median, reaching the PIRLS and ePIRLS International Benchmarks for overall reading

		Low	Intermediate	High	Advanced
PIRLS	Ireland	98	89	62	21
	International median	96	82	47	10
ePIRLS	Ireland	98	90	63	20
	International median	97	84	50	12

4 International values refer to the median, not the mean. Median values are used as they divide countries evenly. For example, the median value for Low is 96%. Thus, in half of participating countries more than 96% of pupils reached this Benchmark, and in half, fewer than 96% did so.

Pupil enjoyment of texts

For each text or project, pupils were asked how much they had liked reading it (a lot / a little / not very much / not at all). One of the main reasons for including the "liking" questions is to ensure that PIRLS content is appealing to most children. However, it also allows some examination of the relationship between interest or engagement and performance on the test, identifies what children find engaging reading material, and flags any national or gender differences in what is considered engaging. There is some evidence that boys' reading scores may be more influenced by the topic and by the main character's gender than girls' scores (e.g., Taube & Munck, 1996) and that boys' enjoyment of a text may have a greater effect on their level of performance (Ainley, Hillman, & Hidi, 2002; Logan & Johnson, 2009). Also, in the context of an international study, there are likely to be cultural differences in what topics are considered interesting.

Pupils were generally quite positive about the texts and projects. For example, across all countries, approximately 90% of pupils liked ePIRLS tasks (*a little or a lot*), while liking of PIRLS texts ranged from 80-93%, with ratings high for the relatively easier PIRLS Literacy texts.[5] In Ireland, pupil ratings for the paper-based texts were slightly less positive than the international averages, and slightly more positive for the computer-based projects (Tables 7 & 8).

For PIRLS texts and the relatively easier PIRLS Literacy texts, Irish pupils' ratings were the same as the international average on two texts, and slightly less positive on most (Table 7). For three texts (*Macy and the Red Hen, Icelandic Horses, Pemba Sherpa*) 7% fewer pupils in Ireland than internationally indicated that they liked reading the texts. These texts represent a mixture of Literary and Informational content, and their ratings are depressed because Irish boys, in particular, were less likely to rate them positively. The two Literary texts (*Pemba Sherpa* and *Macy and the Red Hen*) feature female protagonists. However, Irish boys also gave much less positive ratings than girls to *Flowers on the Roof* (an elderly female and a young male protagonist), *Empty Pot* (almost exclusively male characters), and *Oliver and the Griffin* (exclusively male characters). Thus, it seems that Irish boys are less positively disposed to a variety of Literary texts, irrespective of the gender of the characters. *Shiny Straw* was the only Literary text for which both boys and girls in Ireland expressed similar levels of liking (84% and 85%, respectively). It is also the only Literary text in which the main characters are animal, not human.

In contrast, boys generally, and in Ireland in particular, were more positively disposed than girls towards many Informational texts. Irish girls gave their lowest ratings (73-75% liked) to three Informational texts (*Where's the Honey, Icelandic Horses,* and *Sharks*). While *Icelandic Horses* was also relatively unpopular with Irish boys (71% liked), they were more favourably disposed to *Where's the Honey* (81% liked) and far more likely than Irish girls to have liked a text about *Sharks* (83% liked). The low rating by Irish girls for *Sharks* meant the Irish rating was 8% below the PIRLS average.

5 PIRLS Literacy was an add-on to the main PIRLS study. Countries whose pupils are still developing basic reading skills could take part in PIRLS Literacy and have their results reported on the PIRLS achievement scale. This was possible as some texts were common to PIRLS Literacy and PIRLS. For example, a subset of PIRLS pupils received one easier PIRLS Literacy text and one PIRLS text in their booklet.

Table 7: Percentages of pupils who liked PIRLS and PIRLS Literacy passages either *A little* or *A lot*, Ireland and PIRLS average, by gender

	Type	PIRLS				Ireland			
		Overall	Girls	Boys	Gap	Overall	Girls	Boys	Gap
Shiny Straw	Lit	88	90	87	3	85	85	84	1
Macy and the Red Hen (R)*	Lit	86	89	83	6	79	82	75	7
The Empty Pot	Lit	89	93	86	7	84	89	79	10
Oliver and the Griffin	Lit	86	89	83	6	84	88	81	7
Flowers on the Roof (R)*	Lit	92	96	89	7	90	95	85	10
Leonardo Da Vinci (R)*	Inf	84	83	84	-1	84	82	86	-4
Green Sea Turtles (R)*	Inf	89	90	89	1	89	87	90	-3
Where's the Honey?	Inf	81	79	82	-3	78	75	81	-6
Icelandic Horses	Inf	80	82	77	5	73	74	71	3
Sharks	Inf	86	83	89	-6	78	73	83	-10
Shared PIRLS Literacy Passages									
	Type	PIRLS				Ireland			
		Overall	Girls	Boys	Gap	Overall	Girls	Boys	Gap
Pemba Sherpa	Lit	92	94	90	4	85	89	79	10
Learn to Fly?	Inf	93	92	93	-1	87	84	90	-6

(R)* Text was released in 2017 and can be accessed from www.erc.ie/pirls or https://timssandpirls.bc.edu/

On each of the five ePIRLS projects, slightly more Irish pupils than the international average indicated that they liked the project (Table 8). The gap was largest for a project about *Elizabeth Blackwell* (90% of pupils in Ireland versus an ePIRLS average of 83%). This may be partly attributable to 85% of boys in Ireland indicating that they liked the project, compared to 78% internationally, the lowest rating from boys for any project. Apart from the Blackwell project, where the main theme was the struggle of *Elizabeth Blackwell* to become one of the world's first qualified female doctors, gender differences were relatively small, and both boys and girls were very positive in their ratings for all ePIRLS projects. Two projects were slightly more popular with Irish boys than with girls; a project about *Mars* (discussed later) (liked by 93% of boys versus 89% girls) and one about *The Legend of Troy* (93% versus 91%).

Table 8: Percentages of pupils who liked ePIRLS projects either *A little* or *A lot*, Ireland and ePIRLS average

	ePIRLS				Ireland			
	Overall	Girls	Boys	Gap	Overall	Girls	Boys	Gap
Mars (R)*	88	87	89	-2	91	89	93	-4
Rainforests	93	94	92	2	96	96	95	1
Elizabeth Blackwell (R)*	83	88	78	10	90	94	85	10
Migration	92	93	92	1	93	94	93	1
The Legend of Troy	89	89	90	-2	92	91	93	-2

(R)* Project was released in 2017 and can be accessed as an interactive test from https://timssandpirls.bc.edu/

Difficulty by item characteristics

This section provides an overview of the performance of pupils in Ireland and internationally on different types of items, in terms of the percent of pupils answering items correctly. Overall performance and performance by gender is examined for items grouped by associated text, reading purpose, comprehension process, item type (multiple choice or constructed response[6]), and item content. Because only general item characteristics are discussed, these analyses can draw on *all* test content – that is, including content and items that remain secure and confidential, as they will again be used in the next cycle of PIRLS or ePIRLS. In contrast, the analyses in the subsequent sections draw on the content of specific items and are therefore based solely on items that were released after the completion of PIRLS and ePIRLS 2016.

The analyses presented are based on the PIRLS and ePIRLS achievement data Item Percent Correct Statistics, released as a supplement to the PIRLS 2016 User Guide for the International Database (Foy, 2018). The supplement provides weighted summary statistics for each participating country on each achievement item included in the PIRLS and ePIRLS 2016 assessments. They also display the international averages for each item, with each country weighted equally. All international averages are based on these source documents (or averaged from data in them), whereas some Irish statistics that examine gender differences are calculated from the Irish dataset using the IEA's IDB Analyser (IEA, 2018).

As this review is largely a descriptive analysis of items, standard errors (a measure of the precision of an estimate) are not provided for the percentages reported here. As a rough guide, standard errors for PIRLS international average percent correct values range from 0.1 – 0.3 while Irish national values range from 0.1 – 2.7, with most falling below 2.0. Full details for each item are available from https://timssandpirls.bc.edu/pirls2016/international-database/index.html

Text/Project and Purpose

Looking first at the texts and projects associated with items, it is clear that some were more difficult than others (Tables 9 and 10). On PIRLS, two Informational texts (*Icelandic Horses, Leonardo da Vinci*) proved the most difficult, with only 56% of items on average answered correctly by pupils in Ireland (the international averages were 48% and 47%, respectively). In contrast, for the PIRLS Literacy texts of *How did we Learn to Fly* and *Pemba Sherpa*, 85% and 81%, respectively, of items were correctly answered by Irish pupils. These two texts were also the easiest, internationally. Because some texts are considerably easier than others, it is necessary to look at Irish performance *relative* to the international average, rather than comparing how Irish pupils performed across different texts.

Compared to the PIRLS average, Irish pupils showed a slightly larger advantage on Literary texts than on Informational texts. Overall, the percent of items answered correctly by Irish pupils was 9% higher for Informational texts and 11% higher for Literary texts. This is not surprising, given that the scaled achievement scores showed that Irish pupils had a statistically significant national strength

6 Constructed response items are those that require pupils to write a response, rather than choosing one of four multiple choice response options. The response length can vary from a single word to a few sentences.

on Literary texts, relative to Informational. Nonetheless, Irish pupils demonstrated a relative strength on one Informational text, *Green Sea Turtles*, and a relative weakness on one Literary text, *Pemba Sherpa*. Thus, the nature of the text content as well as the purpose of the text both contribute to performance on the related items. Also, *Green Sea Turtles* is one of only two texts for which the percent of Irish pupils who liked it was as high as the PIRLS international average. It is one of two texts that are presented in full later in this review.

While girls tended to answer more items correctly than boys, there was some variation in gender differences by text and by purpose. The gender gap tended to be slightly smaller for Informational texts, in Ireland and internationally (again, reflecting different gender gaps by subscales). On one Informational text (*Green Sea Turtles*), Irish boys answered marginally more items correctly than did Irish girls, but across PIRLS as a whole, girls did slightly better than boys (by 4%) on this text. On one Literary text (*Macy and the Red Hen*), boys and girls in Ireland obtained the same percent correct answers, whereas internationally, girls outperformed boys by 3%.

The advantage demonstrated by Irish pupils over the international average was most pronounced on two texts, *Flowers on the Roof*, and, *Oliver and the Griffin*. Relative to the PIRLS average, Irish pupils answered 15% more items correctly for *Flowers on the Roof* and 13% more for *Oliver and the Griffin*. *Flowers on the Roof* was also the text with the highest liking rating among Irish pupils. The advantage was smallest for *Pemba Sherpa* (7%), followed by *How did we Learn to Fly* and *Icelandic Horses* (both 8%). As noted earlier, three texts (*Macy and the Red Hen, Icelandic Horses, Pemba Sherpa*) were markedly less popular with Irish pupils. Two of the three are those where Irish pupils show a smaller than average advantage, suggesting that interest and enjoyment may play a part in performance on these texts.

Table 9: Percentages of PIRLS items answered correctly by pupils in Ireland and internationally, by text, text type, with gender difference

		Ireland	PIRLS	IRL advantage	Girl – Boy	
					Ireland	PIRLS
Informational	Learn to Fly	85.0	76.8	8.2	0.9	2.5
	Leonardo da Vinci	56.2	47.4	8.8	3.9	2.6
	Sharks	61.2	51.5	9.8	1.9	3.4
	Where's the Honey	58.7	48.5	10.1	3.2	2.3
	Icelandic Horses	55.8	47.6	8.2	1.9	2.9
	Green Sea Turtle	63.4	52.0	11.4	-0.8	3.7
	Overall	64.1	54.7	9.4	1.7	2.9
Literary	Pemba Sherpa	81.0	73.9	7.1	1.9	3.2
	Flowers	79.3	64.7	14.6	6.7	6.2
	Shiny Straw	66.7	56.1	10.5	1.2	2.3
	Empty Pot	71.1	61.8	9.3	6.0	5.0
	Macy & Hen	63.9	51.9	12.0	0.0	3.4
	Oliver & Griffin	66.0	52.8	13.3	5.5	5.1
	Overall	71.4	60.6	10.8	3.5	4.2

In contrast to PIRLS, the percent of ePIRLS items correctly answered by Irish pupils was closer to the ePIRLS average, and showed less variation across projects (Table 10). The lowest percent correct value for Irish pupils was on the *Elizabeth Blackwell* project (57%) and the highest was on *Migration and Rainforests* (66%), while, internationally, values ranged from 49% (*Blackwell*) to 60% (*Rainforests*). Generally, the percent correct value for Irish pupils was about 6-8% higher than the

ePIRLS average. The advantage demonstrated by Irish pupils was smallest for *Rainforests* (just below 6%), despite this being the project most liked by Irish pupils (96% liked it a little or a lot).

Table 10: Percentages of ePIRLS items answered correctly by pupils in Ireland and internationally, by project, with gender difference.

	Ireland	ePIRLS	IRL advantage	Girl – Boy	
				Ireland	ePIRLS
Mars	62.4	55.2	7.2	0.8	1.0
Rainforests	65.6	59.8	5.8	3.9	3.0
Elizabeth Blackwell	56.8	49.2	7.6	6.2	5.1
Migration	65.7	59.4	6.3	2.8	2.5
Legend of Troy	63.9	56.2	7.7	0.9	1.7
Overall	63.2	56.4	6.8	2.8	2.5

On average, Irish girls answered more items correctly than boys across all the projects, as did girls across all ePIRLS countries. The gender gap was smallest on *Mars* and *The Legend of Troy* (both 1%) and highest on *Blackwell* (6%), differences largely mirrored internationally. *Mars* and *The Legend of Troy* are the two projects that Irish boys liked slightly more than did Irish girls, while *Elizabeth Blackwell* was the project least liked by boys. Thus, as may be the case with some texts in PIRLS, there may be a link between enjoyment and performance on some ePIRLS projects.

Item Type

PIRLS and ePIRLS use two main types of items – constructed response and multiple choice. There is a body of research indicating that item type may be related to differences in performance by characteristics such as gender, language of the test, and country. (e.g., Lafontaine & Monseur, 2009; Le, 2006; Lissitz, Hou, & Cadman Slater, 2012; Reardon, Kalogrides, Fahle, Podolsky & Zarate, 2018). Generally, multiple choice items are more likely to be answered than are constructed response items. Boys do relatively less well on constructed response than on multiple choice items, particularly in reading tests, or where a relatively long response is required. However, some (e.g., Solheim & Lundetræ, 2016) suggest that this may be related to the finding that lower-achieving pupils tend to perform better on multiple choice than on constructed response items. There are also country-level and language differences in performance by item type, and by gender. For example, Lafontaine and Monseur's analyses of PISA 2000 data not only suggested that a test composed of 100% of open-ended items will lead to a gender gap 54% larger than a test including only multiple-choice items, but that the predicted change in gender gap would vary widely by country – from relatively little effect in Portugal to increasing it by 114% in Korea. However, as most research thus far has been based on paper-based assessments, it is less clear if item type effects also manifest on a computer-based test such as ePIRLS.

Looking first at non-response in PIRLS, constructed response items were more likely to be skipped than were multiple choice items, both in Ireland and across all participating countries. In Ireland, 3.4% of constructed response items were not answered, on average, higher than the 0.9% of multiple choice items. Internationally, 6.7% of constructed response were skipped, compared to 1.5% of multiple choice items. For ePIRLS, constructed response items were again more likely to be skipped than multiple choice items both in Ireland and internationally. On average, 3% of constructed response items and 1.4% of multiple choice items were left blank by Irish pupils. Internationally, pupils skipped 5.4% of constructed response items and 2.5% of multiple choice items.

For constructed response items, there was a pattern of higher non-response in lower performing countries, but also, considerable national differences in rates of non-response. Thus, a high rate of skipping was evident in countries such as Iran, Saudi Arabia, Azerbaijan, and Malta (all relatively low performing), but also in certain mid-ranking countries such as Austria, Czech Republic, France and the French-speaking part of Belgium (but not the Flemish part of Belgium). This suggests that, as well as reading skills, cultural factors may have a role to play in how pupils respond – or not – to constructed response items.

Tables 11 and 12 show the percentages of multiple choice and constructed response items answered correctly in Ireland and internationally. As might be expected for a high-performing country, Irish pupils answered more items correctly on both item types on both assessments. The Irish advantage was most pronounced for constructed response items on PIRLS (11% more items answered correctly) and in particular for Irish boys, who answered 12% more items correctly than the PIRLS average for boys. In contrast, the gap between Irish pupils and the international average on ePIRLS was smaller, and was close to 7% on both item types, overall and when split by gender.

In line with previous research, the gender gap – in Ireland and internationally – was slightly smaller on PIRLS items in multiple choice format (1.7% and 2.6%, respectively) than in constructed response format (3.6% and 4.4% respectively). For ePIRLS, the gender gap for multiple choice items was 2.3% in Ireland and 1.8% internationally, whereas for constructed response items it was 3.2% (both in Ireland and internationally). Thus, while ePIRLS also showed a relative poorer performance for boys on constructed response items, the difference between multiple choice and constructed response was slightly less pronounced than on the paper-based assessment. However, as ePIRLS only measures Informational reading, this may be a contributory factor.

Table 11: Percentages of PIRLS items answered correctly by pupils in Ireland and internationally, by item type and by gender

		Ireland	PIRLS	Advantage
Overall	Multiple choice	77.7	68.8	8.9
	Constructed response	59.0	47.7	11.3
Girls	Multiple choice	78.6	70.2	8.4
	Constructed response	60.8	49.9	10.9
Boys	Multiple choice	76.9	67.6	9.4
	Constructed response	57.2	45.5	11.7

Table 12: Percentages of ePIRLS items answered correctly by pupils in Ireland and internationally, by item type and by gender

		Ireland	ePIRLS	Advantage
Overall	Multiple choice	72.1	65.4	6.7
	Constructed response	55.3	48.3	7.0
Girls	Multiple choice	73.2	66.3	6.9
	Constructed response	56.9	49.9	6.9
Boys	Multiple choice	70.9	64.5	6.4
	Constructed response	53.7	46.7	6.9

Item Process

As noted earlier, each PIRLS item requires pupils to use one of four comprehension processes: *focus on and retrieve explicitly stated information; make straightforward inferences; interpret and integrate ideas and information;* and, *evaluate and critique content and textual elements*. Theoretically, it is possible for an item assessing any of the processes to be very easy or very hard. In practice, Retrieve items tend to be easiest and Evaluate items tend to be most difficult. For reporting purposes, the four processes were combined into two subscales (Retrieve/Infer and Interpret/Evaluate), partly because of the relatively small number of items in each.

Bearing in mind a general pattern of lower response rates for constructed response items, and a tendency for lower-achieving pupils to answer fewer constructed response items, Table 13 looks at the interaction between PIRLS item type and item process. For items measuring Retrieve/Infer, there is a small difference in the percent correct responses by item type. In Ireland, multiple choice items tended to be answered correctly about 5% more often than constructed response items, similar to the PIRLS international average gap of 8%. In contrast, for Interpret/Evaluate items, in Ireland and across PIRLS as a whole, there is a very large difference in the percent of items answered correctly, by item type. Whereas 75% of Interpret/Evaluate items that were presented as multiple choice items were answered correctly in Ireland, this was true of only 47% of constructed response items (a gap of 28%). Internationally, the equivalent gap is 30%.

The gap can perhaps be attributed to what is often more difficult content for constructed response items that assess Interpret/Evaluate skills. However, there is also a greater tendency for pupils to skip constructed response items assessing Interpret/Evaluate processes. While skipping items was relatively infrequent amongst Irish pupils (generally high achievers), across PIRLS as a whole over 8% of constructed response items assessing Interpret/Evaluate were skipped.

Table 13: Percentages of PIRLS items answered correctly by pupils in Ireland and internationally, and percent skipped, by process and by item type

		% correct		% skipped	
		Ireland	PIRLS	Ireland	PIRLS
Retrieve/Infer	Multiple choice	78.9	70.6	0.8	1.3
	Constructed response	73.5	62.9	2.1	4.6
	Gap	5.4	7.7	1.3	3.3
Interpret/Evaluate	Multiple choice	75.0	64.8	1.0	1.8
	Constructed response	47.0	35.1	4.4	8.4
	Gap	28.1	29.8	3.4	6.6

Performance across all ePIRLS items is also higher for multiple choice than for constructed response items, for both Ireland and across ePIRLS as a whole (Table 14). However, unlike PIRLS, there is a sizeable difference of almost 12% in favour of multiple choice items in the percent of Retrieve/Infer items answered correctly, both in Ireland and internationally. This is not attributable to more constructed response items being skipped, as the rate of skipping is similar to that in PIRLS. It may be that pupils find it harder to retrieve information from across webpages than they do to retrieve it on paper, or the type of Retrieve/Infer items included in ePIRLS may be more difficult than their paper counterparts.

In contrast, while the gap by item type is more pronounced for items assessing Interpret/Evaluate skills (22% in favour of multiple choice items in Ireland and internationally), it is not as large as the equivalent gaps on PIRLS. Also, the percent of constructed response items skipped is slightly lower

than for the PIRLS equivalent item types. In sum, for the paper-based PIRLS assessment, there is little difference by item type in the percent of items answered correctly for Retrieve/Infer items, and very large differences for Interpret/Evaluate items. In contrast, there are more moderate differences by item type for both processes on ePIRLS, and slightly lower rates of skipping for constructed response items.

Table 14: Percentages of ePIRLS items answered correctly by pupils in Ireland and internationally, and percent skipped, by process and by item type

		% correct		% skipped	
		Ireland	PIRLS	Ireland	PIRLS
Retrieve/Infer	Multiple choice	75.7	70.0	0.3	0.7
	Constructed response	64.2	58.3	2.1	3.9
	Gap	11.5	11.7	1.8	3.2
Interpret/Evaluate	Multiple choice	69.4	62.1	2.1	3.8
	Constructed response	47.5	39.6	3.8	6.7
	Gap	22.0	22.5	1.7	2.8

Item Content – PIRLS

There were 36 PIRLS items on which Irish pupils did extremely well (at least 15% more Irish pupils answered correctly than the PIRLS average percent correct). Of these, 10 were from Informational texts and 26 from Literary texts. Although the 10 Informational text items were about a variety of topics, four required the reader to record at least two pieces of information in their response (e.g., complete a table showing the green sea turtle's size at various points in its life cycle, describe two explicitly stated character actions from different parts of a text). This type of item type is uncommon in PIRLS, suggesting that supplying multi-part answers may be something of a relative strength for Irish pupils.

Among the 26 Literary items on which Irish pupils demonstrated particularly high performance, just over half related to recognising or explaining a character's feelings or motivations (e.g., infer the reason for a character's feelings). Although this type of item is relatively common and included in all PIRLS Literary texts, the large number of such items on which Irish pupils demonstrated a large advantage suggests it is another relative national strength. Irish pupils also performed well on six Literary items that required integration and/or inference in relation to multiple narrative elements, to provide descriptions or explanations (e.g., identify the narrator from a range of clues, infer the significance of a character's action from subsequent events). Performance was also high on two items requiring the interpretation of imagery (e.g., infer an explanation by examining description and imagery) and two items dealing with evaluation of writing style (e.g., recognise that the author's choice of words raises suspense).

There were 11 items on which Irish performance was relatively poor (a less than 2% advantage over the PIRLS average percent correct value, or poorer performance than the international average in the case of three items). The items represent a mixture of types and themes, and there are no obvious common elements that apply to a majority of them. Five were from Informational texts and six from Literary texts, seven were multiple choice and four constructed response items. Three require the reader to interpret a character's motivation or infer their feelings – similar to the skills required in other items on which Irish pupils performed exceptionally well. Two items (both constructed response) were tagged as Above Advanced, meaning that fewer than 50% of the subset of high-performing

pupils at the Advanced International Benchmark were able to answer it correctly. However, both items were also the subject of debate during the finalisation of the coding guide by National Research Coordinators, and it may be that some of the items' difficulty can be attributed to restrictive coding instructions rather than to pupils failing to understand what was being asked.

The item on which Irish pupils' performance was weakest is a constructed response item from a Literary text. It requires pupils to explain why the protagonist's parents advise him to do something that is potentially embarrassing (doing so shows that he is honourable, and demonstrates pride in having done his best). For context, the boy and the members of his village, including his parents, have been tricked by an important figure, as a test of character. As a result, despite his best efforts, the boy had "failed" at a task. The parents nonetheless encouraged him to present his attempt to the important figure. At the end of the story, the trick was revealed, and there was an unexpected positive outcome for the boy. Irish boys, in particular, had difficulty with the item. Some opted for explanations that attributed ulterior (not always honourable) motives to the parents while others assumed that the parents were aware of the trick and that this was the reason for their advice.

There were 14 PIRLS items on which the gender gap in favour of girls among Irish pupils exceeded 10%. Most (11) were constructed response items. Three were from Informational texts and each of the three required the reader to extract information from different parts of the text and supply an answer that contained at least two facts. Irish boys were slightly less likely than girls to answer all parts of these items. The remaining 11 items were from a mixture of Literary texts, and eight of the 11 required the reader to interpret or understand a character's motivation or feelings, while one required them to describe two actions completed by a character. All 11 Literary items were from texts featuring human protagonists, with an emphasis on the resolution of an internal dilemma through the story.

As there was only one PIRLS item on which the Irish gender gap in favour of boys exceeded 10%, items where the gap was at least 5% were examined. Of the 12 items, half were from Literary texts and half from Informational texts. Five were constructed response items and seven were multiple choice. Two items assessed a character's motivation or feelings, while three required pupils to draw together information from different parts of the text to provide a single answer. Two required pupils to provide two pieces of information for a full answer. All but one item was from a text where animals featured strongly, and six items were from texts where humans were minor or non-existent elements of the text.

Item Content – ePIRLS

As ePIRLS has a smaller number of items than PIRLS and Irish performance is closer to the ePIRLS average, there are only six items on which the Irish advantage exceeds 15%. Therefore, the 10 items showing the largest gaps between Irish and international values are examined. Of the 10 on which Irish pupils did best (an at least 13% advantage), five were from the *Mars* project, and two were from *The Legend of Troy*. Four of the items needed the reader to draw on the theme of the project as a whole, or to integrate information across a number of webpages. Two examined the author's view (how content exemplified the writer's point of view and the appropriateness of the choice of title on a webpage).

Of the 10 items on which Irish performance was (relatively) poorest, four were from *Mars* and three from *The Legend of Troy*, meaning that Irish performance on these two projects varied considerably, depending on the items. Three of the items were partly based on information contained in a graphic or an image. As very few ePIRLS items required the pupil to use information from images or graphics,

this suggests a weak area for Irish pupils, perhaps because they are unfamiliar with interpreting images and graphics in an internet environment. Most of the other items on which performance was relatively poor were based on a single webpage, and typically required pupils to infer the correct response (e.g., inferring the most relevant website from a list of sites generated by a search engine).

Looking next at the 10 ePIRLS items with the largest gender gap in favour of girls among Irish pupils, seven were constructed response items, four were from *Elizabeth Blackwell* (the least popular project among Irish boys), and three from a project on Rainforests. There were no obvious common features among the items in terms of content, although six assessed pupils' ability to integrate information.

Of the 10 ePIRLS items with the largest gender gap in favour of boys among Irish pupils, six were multiple choice items, four were from *The Legend of Troy* and three from a project on *Zebra and Wildebeest Migration*. Two items related to maps – one required pupils to locate information in a map while the other required an evaluation of the usefulness of a map with interactive features. Two items related to aspects of an online environment – accessing a popup box, and inferring the most relevant website from a list generated by a search engine. However, as one of the items on which Irish girls performed very well also required inferring the best search result, it may be that the content of the search matters more than any technical understanding of the nature of an internet search engine.

Sample PIRLS texts and items

Four of the texts (and their associated items) used in PIRLS 2016 were released after the results were published, and can be accessed from the PIRLS international website (https://timssandpirls.bc.edu/). In this section, two of the four released texts (*Flowers on the Roof* and *Green Sea Turtles*) are discussed. In addition, two texts from the PIRLS Reader – a booklet containing two texts formatted in the style of a real book – with associated pupil answers, can be seen at www.erc.ie/pirls.

Flowers on the Roof and *Green Sea Turtles* are the selected texts for analysis, as the former is a Literary text and the latter an Informational one. Also, in Ireland, *Flowers on the Roof* has a slightly larger than average gender difference, while *Green Sea Turtles* is a text on which boys in Ireland perform on a par with Irish girls. *Flowers on the Roof* was part of a normal PIRLS booklet whereas *Green Sea Turtles* was part of the PIRLS Reader. Pupils assigned the Reader received a separate answer booklet in which to note their responses. Unlike the Reader, the other 15 booklets contained a text followed by related questions, or (in the case of PIRLS Literacy texts) a text interspersed with related questions. As such, the Reader might be closer to an authentic reading experience than is possible when texts and test items are together in test booklets.

The two texts provide a wide variety of the types of presentation, content, and items used in PIRLS. They also vary slightly on length. *Flowers on the Roof* (Figure 1), is just over 800 words while *Green Sea Turtles* is longer (932 words and some accompanying simple diagrams). As Irish pupils performed very well on PIRLS, performance tended to be above the PIRLS average on most items. Nonetheless, there are some items that were relatively more challenging than others for Irish pupils, and some on which Irish pupils showed quite significant strengths. The two texts contain a total of 29 items, some of which are presented after the relevant text, along with details about Irish performance on the item relative to the international PIRLS average, and information on the performance of boys and girls. Items are discussed in a subsequent section, using each item's unique ID, which identifies the source text, the item's location in the sequence of the text and whether it is a multiple-choice or constructed-response item. For example, the item ID F10C is from *Flowers on the Roof*. It is the tenth item related to that text (10) and it is a constructed response item (C).

> **NOTE: Both PIRLS and ePIRLS assess reading comprehension, not spelling. Once the intended meaning of a spelling can be ascertained, the spelling itself is not considered when deciding if the pupil demonstrated understanding in their response.**

Figure 1: *Flowers on the roof* text, formatting compressed

Flowers on the Roof

by Ingibjörg Sigurdardóttir

Shall I tell you about a granny I know? She's a really strange old lady, and so full of life! Her real name is Gunnjona, but I call her Granny Gunn. Before she moved into our apartment block she lived in the country. Her farmhouse was just like a doll's house. It had tiny little windows and the roof was covered with grass. And there were flowers growing on the roof too!

Granny Gunn had lived all on her own in the farmhouse but she was never lonely because she had many animals to play with: a cow, seven hens, two sheep and a cat.

One day Granny Gunn became ill.

"You aren't seriously ill, but you should move into town," the doctor had said. "It's not very wise to live here all alone. Your cow can't call me if you break your leg out in the yard!" "I can look after myself!" Granny Gunn answered. But then she thought that maybe it would be fun to live in town. "All right!" she said suddenly. "I'll move to town."

Soon she had sold her farm and bought an apartment in our block.

But what was she going to do about the animals? She couldn't take them to town with her, could she? Luckily, the people on the next farm kindly said that they would look after them. It was still very difficult for Granny Gunn to say goodbye to her animal friends. She was so sad that in the end she decided to take her cat, Robert, with her.

Granny Gunn packed all her things into a van and was soon on her way to her new home. She was very excited and really looking forward to seeing the town.

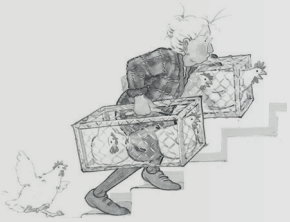
I was very excited, too! I couldn't wait to see who was going to move into the apartment opposite ours. Perhaps it would be another little boy for me to play with. But it was Granny Gunn. Still, at least she had a cat.

Granny Gunn wasn't too happy when she looked around her new apartment. "This is just dreadful!" she said. "The walls are all smooth and white. And just look at those windows! They're far too big!" She became very quiet. "I'm off back home!" she said, and turned to leave. Then she suddenly gave a little scream. Robert the cat had jumped out of the window!

"Don't worry," I said quickly. "He's only jumped out onto the balcony. Look."

Granny Gunn rushed past me onto the balcony. But when she got there, she forgot all about Robert. The balcony was huge, and she could see the mountains far away and even a bit of the sea. Granny Gunn crouched down so that she couldn't see any of the rooftops—only the mountains and the sky. Granny Gunn decided to stay after all.

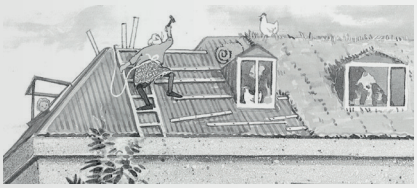
But the next day when I went around to help her unpack, she still looked very unhappy. "Are you upset because all your animals are so far away?" I asked her. "I do rather miss them," she sighed. "Then why don't you go and fetch them?" I asked. Granny Gunn winked at me and gave me a funny grin.



There was no one at home when I came to visit her the next day. Granny Gunn had taken the bus out into the country. That night I woke up to hear a strange cackling sound coming up the stairs. What could it be? Of course! The hens! They must have been too frightened to go in the lift!

The next morning, I helped Granny Gunn feed the hens. "I feel as if I'm back home," she said. "The hens are cackling all around me, and if I squint, I can easily imagine that the mountains I see are those near my farm. All that's missing is the smell of earth and grass." Suddenly she opened her eyes wide and sat up. Granny Gunn had clearly thought of something new.

"Well now," she said. "Don't you think it would be rather nice to have some grass on the roof? I think we'll have to go to town tomorrow!" And that's exactly what we did.



When we got home, Granny Gunn carried the pieces of lawn turf up onto the roof. She laid them out carefully, and fixed them so that they wouldn't fall off. Granny Gunn is much happier now. She's made a bit of countryside here in the town. She's now as fond of her rooftop garden as she had been of her old farm. And there are flowers growing on the roof once more.

Granny Gunn is not like anyone else I know. She can do anything! There's only one thing that bothers her now. How is she going to get the cow into the lift?!

Figure 2: Selected items from *Flowers on the Roof*, with item characteristics and percentages of pupils answering correctly, by gender, Ireland and PIRLS	
Item Characteristics	**Item**
ID: F10C **Process:** Straightforward Inference **Benchmark:** Low	**At the end of the story, how did Granny Gunn feel about her new home?** Guide: Answer must show that she liked new home / not homesick *She felt like she was back home / She decided that she liked it after all* **Correct:** Ireland: 95% (Irish Girls: 96%) PIRLS: 83% (Irish Boys: 94%)
ID: F06C **Process:** Straightforward Inference **Benchmark:** Intermediate	**Why did Granny Gunn scream when the cat jumped out of the window?** Guide: Answer must show she did not know there was a balcony outside of her window / was afraid for the cat *She did not know there was a balcony / She was afraid he would get hurt* **Correct:** Ireland: 79% (Irish Girls: 78%) PIRLS: 57% (Irish Boys: 81%)
ID: F05M **Process:** Straightforward Inference **Benchmark:** High	**Granny Gunn did not like the walls and windows in her new apartment. Why else was she unhappy?** a) She was ill. b) She missed her cat. c) She did not like the balcony. **d) She felt homesick.*** **Correct:** Ireland: 67% (Irish Girls: 77%) PIRLS: 67% (Irish Boys: 56%)
ID: F11M **Process:** Evaluate and Critique Content **Benchmark:** High	**The last line in the story is: 'How is she going to get the cow into the lift?!' Why does the story finish with this question?** **a) to add a joke to the story.*** b) to explain the moral of the story. c) to make the story believable. d) to help the reader understand what happened. **Correct:** Ireland: 79% (Irish Girls: 78%) PIRLS: 59% (Irish Boys: 79%)
ID: F07C **Process:** Interpret and Integrate **Benchmark:** Advanced	**When Granny Gunn was on the balcony, she crouched down so that she could not see any of the rooftops – only mountains and the sky. Why did she do this?** Guide: For full credit, answer must show complete comprehension by integrating ideas from across the text to interpret Granny's feelings about the mountains and sky. OR suggest she could actually see the countryside where she had lived. *because they reminded her of the country / She was thinking about her farm and missing it* **Full credit:** Ireland: 52% (Irish Girls: 59%) PIRLS: 36% (Irish Boys: 45%) Guide: For partial credit, describes Granny's feelings about the view but does not make a connection to her feelings about her home in the country. *So she could see the countryside / The mountains were beautiful* **Partial credit:** Ireland: 73% (21% PC only) (Irish Girls: 75% (16%% PC only) PIRLS: 51% (15% PC only) (Irish Boys: 70% (25% PC only)

Most released items from PIRLS 2016 that were categorised as being at the Low International Benchmark were not from PIRLS texts, but from PIRLS Literacy (see footnote 5). However, one item from *Flowers on the Roof* – one of the easiest texts, and a text shared with PIRLS Literacy – was

categorised as a Low International Benchmark item. As shown in Figure 2, item F10C is a fairly straightforward constructed response item, and pupils only needed to write a few words, inferring Granny Gunn's feelings at the end of the story. Almost all Irish pupils (95%) answered correctly, with no notable gender differences. The PIRLS average of 83% correct was also high, but some 12% lower than in Ireland.

Examples of correct responses range from the very short "Happy" to the longer "Granny Gunn was proud of what she did and happy with her new home", or more nuanced "She felt the same way she felt at the farm" / "She felt at home". A smaller number of pupils also contrasted her eventual happiness with her unhappiness at the start of the story. Incorrect answers were quite varied, but included examples such as saying that she still missed her animals or was homesick (thereby failing to recognise how her feelings had changed by the end of the story) or contained information that did not address her feelings ("she got her lawn on the roof") or related to another part of the story ("She did not like her new home at first").

F06C is an example of an item at the Intermediate International Benchmark, and another constructed response item tapping pupils' ability to make straightforward inferences. Pupils have to infer why Granny screamed when Robert (the cat) jumped out the window, as the reason is not explicitly stated in the text. They must draw on the response of the young boy to the scream in order to identify the reason. The answer must explain that she was not aware that there was a balcony there or (in a related vein) focus on the fact that she was afraid her cat would be hurt. For example, "She tought he fell" or "Granny Gunn screamed because she thought Robert would die".

It is an item on which Irish pupils did very well – the gap between the percent answered correctly in Ireland versus internationally is 23%. However, it is also one of the minority of items where Irish boys did slightly better than Irish girls, whereas, internationally, girls performed marginally better than boys. Incorrect answers did not grasp that Granny did not know about the balcony, or were too general to address the question ("She loved that cat" / "She screamed because she got a fright") and some assumed she thought the cat had run away or run home, not demonstrating her fear for the cat's safety.

F05M is an example of an item on which national performance only matches the PIRLS average (67% answered correctly). It is a multiple choice item at the High International Benchmark that asks pupils to infer one reason why Granny was unhappy. In Ireland, there is a large gender gap (21%) in favour of girls, much larger than the international gender gap of 7%. Irish boys also performed poorly on the item relative to boys internationally. Irish boys were twice as likely as Irish girls to choose option C (She did not like the balcony), focussing more on issues to do with the physical environment than on Granny's emotional state. To identify the correct response, pupils have to connect her complaints about the new apartment with the next lines ("She became very quiet. 'I'm off back home!' she said") or with the comment later about her missing her animals, and infer that she is homesick.

Item F11M assesses pupil's ability to evaluate and critique text content, and is an example of an item at the High International Benchmark. Pupils need to be able to recognise that the author's intent is to be humorous and to end the story on a light-hearted note. The item requires pupils to draw on their understanding of use of language. It marks a shift from constructing meaning from the text, as some easier items required, to reflecting on "the author's language choices and devices for conveying meaning" (Mullis, Martin & Sainsbury, 2015, p.23). Again, Irish pupils performed well above the PIRLS average on this item, with 79% answering correctly, compared to an international average of 59%. Irish boys and girls performed equally well on this item, and the gender differences for PIRLS as a whole were also quite small (2%).

Item F07C is at the Advanced International Benchmark, and assesses pupils' ability to interpret

and integrate what they read. It is also an example of a full/partial credit item, whereby pupils who provide a full answer get two points, and those whose answer is partially correct receive one point. Constructed response items assessing this comprehension process can produce quite varied responses, and potentially, multiple acceptable responses because they often require the reader to draw on their own experience and perspectives. For F07C, which asks why Granny crouched down, a number of explanations are accepted. For example, some pupils indicated that she was reminded of her home in the country when she saw the mountains and the sky, while others felt she could pretend she was back in the country if she didn't see the buildings ("So then she would think that she was in her farm"), and others responded that doing so let her see the countryside where she had lived ("she could see her farm animals and where she youst to live").

Partial credit was given to answers that explained that crouching meant she could only see countryside, but without the important reference to her feelings about her farm. For example, "She done it to get a better view of the mountains and a bit of the see" or "Because she did not want to see rooftops". Examples of incorrect answers included her not liking the view, or needing to sit down because she was old. Irish pupils again performed above the PIRLS average on this item (52% received a full credit versus 36%, respectively), and it was an item on which Irish girls in particular performed quite well (59% full credit, versus 45% for Irish boys). A total of 73% of Irish pupils received a partial credit (PIRLS: 51%), including 70% of boys.

Figure 3: *Green Sea Turtles* text, formatting compressed

The Green Sea Turtle's Journey of a Lifetime

From Turtle Travels By Gary Miller

Out From the Sand

It's a starry night in August. A nest of eggs lies buried in the sand more than half a metre below the surface of a Costa Rican beach. The nest holds more than 100 green sea turtle eggs, each about the size of a golf ball. One of the baby sea turtles begins to stir and hatch from her egg. The hatchling tears at the shell of her egg with the sharp point on her beak. Still buried beneath the sand, the baby sea turtle breaks free. Soon, the whole nest is alive with motion. The baby turtle uses her flippers to climb up and up. It can take more than a day to reach the surface of the sand.

Into the Water

When the hatchling reaches the surface of the sand, she is drawn to the moonlight reflecting off the ocean. Luckily, there are no lights shining from a nearby street or house. These lights can confuse a baby turtle. They can make it go the wrong way, away from the sea.

The hatchling's journey to the water is a race for survival. She is no bigger than a walnut. Crabs and birds, such as night herons, snatch up some of the other baby turtles on the beach. This baby turtle makes it to the water. The frothy surf pushes the baby turtle back. She fights to swim against the breaking waves. The hatchling continues to swim through the first day and night, and she does not slow down for two days.

Out to the Open Sea

The baby turtle's journey through the open sea is often called the "lost years." Scientists know little about this phase of a green sea turtle's life. She may move with the currents, floating with mats of seaweed. The hatchling may snack on prawns, small jellyfish, and snails that drift in and around the seaweed. Unfortunately, the sea also contains plastic and rubbish that people throw away. Eating these could be deadly for the turtle.

The sea has many other dangers, too. Predators such as sharks swim below the small turtle and large birds fly above. Fortunately, she gets some protection from the colouring of her shell. The bottom is almost white, so sharks swimming below may not spot her in the sunlight. The top of her shell is dark, so from above the turtle blends into the dark water.

Growing Up Green

After several years, she has become a juvenile. She is no longer a hatchling, but she is not yet an adult. Her shell has become about the size of a dinner plate. It is now time to leave the open sea for the warm coastal waters of Florida, in the United States. With her larger shell she is safer than she was as a hatchling. Although she sometimes slurps up a jellyfish, now she mostly eats algae and sea grass.

Back to Contents

Years pass as she slowly grows. She moves farther off shore to feeding grounds where she becomes an adult. At night, she rests in the water under rocks and ledges, holding her breath for up to five hours. Each day, she returns to the same patch of sea grass called turtle grass. Like a lawnmower, the turtle keeps this sea grass pasture cut short. Eating sea grass and algae turns her body fat a green colour. In fact, this is how green sea turtles get their name!

Back to the Sand

When the turtle is about 26 years old, her adult shell is nearly a metre long and she weighs around 140 kilogrammes. Now she sets out on a new adventure. She begins her long trek back to the beach where she was born. She is going to lay her own eggs.

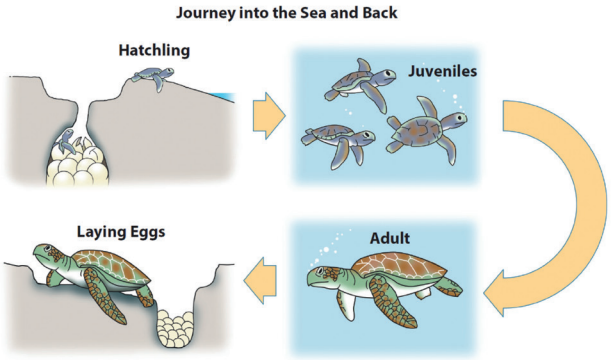
The sea turtle may have to travel over 1000 kilometres, but she is well equipped for the journey. Her flippers are like wings. She flies through the water. Scientists are still learning how a sea turtle can find its way through the ocean. They think the turtles may sense changes in Earth's magnetic field. That may help the turtles create a kind of mental map. Their memory of chemicals or smells in the water may also help them find their way. Once she returns to her birthplace, she finds a mate. A few weeks later, she waits until it is dark, and then climbs onto the beach.

The Next Generation

Out of the water, she struggles to move on land. She crawls to a place where high tides will not wash away her eggs. Using her front flippers, she digs a wide pit. This will become her nest. With her rear flippers, she scoops out a smaller hole inside the pit. After two hours of hard work, she is ready to lay more than 100 leathery white eggs inside the smaller, deeper hole. She packs sand over them. Then she tosses sand over the whole nest. During the following two months, she will dig and lay eggs in three more nests. After two months, the new hatchlings break out of their shells to begin their own journeys.

Turtles Live On

After laying all of her eggs, this adult sea turtle once again sets out for her feeding grounds off the coast of Florida. Every few years she and other adult turtles will return to this beach to lay more eggs. Every green sea turtle does this throughout its entire life, which could last up to 80 years. Over this time, thousands of baby green sea turtles will be born and set out into the open sea.

Figure 4: Selected items from *Sea Green Turtles*, with item characteristics and percentages of pupils answering correctly, by gender, Ireland and PIRLS	
Item Characteristics	**Item**
ID: T01M **Process:** Straightforward inference **Benchmark:** Intermediate	**What is the first section "Out From the Sand" about?** a) what different sea turtles look like b) how sea turtles learn to swim c) what sea turtles like to eat **d) how sea turtles' eggs hatch*** **Correct:** Ireland: 92% PIRLS: 84% Irish Girls: 92% Irish Boys: 93%
ID: T05M **Process:** Retrieve Explicitly Stated Information **Benchmark:** High	**What is the first thing the hatchling does when she finally gets past the breaking waves?** a) searches for the other hatchlings **b) keeps swimming far out to sea*** c) rests in the seaweed d) finds food to eat **Correct:** Ireland: 57% PIRLS: 57% Irish Girls: 58% Irish Boys: 57%
ID: T14C **Process:** Evaluate and critique **Benchmark:** High	**A diagram from the article is shown below. What does this diagram help you to understand?**  Guide: Answer must state that it indicates the life cycle or stages in a turtle's life *The circle of life for turtles / helps us know how they grow up / different stages in its life* **Correct:** Ireland: 54% PIRLS: 47% Irish Girls: 53% Irish Boys: 54%
ID: T13M **Process:** Straightforward inference **Benchmark:** Advanced	**Which activity in an adult female green sea turtle's life is not fully understood by scientists?** a) how she can swim over 1000 kilometres b) how she makes a nest for her eggs c) how she avoids being eaten by predators **d) how she finds the right beach to lay her eggs*** **Correct:** Ireland: 52% PIRLS: 45% Irish Girls: 48% Irish Boys: 56%

ID: T07C **Process:** Interpret and integrate **Benchmark:** Above advanced	**The colour of a hatchling's shell protects it from predators. Give a way it is protected from birds AND Give a way it is protected from sharks.** Guide: answer recognizes that the dark shell colour of the top conceals the turtle by blending in with the dark water AND that light shell colour of the bottom conceals the turtle by blending in with the sunlight *Birds: It camouflages them / It blends in with the sea because its shell is black* *Sharks: It is the same colour as the sunlight / The shell blends into the water* **Full credit:** Ireland: 26% PIRLS: 25% Irish Girls: 29% Irish Boys: 24% Guide: recognises one of the two ways the shell protects **Partial credit:** Ireland: 50% (24% PC only) PIRLS: 44% (19% PC only) Irish Girls: 53% (25% PC only) Irish Boys: 47% (23% PC only)

While *Flowers on the Roof* was a relatively easy text, *Green Sea Turtles* is a good deal more difficult. No items in the text were associated with the Low International Benchmark level, and only two of the 16 related items were at the Intermediate International Benchmark. Item T01M is an example of an Intermediate level item, on which Irish pupils perform slightly better than the international average (92% of Irish pupils answered correctly, versus 84% of pupils, internationally). Internationally, girls had a 4% advantage over boys on the item, whereas in Ireland, boys had a 1% advantage over girls. The item is multiple choice and requires a very straightforward inference that the section "Out From the Sand" is about the turtle eggs hatching, information that is clearly contained in that part of the text.

Item T05M is another multiple choice item, but at the High International Benchmark. There is no difference between the performance of Irish pupils and the PIRLS international average on this item (57% answered correctly in each case), nor are there any gender differences in performance among Irish pupils. Across PIRLS as a whole, girls are marginally more likely than boys (by 2%) to identify the correct answer as option B (the hatchling keeps swimming out to sea). The item requires pupils to retrieve explicitly stated information, although the phrasing in the text is not an exact match to the question asked and correct response option ("She fights to swim against the breaking waves. The hatchling continues to swim though the first day and night, and she does not slow down for two days"). About one quarter of pupils, in Ireland and internationally, chose option D for this item (finds food to eat), perhaps drawing on their opinion of what they might expect the hatchling to do rather than what is stated in the text.

Item T14C is a constructed response item that assesses pupils' ability to evaluate and critique content. It is at the High International Benchmark, and only 54% of Irish pupils (slightly higher than the PIRLS average of 47%) correctly indicated that the diagram helped to understand the life cycle or stages in a turtle's life. Gender differences on this item were minimal in Ireland (a 1% advantage for boys), while there was a 3% advantage for girls across PIRLS as a whole. Examples of correct answers provided by Irish pupils include "How sea green turtles grow up", "the circl of life" and "It helps you understand their lifecircle". Incorrect answers tended to focus on one or two elements of the diagram only ("How they lay eggs" or "A turtle laying eggs and baby turtles"), or to ignore the cyclical nature of the diagram "How a turtle is born and what type of turtle it is". A small number of (usually) lower achieving pupils copied the diagram title as their answer (e.g., "It shows the journey into the sea and back"), which, while true, does not provide any additional information.

Item T13M assesses pupil's ability to make straightforward inferences. It is a relatively difficult (Advanced International Benchmark) multiple choice item answered correctly by 52% of pupils in Ireland and by 45%, internationally. Irish boys outperform Irish girls on this item by 8%, making it one of a small number of items (12 of 181) where Irish boys outperformed their female counterparts by

5% or more. In contrast, across PIRLS as a whole, there were no gender differences in performance on this item. Pupils are asked which part of the female turtle's life is not fully understood by scientists, the correct answer being how she manages to find her birthplace beach. The phrasing in the text is a little less precise ("how a sea turtle can find its way through the ocean") and the reader must infer that this relates to finding the way to the specific beach. Over one-third (35%) of Irish girls selected the incorrect option A (how she can swim over 1000 km).

Item T07C is a particularly difficult constructed response item that assesses pupils' ability to interpret and integrate text. Scoring was applied using either a full credit (2 points) or partial credit (1 point) model, whereby both parts of the question had to be answered correctly to get full credit. The item is tagged as Above Advanced International Benchmark, meaning that fewer than 50% of the subset of high-performing pupils at the Advanced International Benchmark were able to answer it correctly. Overall, only 26% of pupils in Ireland and 25% internationally obtained a full credit on this item, whereas 50% of Irish pupils received a partial credit, as did 44% of pupils across PIRLS as a whole. In Ireland, girls did slightly better on this item than boys (29% received a full credit, versus 24%), broadly in line with gender differences overall on this item.

Pupils are asked to identify how the turtle's shell protects from predators above and below. The answer is obtained from a few consecutive lines of text "*Predators such as sharks swim below the small turtle and large birds fly above. Fortunately, she gets some protection from the colouring of her shell. The bottom is almost white, so sharks swimming below may not spot her in the sunlight. The top of her shell is dark, so from above the turtle blends into the dark water.*" As such, there is no need to integrate information from different parts of the text, or to step back and critique the text. Nonetheless, only a minority managed to identify the two methods of protection.

This is also an example of an item that was often difficult to score, because there were a number of instances where pupils showed some understanding of the issue, but did not quite explain their answer in sufficient detail to meet the criteria as set out in the coding guide (a shared guide, used by all scorers in all countries when assigning scores). Frequently, the issue was either that the pupil did not demonstrate understanding of how the shell camouflaged the turtle, or may have understood the nature of the camouflage but did not explicitly reference it. Examples of correct responses by Irish pupils as to how the shell protected from birds include "All of her shell on the top is dark so above she blends into the dark water" and "the top is dark so from above it blends in to the dark water". Examples of correct responses to how the turtle is protected from sharks include "the shark wont be able to see because bottem of the shell will blend in with the sun light" and "Sharks cant see the turtles because they blend in with the sun and water".

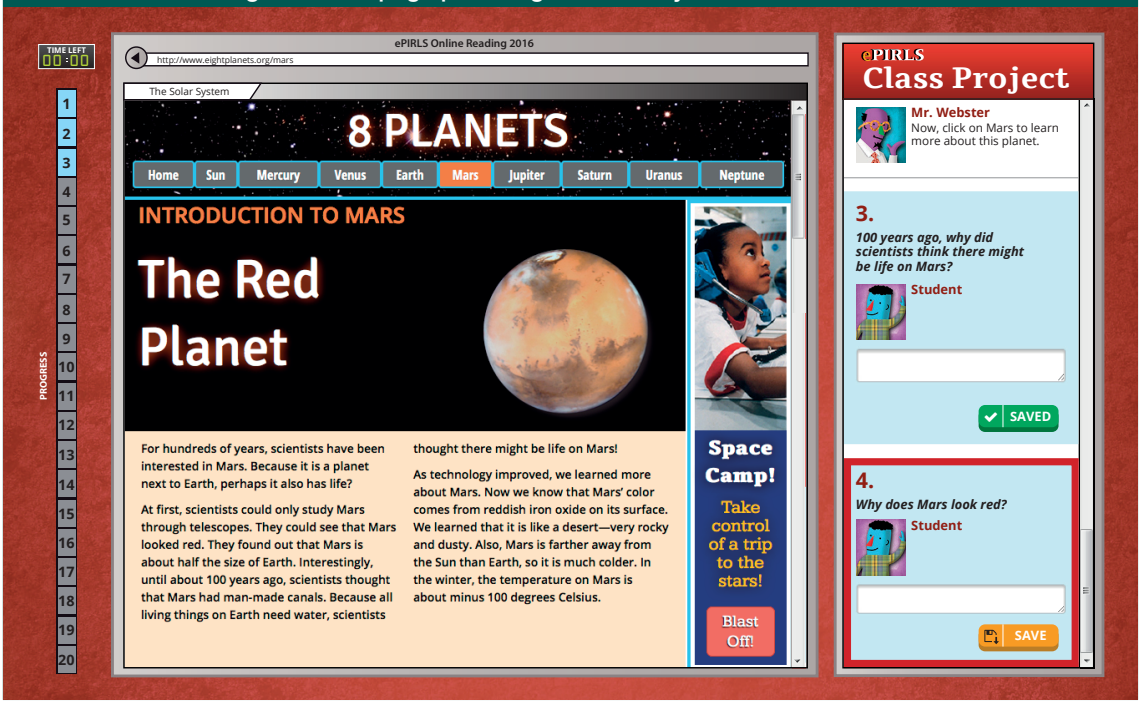
Incorrect answers included a mixture of facts without explanation, incorrect information, and (in the case of the second part) a sizeable number of missing or rather random responses. Examples of factually correct, but non-credited answers include "protects them from birds so they can't eat them" (true and a common response, but does not explain how the shell protects) and "The bottem of the turtles shell is white" (true and another common response, but fails to include any reference to the camouflage aspect). Incorrect answers include "She is white so sharks from above may not spot her" (sharks are below, while birds are above), "birds can't get at them because they are too deep" (not stated anywhere in the text), "She will curl up into it" and "sharks will think there a rock" (both of which seem to draw on best guesses rather than to what is in the text).

Sample ePIRLS project and items

Two ePIRLS projects (*Mars* and *Elizabeth Blackwell*) and their associated items were released after the results were published. They can be accessed from the PIRLS international website (https://timssandpirls.bc.edu/) as live, interactive tests. In this section, one released project (Mars) is discussed, and some items and screenshots are used to illustrate the nature of the project.

As noted earlier, ePIRLS projects are designed to assess Informational reading skills and pupils' ability to use the internet for schoolwork. Content is presented as interactive online class projects, guided by a teacher avatar named Mr or Ms Webster. The teacher communicates through an online chat dialogue visible as a sidebar next to a simulated internet window. A sequence of simulated webpages containing texts, graphics, links, pop-up windows and animation appear in the main window alongside the sidebar. Pupils are prompted to select or enter responses in the dialogue area at appropriate times. Webster also navigates pupils through the exercise by indicating which links to click in the main window to proceed through the assignment. In the event of pupils selecting an incorrect link, the correct page is loaded, meaning that pupils can still continue with the assessment. A progress bar is visible throughout the assignment on the left of the screen. Figure 5 provides an example of a screen as seen by pupils.

Figure 5: Webpage providing introductory information on Mars



Mars was one of two ePIRLS projects that was slightly more popular with Irish boys than girls (93% of boys versus 89% of girls liked *Mars* a little or a lot). The purpose of the Mars project is to learn about scientists' efforts to explore the planet Mars. Reading content is presented in a number of different webpages. There are two Google search results pages in which pupils select the most appropriate search result for a given aim. Six Informational webpages are also presented, from two

different sites. These pages vary in their presentation, with text accompanied by illustrative graphics ranging from still images to animated and interactive graphics and pop-up boxes. The pages also contain simulated content-related advertisements, serving as distractors and adding authenticity. The first two webpages are from a site about the solar system, containing information about the solar system, generally, and some facts that are specific to the planet Mars. The remaining four pages are from a site about Mars exploration, containing detailed information on Mars exploration, the challenges in locating Mars, an overview of four types of technologies used during missions, and a description of a Mars "Rover" called Curiosity. The final webpage of the Mars project is a more subjective online news article about Curiosity, the Mars rover.

Some selected *Mars* content relevant to the items to be discussed in this section is presented in Figures 5 to 8. Figure 5 shows a screenshot of the entire screen as it appears to pupils. Figures 6 to 8 display main reading content only.

> **NOTE: The screenshots show the international version of the Mars project. Some spelling and phrasing adaptations were made for the Irish version, but are not shown here.**

Figure 6: Webpage describing challenges locating Mars for space missions with animated graphic of the Earth and Mars orbiting the sun on bottom right

Figure 7: Webpage describing three types of missions to Mars

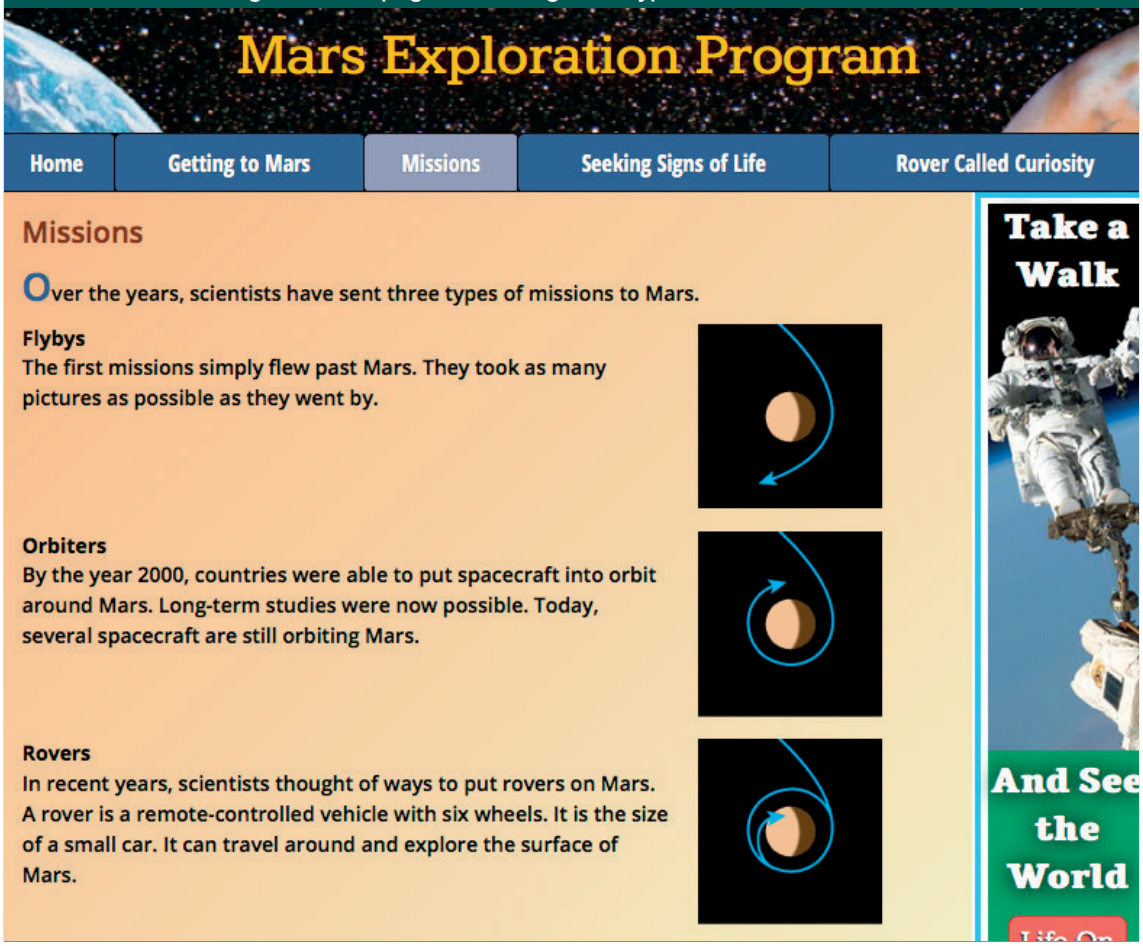
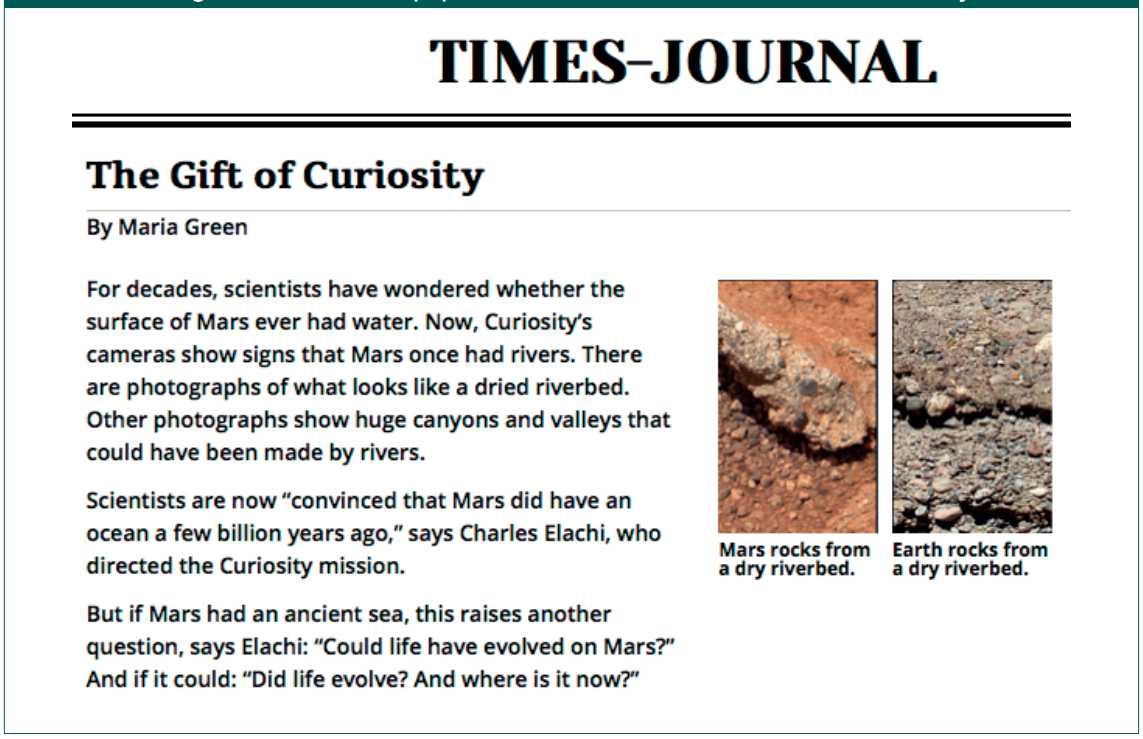


Figure 8: Online Newspaper Article about the Mars rover called Curiosity



Ireland's performance on most *Mars* items was higher than the ePIRLS average, mirroring Irish pupils' performance above on ePIRLS generally. However, Irish performance on some items was atypical – ranging from at or below the international average to much higher than the norm, or demonstrating unusual patterns of gendered performance. A selection of *Mars* items is presented in Figure 9, chosen to demonstrate a variety of item processes across the different Benchmark levels along with varying performance by Irish pupils. Figure 9 cross-references the corresponding content relevant to each item in Figures 5 to 8.

Item M05M is an example of one of three items at the Low International Benchmark within the Mars project. It is a Retrieve item, requiring pupils to locate and recognise an explicitly stated reason why Mars is colder than the Earth. The answer is found in the webpage shown in Figure 5, in a single, discrete sentence ("Also, Mars is farther away from the Sun than Earth, so it is much colder"). Performance on this multiple choice item was very high both in Ireland and internationally (90% and 87% respectively), suggesting that this was a relatively easy item for most pupils. In Ireland, there was a gender gap of 3% in favour of girls, compared to a 1% gap, internationally.

M03C is an example of an item at the Intermediate International Benchmark and is a constructed response item. Pupils must infer from the text why scientists 100 years ago thought there might be life on Mars. Again, the source of the possible answers is in Figure 5, which refers to what looked like canals on Mars (meaning a potential water supply) and to Mars being the planet next to Earth, with a potentially similar capacity to support life. Irish pupils performed particularly well on this item (85% correct responses versus 66%, internationally). There was a gender gap of 4% in favour of Irish girls compared to a gap of less than 1% internationally. Examples of correct responses provided by Irish pupils included "they thought that someone had built canals on mars" and "Because they think that Aliens might have built canals" (both answers refer to the potential water source) and "they thought there could be life because it was so close to earth" (referring to the similar capacity to support life).

A very common wrong answer supplied by Irish pupils was "because all living things on earth need water". It is likely that this exact wording was common because on the webpage it directly precedes the words "scientists thought there might be life on Mars!". Other relatively common incorrect responses included "they thought that they had seen someone building canals on mars" (untrue) and "They had worse technology" / "They thought because they had no gagets back then" (both of which indirectly refer to the piece in the text stating that scientists then relied on telescopes alone, which, while true, does not answer the question).

M10M is an example of an item at the High International Benchmark that requires the skills of evaluating and critiquing text and visual content of a webpage (Figure 6). Pupils are asked to determine the purpose of the animated diagram depicting the Earth and Mars orbiting around the Sun. The diagram shows the planets moving around the sun within their differently-sized orbits, showing that the distance between the planets varies, depending on their position at different points in time. The correct option ("to show that the distance between the Earth and Mars changes") was chosen by 78% of Irish pupils, slightly higher than the ePIRLS average of 72%. Gender differences were minimal (1% in favour of boys), both in Ireland and across ePIRLS as a whole. The most common incorrect response, chosen by 12% of Irish girls and 10% of boys, was "to show that the Earth and Mars are far from the Sun", which also relates to distance, albeit to the Sun rather than between the planets.

Figure 9: Selected items from *Mars*, with item characteristics and percentages of pupils answering correctly, by gender, Ireland and ePIRLS	
Item characteristics	**Item**
ID: M05M **Process:** Retrieve **Benchmark:** Low **Related Content:** Figure 5	**Why is Mars colder than the Earth?** **a) It is farther from the Sun.*** b) It is smaller than the Earth. c) It is very rocky. d) It has canals. **Correct:** Ireland: 90% ePIRLS: 87% Irish Girls: 91% Irish Boys: 88%
ID: M03C **Process:** Straightforward Inference **Benchmark:** Intermediate **Related Content:** Figure 5	**100 years ago, why did scientists think there might be life on Mars?** Guide: Answer must refer to man-made canals / water on Mars / that the presence of water might indicate life / or Mars being close to the Earth. *Because scientists thought that someone built canals on mars.* *Because it is next to the Earth so it could have life on it.* *They thought Mars had water.* **Correct:** Ireland: 85% ePIRLS: 66% Irish Girls: 87% Irish Boys: 83%
ID: M10M **Process:** Evaluate and critique **Benchmark:** High **Related Content:** Figure 6	**Look at the moving diagram of the Earth and Mars orbiting around the Sun. What is the main purpose of the diagram?** a) to show the colours of the Earth and Mars. b) to show that the Sun is part of the solar system. **c) to show that the distance between the Earth and Mars changes. *** d) to show that the Earth and Mars are far from the Sun. **Correct:** Ireland: 78% ePIRLS: 72% Irish Girls: 78% Irish Boys: 79%
ID: M13C **Process:** Interpret/Integrate **Benchmark:** Advanced **Related Content:** Figure 7	**Now that there are rovers on Mars, why are orbiters still useful?** Guide: Answer must indicate either that orbiters can remain in orbit for a long time, or have the capability of seeing different things about the planet. *For long-term studies.* *The orbiters can take photos and tell if the rover gets to Mars.* *They can take pictures from space. / They have a more complete view of the planet.* *They can go around the planet faster and they can possibly find different things compared to rovers.* **Correct:** Ireland: 27% ePIRLS: 27% Irish Girls: 24% Irish Boys: 29%
ID: M20C **Process:** Evaluate and critique **Benchmark:** Advanced Related **Content:** Figure 8	**The writer of "The Gift of Curiosity" is in favour of exploring Mars. Explain how the writer shows this.** Guide: Answer must refer to the title, or gaining knowledge, or the writer's interest or enthusiasm on the subject. (Title) *She calls it "the gift."* (Gaining knowledge) *Yes because now we know more about Mars and what's on it.* (Interest) *At the end she said a lot of questions so I think she wants to know more/ Because by the language of her article he seems very interested* (Enthusiasm) *I think it because it sounds like she likes Curiosity/ Because Curiosity gave so many things/ Because she wrote about it* **Correct:** Ireland: 48% ePIRLS: 25% Irish Girls: 50% Irish Boys: 46%

ID: M11C **Process:** Interpret/ Integrate **Benchmark:** Full Credit: Above advanced Partial Credit: Intermediate **Related Content:** Figure 6	**You have to plan a long time ahead to get to Mars. Explain why.** Guide: For full credit, answer must show that it is difficult to know where Mars will be by explaining either of the following: Mars is moving and the rocket takes a long time to get to Mars/Mars is far away *or* Mars and Earth have different orbits *Mars is moving around the Sun you need to know where Mars will be when the rocket reaches it.* *You have to aim your rocket for where Mars will be in 8 months.* *Mars and the Earth have different orbits.* *The two planets move farther apart and then come closer together.* **Correct:** Ireland: 18% ePIRLS: 20% Irish Girls: 22% Irish Boys: 15% Guide: For partial credit, response must provide one part of the reason for planning ahead. *Mars is orbiting/moving.* *It takes a long time to get there.* *Calculating where Mars will be is difficult.* **Partial credit:** Ireland: 60% (42% PC only) ePIRLS: 57% (37% PC only) Irish Girls: 58% (36% PC only) Irish Boys: 63% (48% PC only)
ID: M14C **Process:** Interpret/Integrate **Benchmark:** Full Credit: Above advanced **Related Content:** Figures 5 and 7	**New inventions have helped scientists look at Mars from locations closer and closer to the planet. In the boxes below each location, write the name of the invention that scientists used to look at Mars. You can look back through the webpages you have read.** **Looking at Mars from the Earth:** _____ **Looking at Mars from space:** _____ **Looking at Mars from its surface:** _____ Guide: For full credit, answer must show complete comprehension by providing all three responses shown below: Looking at Mars from the Earth: *Telescope(s)* Looking at Mars from space: *Flyby(s) (flyby missions), orbiter(s)/satellite(s), or spacecraft/rocket(s)* Looking at Mars from the surface of the planet: *Rover(s)* **Fully correct (3 points):** Ireland: 37% PIRLS: 23% Irish Girls: 30% Irish Boys: 45% Guide: For partial credit (2/3 inventions correct), response shows satisfactory comprehension by providing two correct inventions **Partial credit (2 points):** Ireland: 61% (24% PC only) PIRLS: 47% (24% PC only) Irish Girls: 55% (25% PC only) Irish Boys: 68% (23% PC only) Guide: For partial credit (1/3 inventions correct), response shows minimal comprehension by providing one correct invention **Partial credit (1 point):** Ireland: 75% (14% PC only) PIRLS: 60% (14% PC only) Irish Girls: 70% (15% PC only) Irish Boys: 80% (13% PC only)

M13C is a constructed response item at the Advanced International Benchmark that requires pupils to integrate information from a web page (Figure 7) to explain why orbiters are still useful, now that there are also rovers on Mars. Answers need to highlight the different capabilities of orbiters compared to rovers by indicating that orbiters can remain in orbit for a long time, or that they can see different things about the planet. Ireland's overall rate of 27% correct answers was identical to that

of the ePIRLS average, suggesting that pupils in general found this a difficult item. Internationally, the gender gap was only 1% (in favour of boys). However, Irish boys outperformed girls by a 5% gap, and Irish girls' performance was 2% below the ePIRLS girls' average, indicating a relative weakness for Irish girls.

Correct answers by Irish pupils commonly referred to orbiters' capacity for long-term studies, often quoting directly from the text "long term studies were now possible". Another very common correct response referred to the different views captured by the two types of spacecraft (e.g., "They are useful because you can get a view of mars from a different perspective then on the land" and "to have a birds eye vew?"). Some interpreted the question as determining how orbiters might help support rovers, and inferred more indirectly from the text that orbits can help to keep track of rovers e.g. "you can use orbiters to check if the rover is okay" and "So they can track the rovers from mission control to see if they are active". Another correct inference was in relation to the speed with which orbiters can explore Mars in comparison to rovers e.g. "To explore faster than rovers" and "you can explore alot of mars quicker".

Incorrect responses often focused on the capabilities of orbiters without comparison to rovers (e.g., "to take pictures" and "Because they can still orbit around mars"). A common wrong answer was to repeat all or part of either of the following two sentences from the text: "By the year 2000, countries were able to put spacecraft into orbit around Mars" or "Today, several spacecraft are still orbiting Mars". Pupils also commonly made inadequate attempts to convey the vantage point of orbiters compared to rovers (e.g., "they are till useful because they are still floating around mars and can collect more information", "beacause they are higher than rovers", "orbiters are still used because they are around mars" and "to see far away"). Others correctly inferred that orbiters might act as an aid to rovers but misjudged how, with some pupils describing them as a means of transport (e.g., "they can bring more rovers and fix the broken ones" and "you could put spacecraft into them"). Another type of incorrect response was to describe the purpose of orbiters as related to orbits "cause it can show the orbit" or "to get to mars you have to caculate the mars orbit".

M20C is another item at the Advanced International Benchmark. It relates to the content of the online news article "The Gift of Curiosity" (Figure 8). Pupils must evaluate and critique the content and textual elements in order to explain how the writer shows that she is in favour of exploring Mars. Answers must refer to the title of the article, gaining knowledge, or the writer's interest or enthusiasm on the subject. Ireland did particularly well on this item relative to other countries, with 48% of Irish pupils compared to only 25% internationally, answering correctly. The 23% advantage for Ireland on this item was the largest advantage across all ePIRLS items. The gender gap between Irish boys and girls was 4% in favour of girls, similar to the average ePIRLS 3% advantage for girls on the item.

Most pupils in Ireland who answered correctly referred to the writer's interest, enthusiasm, or tone of the article as evidence that she was in favour of exploration. For example, "The writer shows she is in favour of exploring Mars by being so enthuasiastic", "the writer shows this by being interested in mars", "She said it was a gift to explore and she sounded so happy", and "he[7] loves curiosity",. Another common type of response was to infer that the writer had a favourable attitude because she provided lots of information ("by telling us stuff about Mars", "she has loads of facts and explains them well"). A less common response was to refer specifically to the phrasing of the title "because he calls curiosity a gift".

7 It was relatively common for pupils to refer to the writer as male, even though the author is listed as "Maria Green". Similarly, in the Pemba Sherpa PIRLS text, the main protagonist (who is female and trying to enter a male-dominated job) was frequently referred to as male.

Common incorrect responses from Irish pupils referred to the images used in the article, with some believing the writer to have taken the picture of Mars herself (e.g., "he shows this becuase he photographs that dont look like be got them from the internet"), or misinterpreting the question as referring to the writer wishing to explore Mars in person ("I think she wants to explore mars because she holds an interest"). A small number believed the writer to have already been on Mars, hence her enthusiasm ("The writer shows this because he directed the Curiosity mission", "the wrighter of it sounds like he was on mars before because he said a few billion years ago there was a ocean"). Other types of incorrect responses included just mentioning facts from the article (e.g., "Scientists are convinced that mars did have an ocean a few billion years ago") or vaguer descriptions of the article content (e.g., "he tells people that there could be life on mars and is telling us to think about it did life evolve? and where is it now?").

Item M11C is benchmarked at *above* Advanced level. Answering correctly requires integrating information in the text with an animated graphic to explain why it is necessary to plan a long time ahead to get to Mars. This was another item in which Ireland performed below the international average (18% of pupils provided full credit responses versus 20% internationally) with a gender gap of 7% in favour of girls in Ireland. Irish boys lagged slightly behind the PIRLS average (3%) for boys.

For full credit, pupils must say that Mars is moving and is far away, or that Mars and Earth have different orbits. Examples of full credit responses from Irish pupils include "because you need to calculate Marses orbit then it will take about eight months to get there" and "Because they have different orbits". Partial credit was awarded for responses providing one part of the reason for planning ahead, and such responses were considered to be at the Intermediate International Benchmark. Irish boys were more likely to get a partial credit than were Irish girls (48% versus 36%). Common partially correct answers reproduced one piece of relevant information from the text, focusing on either the length of time taken to get to Mars (e.g., "Because it will take u eight months to get to mars so you need a plan eight months ahead of you") or referred to locating Mars within its orbit but not relative to the Earth's orbit (e.g., "Because you need to calculate mars orbit" or "Because somtimes it is closer and somtimes farther away").

Some pupils attempted to convey the difficulty in locating Mars but with insufficient clarity ("because earth and mars changes when going around the sun", "because they move differantly") to merit a partial credit. Many other incorrect responses referred to trip practicalities (e.g., "you can be going for a long time so you need food and a drink", "first you need a powerful rocket" and "because well you need a roccet and some ocsagen and a helmet and some space clothes"). Others referred to the importance of planning, especially for a potentially dangerous trip (e.g., "You have to plan a long time ahead before you go to Mars because the people risking their lives need to train and be fit. And the scientists must have to set up camaras to see if there is any scine of life." and "So you do not die while recklessly try to go to mars without a good plan".

Item M14C is an item that can be credited as 3, 2 or 1 point. Full credit responses fall above the Advanced International Benchmark, meaning it is a very difficult item, even for high-achieving pupils. It is a constructed response item that assesses pupils' ability to interpret and integrate information from two web pages, naming three inventions used for looking at Mars. Although Irish pupils in general performed above the ePIRLS average for this item (37% versus 23% receiving full credit, respectively), Irish boys appeared to do particularly well on it, with 45% achieving full credit, compared to 30% of Irish girls. Their advantage on M14C was the largest of any ePIRLS item, and they also outperformed the ePIRLs boys' average by 18%. The very good performance of Irish boys on this item, and on Mars, generally, may relate to existing knowledge of space exploration, to interest in the topic (as they rated the project very positively), or a combination of both.

Supplying all three responses requires integrating information from two different web pages. Answers to the last two responses were on the web page beside which the item is located. However, the answer to the first part of the item was on another page accessed by returning to the initial website about the solar system via an open tab link and then clicking on the Mars tab within that website. Remembering the location of the information about the first part of the item and finding it again proved difficult for many pupils. The answer (*telescopes* are used to look at Mars from Earth) was not supplied by well over one-third of pupils, who either made no response or provided an incorrect response. Also, it may be the case that some who answered this part correctly did so by drawing on their external knowledge rather than accessing the information from the webpage. In contrast, most pupils answered the second two parts of the item, identifying flybys, orbiters or spacecraft as ways to view Mars from space, and rovers as ways to view Mars from its surface. In Ireland, 61% answered at least two parts of the item correctly (PIRLS average: 47%), with boys again outperforming girls by a large margin (68% versus 55%, respectively).

A small number of pupils reversed the order of the responses, meaning they received no credit for their answers, but a majority answered at least the last part (rovers look at Mars from its surface) correctly. A number of pupils mentioned cameras and pictures in response to looking at Mars from space (e.g., "they took pictures as they flew by", "the first mission simply flew past mars and took as many pictures as possible"). Such answers are not credited as, while it is true that the flybys were used to take pictures, a correct response must indicate the vehicle hosting the camera. Another type of incorrect response was when pupils misinterpreted the item as asking for the view/experience/size of Mars from the different locations (e.g., "(1) mars looks reddish from the earth, (2) mars is very rocky from space, (3) mars is cold when your on it" and "(1) a small dot, (2) big, (3) huge").

Summary and conclusions

In PIRLS 2016, pupils in Fourth class in Ireland performed at a very high level. Only two countries (Russian Federation and Singapore) achieved statistically significantly higher mean scores, four countries obtained means that did not differ significantly from Ireland, while all remaining 43 countries were outperformed by Ireland. Performance was equally high on ePIRLS.

Frameworks for both curriculum and assessment in Ireland are broadly aligned with the PIRLS assessment framework. This alignment is apparent in the definition of reading literacy used in PIRLS and in the National Assessments – both focussed on reading as a constructive and social process, and on children's use of reading for information and for pleasure (Eivers et al., 2005, Mullis, Martin & Sainsbury, 2015). PIRLS reading purposes are also covered in the PSEC, although classroom practice in Ireland tends to place more emphasis on Literary than Informational texts (Eivers et al., 2005, Kavanagh et al., 2015). PIRLS 2016 data also suggest that Irish pupils were more likely than the PIRLS average to be regularly exposed to a variety of fiction materials in class. This may have contributed to the relative strength on Literary texts demonstrated by Irish pupils.

Like most pupils in most countries, Irish pupils reported that they liked reading the texts and projects presented to them. In Ireland, ratings for the paper-based texts were slightly less positive than the international averages, and slightly more positive for the computer-based projects. Irish boys rated a variety of Literary texts less positively than did Irish girls, irrespective of the gender of the main protagonists. The only Literary text rated equally positively by Irish boys and girls featured animals, not humans, as the main protagonists. In contrast, boys generally, and Irish boys in particular, were more positively disposed than girls towards many Informational texts.

For PIRLS, the performance of Irish pupils (in terms of the advantage demonstrated over the international average) varied considerably, depending on the text. There was a weak relationship between pupils' enjoyment of a text and performance on that text. In contrast, for ePIRLS Irish pupils showed relatively little variation on performance or enjoyment by project.

In Ireland and internationally, multiple choice items were more likely to be correctly answered than constructed response items, a finding in line with previous research indicating that constructed response items are more likely to be skipped, particularly by weaker pupils (e.g., Lissitz et al., 2012; Reardon et al., 2018). However, slightly different patterns emerged for the paper and digital tests. On PIRLS there was little difference in the percentage of correct responses for constructed response and multiple choice items assessing the Retrieve/Infer subscale, but large differences on items assessing Interpret/Evaluate (in favour of multiple choice items). For ePIRLS, relative to PIRLS, there was a moderately larger difference in percent correct by item type for Retrieve/Infer items and a smaller gap for Interpret/Evaluate items. Also, there was slightly lower rates of skipping for constructed response items on ePIRLS. Thus, the manner in which pupils deal with constructed response and multiple choice items may be slightly different in a digital environment than on paper. However, it may also be due to differences in test content and item difficulty. Further assessment cycles may be able to clarify if the differences arise from assessment mode or content.

Of the material released in December 2017, two PIRLS texts (*Flowers on the Roof* and *Green Sea Turtles*) and one project (*Mars*) were selected for review. Both PIRLS texts were ones on which Irish pupils performed very well, but on which gender differences varied – from a 7% difference in

favour of girls on *Flowers on the Roof* to a 1% advantage for boys on *Green Sea Turtles*. On *Mars*, overall Irish performance was also very good, with no notable gender differences. The items selected illustrated the types of items on which Irish pupils displayed relative strengths and weaknesses, and where the gender gap was larger or smaller than the average.

For Literary texts, Irish pupils displayed a particular strength in recognising or explaining a character's feelings, motivations or actions as well as in the ability to infer and integrate multiple elements within narratives to provide descriptions, explanations and evaluations. A particular strength in Informational texts was the ability to answer items requiring response to two pieces of information within a text. Although there were a number of items on which Irish performance was much lower than expected, no common thread was discernible.

Regarding items on which Irish girls performed substantially better than boys, most were constructed response, some required provision of multiple pieces of information, and many required the reader to interpret a character's motivation or feelings. All Literary items on which girls substantially outperformed boys were from texts with human protagonists. Items on which Irish boys outperformed girls (although to a lesser degree) represented a mixture of Informational and Literary texts, and all but one came from texts in which animals featured strongly.

ePIRLS items on which Irish pupils did particularly well included items that assessed pupils' abilities to recognise an overall theme, examine the author's viewpoint, and integrate information across a number of webpages. Poorer performance was demonstrated on items requiring assimilation of information from images or graphics and inferring information based on a single webpage. Irish girls again showed a tendency to outperform boys on constructed response items, while over half of the items in which girls did better required the integration of information. Irish boys showed relative strengths on two items that related to dealing with visual information presented in maps, and interacting with particular elements of an online environment.

The item analyses confirm that, while overall reading skills in Ireland are very high relative to most other countries, there remains some room for improvement. Greater exposure to non-fiction, Informational texts would benefit Irish pupils, and might help to bring national performance on Informational texts up to the level of Literary texts. In particular, boys might benefit from a greater focus on Informational texts, as the pupil ratings indicated that boys were more likely to enjoy reading Informational texts and projects. In this regard, we echo previous recommendations (DES, 2011; Eivers et al., 2010) for greater use of non-fiction texts in Irish classrooms.

Irish pupils' performance (girls' in particular), was also lower than might be expected on some ePIRLS items incorporating visual content and some online navigational skills. While Irish pupils had good access to digital devices at home, they reported less frequent use in a school setting of digital devices than the ePIRLS average (Mullis et al., 2017b). Increased exposure to digital and graphical texts within a school context may be of benefit, not only in terms of performance on ePIRLS, but in supporting the types of skills needed to negotiate real life complex online environments.

In Ireland, girls outperformed boys on the overall PIRLS reading scale, and on many of the released items, particularly Literary texts. However, the gender gap on Literary texts has narrowed since PIRLS 2011. The data from 2016 also show that the gender gap can vary considerably, depending on the text in question. This underscores the importance of using a variety of text types and text themes in order to get pupils to engage with what they read.

It is often assumed that boys do not like to read texts about girls or women. The evidence from the pupil ratings of projects and texts show that this is largely untrue. In a related vein, a slightly unexpected finding was that Irish boys' performance (and interest) was relatively better on Literary

texts and items where the main protagonists were non-human. In contrast, Literary texts items on which Irish girls performed best featured human protagonists and most asked the reader to interpret or understand a character's motivation or feelings. On ePIRLS, Irish girls were slightly less positive than boys about the more science-based projects, but slightly more positive about the social studies or legends projects. As such, using varied texts and themes increases the possibility of pupil engagement with at least some texts.

Finally, it was notable that an incorrect assumption of maleness in particular contexts was still evident in a sizeable number of pupil responses. When pupils were asked how the writer of "The Gift of Curiosity" showed that they were in favour of Mars exploration, most wrote an answer beginning "the writer shows this by …" and did not specify a gender for the writer. However, even though she was clearly tagged in the by-line as Maria Green, almost one quarter of pupils identified the writer as male, slightly more than identified her as female. Assigning maleness to the more heroic character also occurred frequently in the text, *Pemba Sherpa*, despite the character's femaleness being a key element of the plot. It would seem that there remains work to be done in exploring gender stereotypes and we repeat the recommendation in Concannon-Gibney and Shiel (2012) about the "value in directing the attention of pupils (both boys and girls) to the ways in which gender is constructed socially, both in and out of school, how this impacts on pupils' own lives and is endorsed by others, and how gender is portrayed in texts" (p. 151).

References

Ainley, M., Hillman, K., & Hidi, S. (2002). Gender and interest processes in response to literary texts: Situational and individual interest. *Learning and Instruction, 12*, 411–428. DOI:10.1016/S0959-4752(01)00008-1

Concannon-Gibney, T. & Shiel, G. (2012). Reading Literacy in PIRLS 2011. In E. Eivers & A. Clerkin (Eds.), *National Schools, international contexts: Beyond the PIRLS and TIMSS test results* (pp. 77-104). Dublin: Educational Research Centre.

DES (Department of Education and Science) / NCCA (National Council for Curriculum and Assessment) (1999a). *Primary school English curriculum.* Dublin: Stationery Office.

DES (Department of Education and Science) / NCCA (National Council for Curriculum and Assessment) (1999b). *Primary school English curriculum: Teacher guidelines.* Dublin: Stationery Office.

DES (Department of Education and Skills) / NCCA (National Council for Curriculum and Assessment) (2015). *Primary language curriculum: English-medium schools.* Dublin: Department of Education and Skills.

DES (Department of Education and Skills). (2011). *Literacy and numeracy for learning and life: The national strategy to improve literacy and numeracy among children and young people 2011-2020*. Dublin: Government Publications.

Eivers, E., Close, S., Shiel, G., Millar, D., Clerkin, A., Gilleece., L. & Kiniry, J. (2010). *The 2009 National Assessments of Mathematics and English Reading.* Dublin: Stationery Office.

Eivers, E., Gilleece, L., & Delaney, E. (2017). *Reading achievement in PIRLS 2016: Initial report for Ireland.* Dublin: Educational Research Centre.

Eivers, E., Shiel, G., Perkins, R., & Cosgrove, J. (2005). *The 2004 National Assessment of English Reading.* Dublin: Educational Research Centre.

Foy, P. (2018). *PIRLS 2016 User Guide for the International Database.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

IEA (International Association for the Evaluation of Educational Achievement). (2018). *International Database Analyzer, version 4.0.* Hamburg: IEA Hamburg.

Kavanagh, L., Shiel, G., Gilleece, L., & Kiniry, J. (2015). *The 2014 National Assessments of English Reading and Mathematics. Volume 2: Context Report.* Dublin: Educational Research Centre.

Lafontaine, D. & Monseur, C. (2009). Gender gap in comparative studies of reading comprehension: To what extent do the test characteristics make a difference? *European Educational Research Journal, 8,* 69–79.

Le, L. (2006). *Analysis of Differential Item Functioning.* Paper prepared for the Annual Meetings of the American Educational Research Association in San Francisco, 7-11 April 2006. Retrieved from https://www.acer.org/files/analysis_of_dif.pdf

Lissitz, R.W., Hou, X., Cadman Slater, S. (2012). The Contribution of Constructed Response Items to Large Scale Assessment: Measuring and Understanding their Impact. *Journal of Applied Testing Technology, 13(3).* Retrieved from http://www.jattjournal.com/index.php/atp/article/view/48366

Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading, 32,* 199–214.

DOI:10.1111/j.1467-9817.2008.01389

Martin, M. O., Mullis, I. V. S., & Foy, P. (2015). PIRLS 2016 assessment design and specifications. In Mullis, I.V.S., & Martin, M.O. (2015). *PIRLS 2016 assessment framework* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S. (2015). Introduction. In Mullis, I.V.S., & Martin, M.O. (2015). *PIRLS 2016 assessment framework* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017a). *PIRLS 2016 international results in reading.* Retrieved from http://timssandpirls.bc.edu/pirls2016/international-results/

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017b). *ePIRLS 2016 international results in online informational reading.* Retrieved from http://timssandpirls.bc.edu/epirls2016/international-results/

Mullis, I.V.S., Martin, M.O., & Sainsbury, M. (2015). PIRLS 2016 reading framework. In Mullis, I.V.S., & Martin, M.O. (Eds.), *PIRLS 2016 assessment framework* (2nd ed.). (pp. 13-31). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S. & Prendergast, C.O. (2017). Developing the PIRLS 2016 achievement items. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in PIRLS 2016* (pp.1.1-1.29). Retrieved from https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-1.html

NCCA (National Council for Curriculum and Assessment). (2005). *English Curriculum: Additional Support Material.* Dublin: NCCA.

Reardon, S.F., Kalogrides, D., Fahle, E.M., Podolsky, A., & Zárate, R.C. (2018). The relationship between test item format and gender achievement gaps on math and ELA tests in 4th and 8th Grade. *Educational Researcher, 47.*

DOI: 10.3102/0013189X18762105

Shiel, G., Kavanagh, L., & Millar, D. (2014). *The 2014 National Assessments of English Reading and Mathematics. Volume 1: Performance report.* Dublin: Educational Research Centre.

Solheim, O.J., & Lundetræ, K. (2018). Can test construction account for varying gender differences in international reading achievement tests of children, adolescents and young adults? – A study based on Nordic results in PIRLS, PISA and PIAAC, *Assessment in Education: Principles, Policy & Practice, 25(1)*, 107-126, DOI:10.1080/0969594X.2016.1239612

Taube, K., & Munck, I. (1996). Gender differences at item level. In H. Wagemaker, K. Taube, I. Munck, G. Kontogiannopoulou-Polydorides, & M. Martin (Eds.), *Are girls better readers? Gender differences in 32 countries* (pp. 53–98). Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

ISSN 2565-5604