Reading achievement in PIRLS 2016: Initial report for Ireland

Eemer Eivers
Lorraine Gilleece
Emer Delaney

2017

Educational Research Centre
Foras Taighde ar Oideachas

Reading achievement in PIRLS 2016: Initial report for Ireland

Eemer Eivers
Lorraine Gilleece
Emer Delaney

2017

Educational Research Centre
Foras Taighde ar Oideachas

www.erc.ie

Cataloguing-in-publication data:

Eivers, Eemer.

Reading achievement in PIRLS 2016: initial report for Ireland
Eemer Eivers, Lorraine Gilleece, Emer Delaney.
Dublin: Educational Research Centre.

iv, 47 p.; 30cm

Includes bibliographical references.

1. Progress in International Reading Literacy Study
2. Reading (Primary) – Ireland – Evaluation
3. Literacy (Primary) – Ireland – Evaluation
4. Academic achievement
5. Educational tests and measurements

2017
I Title. II Gilleece, Lorraine. III Delaney, Emer.
372.4 072 – dc23

ISBN-10: 0-900440-570
ISBN-13: 978-0-900440-571

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

Preface

PIRLS (**P**rogress in **I**nternational **R**eading **L**iteracy **S**tudy) is a project of the International Association for the Evaluation of Educational Achievement (IEA). It is designed to assess the reading achievement of Fourth Class pupils. It was first conducted in 2001, while Ireland took part for the first time in 2011. The 2016 study included an optional component assessing online reading skills, called ePIRLS.

The present volume is published at the same time as the IEA's main reports on PIRLS (Mullis, Martin, Foy, & Hooper, 2017a) and ePIRLS (Mullis, Martin, Foy, & Hooper, 2017b). It summarises the main achievement-related findings from both studies, focussing on Irish performance and findings most likely to be of interest to an Irish audience. A more in-depth series of reports on PIRLS and ePIRLS from an Irish perspective will be released in 2018.

Acknowledgements

We acknowledge the support of the National Advisory Committee which was appointed by the Minister of Education and Skills to assist in implementing PIRLS. As well as the Chair, Yvonne Keating (Department of Education and Skills), the members of the Committee were: Aedín Ní Thuathail (Irish Primary Principals' Network); Áine Lynch (National Parents Council-Primary); Arlene Forster (National Council for Curriculum and Assessment); Deirbhile Nic Craith (Irish National Teachers' Organisation); Eddie Fox (Educate Together); Máirín Ní Chéileachair (Gaelscoileanna); Mia Treacy (Professional Development Service for Teachers); Seán Delaney (Marino Institute of Education); Suzanne Cobbe (Catholic Primary School Management Association); and Eamonn Moran and Breda Naughton (Department of Education and Skills).

We thank the CEO of the Educational Research Centre, Peter Archer, for guidance throughout the study. Thanks are due to many ERC staff, but particularly to Brenda Donohue, John Coyle, Anne Comey, Patricia Gaffney and David Millar. We also thank members of the Department of Education and Skills Inspectorate for their help in administering the National Quality Monitoring Programme.

A special thanks to the pupils of St Fiachra's Senior National School, Beaumont, for making our report cover so appealing, and to their teacher, Ms Mathews, and principal, Mr Creaner, for facilitating the photos. Finally, and most importantly, we thank all school principals, teachers, pupils and parents who participated in PIRLS 2016. Without their help, the study and this report would not be possible.

Table of contents

Chapter 1:

Introduction

Chapter 1
Introduction

The Progress in International Reading Literacy Study (PIRLS) is the world's largest comparative study of reading achievement at primary level. First run in 2001, it takes place every five years. Ireland participated for the first time in 2011 (Eivers & Clerkin, 2012; Mullis, Martin, Foy, & Drucker, 2012). The most recent cycle included a new element called ePIRLS, an assessment of online reading skills and digital literacy. This report summarises the main findings from PIRLS and ePIRLS 2016, with a particular focus on performance of pupils in Ireland. It is designed to complement the two main international reports on PIRLS and ePIRLS, and the PIRLS encyclopedia (all available from https://timssandpirls.bc.edu/).

The international reports for PIRLS (Mullis, Martin, Foy & Hooper, 2017a) and ePIRLS (Mullis, Martin, Foy & Hooper, 2017b) describe the performance of pupils in all participating countries, including some analyses of how performance relates to characteristics of the home, school and classroom environment, as well as to characteristics of individual pupils. The PIRLS encyclopedia (Mullis, Martin, Goh & Prendergast, 2017) provides contexts within which national performance can be interpreted. It contains information about all participating countries on key characteristics – e.g., structure of the education system, nature of the reading/language curricula, and, information about teacher qualification and certification.

The present report is divided into six chapters, the first of which provides an overview of the study, outlining which countries took part, what was assessed and the nature of the assessment materials. It also identifies a common subset of participating countries – comparison countries – on which we focus more closely as they are considered to be of particular interest. Chapter 2 describes how PIRLS and ePIRLS were administered in Ireland, and how many pupils took part, both in Ireland and overall. Chapter 3 provides the main PIRLS and ePIRLS results for all countries, as well as summarising gender differences and trends in performance (where available) for Ireland and comparison countries.

Chapter 4 outlines the types of skills exemplifying performance at each of the International Benchmarks. It includes the percentages of pupils in Ireland and comparison countries reaching each Benchmark level, and some trend data for PIRLS. Chapter 5 looks at performance on the subscales (reading processes and purposes), overall and by gender. Chapter 6 summarises the main findings and flags forthcoming reports and analyses.

What are PIRLS and ePIRLS?

PIRLS assesses the reading achievement of pupils in Fourth Class (or the equivalent grade level) in Ireland and in many other countries. In 2016, 50 countries/regions and 11 benchmarking

participants[1] took part in at least one element of PIRLS (Table 1.1). In total, more than 340,000 pupils, 330,000 parents, 16,000 teachers, and 12,000 schools took part. PIRLS is conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA). The PIRLS International Study Center at Boston College is responsible for the management of the study at an international level. Within each participating country, management of the study is the responsibility of a national study centre. In Ireland, the Educational Research Centre fulfils this role.

In 2016, the design of PIRLS was adjusted to include two additional components –PIRLS Literacy and ePIRLS. This report largely focusses on PIRLS and ePIRLS, as these are the elements in which Ireland took part.

Table 1.1: Countries and benchmark regions participating in PIRLS, with those also participating in ePIRLS shown in orange

Australia	**Ireland**	Slovak Republic
Austria	**Israel**	**Slovenia**
Azerbaijan	**Italy**	South Africa
Bahrain	Kazakhstan	Spain
Belgium (Flemish)	Kuwait	**Sweden**
Belgium (French)	Latvia	Trinidad & Tobago
Bulgaria	Lithuania	**United Arab Emirates**
Canada	Macao SAR	**United States**
Chile	Malta	
Chinese Taipei	Morocco	**Benchmarking participants**
Czech Republic	Netherlands	*Buenos Aires, Argentina*
Denmark	New Zealand	*Ontario, Canada*
Egypt	Northern Ireland	*Quebec, Canada*
England	**Norway (Grade 5)**	*Denmark (Grade 3)*
Finland	Oman	*Norway (Grade 4)*
France	Poland	*Moscow City, Russia*
Georgia	**Portugal**	*Eng./Afr./Zulu – RSA (Grade 5)*
Germany	Qatar	*Andalusia, Spain*
Hong Kong SAR	Russian Federation	*Madrid, Spain*
Hungary	Saudi Arabia	***Abu Dhabi, UAE***
Iran	**Singapore**	***Dubai, UAE***

PIRLS Literacy is a version of PIRLS designed for use in countries where a majority of pupils are still developing fundamental reading skills. It is characterised by texts that are shorter and less complex than those used in PIRLS, although there is some overlap in test content – that is, some PIRLS texts appear in PIRLS Literacy and some PIRLS Literacy texts appear in PIRLS. Overlapping content means that scores for countries that took part in PIRLS Literacy can be

1 Benchmarking participants are regional or national entities that follow PIRLS procedures, but are not a country-level Fourth grade sample. For example, Canada included two additional samples to provide more detail on performance in Ontario and Quebec, while Denmark trialled the PIRLS Literacy assessment on Third grade pupils.

reported in the same tables and on the same scales as countries that took part in PIRLS. Six countries took part in PIRLS Literacy, of which two (Iran and Morocco) also took part in PIRLS.

ePIRLS is an assessment of digital literacy, measuring pupils' ability to read in an online environment, and in a non-linear fashion. It was an optional addition to the main PIRLS assessment. Fourteen countries and two benchmarking regions took part in ePIRLS. In contrast to PIRLS Literacy, achievement scores for PIRLS and ePIRLS are reported separately, but a country's score on each of the two assessments can be compared directly with each other.

What do PIRLS and ePIRLS assess?

The content of the PIRLS and ePIRLS tests and all background data collected is guided by a framework, summarised here and available in full in Mullis and Martin (2015) (see http://timss.bc.edu/pirls2016/framework.html).

Assessment frameworks

An assessment framework defines who and what is to be measured, and how measurement will take place. The PIRLS framework identifies children in their fourth year of formal schooling as the appropriate target for the assessment. Fourth grade (equivalent to Fourth Class in Ireland) was chosen as it typically marks the point at which pupils have learned how to read and are beginning to read to learn (Mullis & Martin, 2015). Regarding what is to be assessed, Mullis and Martin (2015, p.12) define reading as follows:

> *Reading literacy is 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.*

Thus, the definition emphasises reading comprehension rather than reading fluency, recognises that there are many types of text (including digital forms), and that social interaction allows pupils to understand, appreciate and share what they have read or will read. With that definition as a starting point, the framework identifies the two main purposes for which children read, and four 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*.

The reading purposes and processes are used to guide the assessment content. Thus, the PIRLS assessment is equally divided between Literary and Informational texts, whereas ePIRLS only assesses pupils' ability to read for Informational purposes (Table 1.2). However, although ePIRLS uses a simulated Internet environment and incorporates a set of navigation skills

required to locate and use information on the Internet, the emphasis remains on assessing reading comprehension rather than navigation skills. Thus, both PIRLS and ePIRLS place the same relative emphases on the four comprehension processes.

In PIRLS and ePIRLS, 60% of the assessment is directed at examining pupils' ability to make straightforward inferences and to interpret and integrate ideas and information. Only 20% of each assessment examines the (generally easier) process of retrieving explicitly stated information, while a further 20% assesses the (generally more difficult) process of evaluating and critiquing content and textual elements. In contrast, PIRLS Literacy is designed to be an assessment for children with less developed reading skills. Thus, 50% of content assesses basic retrieval processes, 25% examines ability to make straightforward inferences, and only 25% examines the two more complex processes (interpreting and integrating, and evaluating and critiquing).

Table 1.2: Percentages of the PIRLS, ePIRLS and PIRLS Literacy reading assessments devoted to each reading purpose and comprehension process

	PIRLS	ePIRLS	PIRLS Literacy[1]
Purposes for reading			
Literary Experience	50%	0%	50%
Acquire and use information	50%	100%	50%
Process of comprehension			
Focus on and retrieve explicitly stated information	20%	20%	50%
Make straightforward inferences	30%	30%	25%
Interpret and integrate ideas and information	30%	30%	25%
Evaluate and critique content and textual elements	20%	20%	

Adapted from Mullis & Martin (2015).
[1]Not administered in Ireland

PIRLS also has a detailed context questionnaire framework, which guided the development of all questionnaires and the PIRLS 2016 Encyclopedia. The context framework is an essential element of PIRLS and is designed to provide policymakers with insights into how broader elements of education systems can support reading development. The context questionnaire framework consists of five areas that are known to be relevant to the acquisition and development of reading skills. These are:

- National and community contexts (e.g., national language(s), demographics, structure of the education system, the teaching of reading).
- Home contexts (e.g., language(s) spoken in the home, home resources for learning, early literacy activities).
- School contexts (e.g., location, size and socioeconomic composition, school resources, teacher job satisfaction, school "climate").
- Classroom contexts (e.g., teacher qualifications and experience, classroom resources, instructional time and practices).
- Pupil characteristics and attitudes towards learning (e.g., pupil reading activities, motivation, self-concept in relation to reading).

Full detail of the context framework is provided in Mullis and Martin (2015).

What are the tests like?

Both PIRLS and PIRLS Literacy were presented to pupils in paper test booklets, whereas ePIRLS was presented digitally, using a simulated online environment. On paper, each pupil was randomly assigned one of 16 different booklets. Each booklet had one Literary text (approximately 800 words in length) and one Informational text (600-900 words), each with related questions. Half of booklets had the Literary text first, while the other half had the Informational text at the start. Overall, there were five Literary and five Informational texts in PIRLS. These, and an Informational text and a Literary text from PIRLS Literacy (each of which were shorter and less complex than the PIRLS texts) were rotated across all 16 booklets.

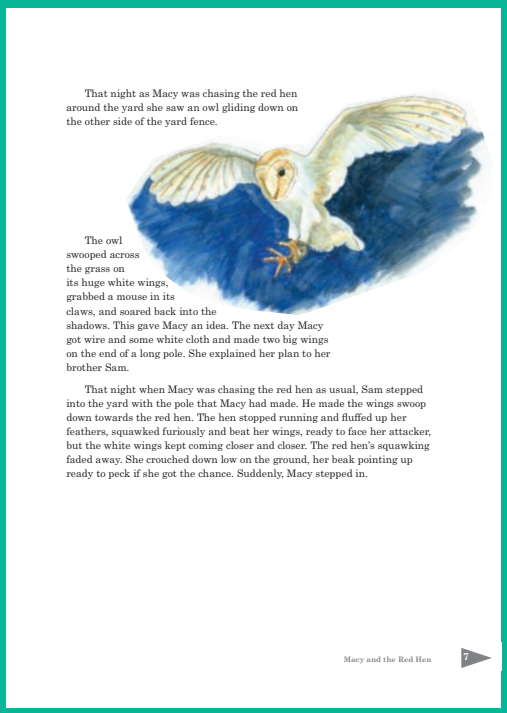
ePIRLS comprised five "projects", each of which involved two to three different websites totalling about 1000 words across 5 to 10 webpages (Mullis & Martin, 2015). Pupils completed two ePIRLS projects. Pupils were guided through the test by either a male or female "teacher" avatar who gave directions. All ePIRLS projects related to science or social science topics.

Sample test content

Four PIRLS texts and two ePIRLS projects were released after the study. The PIRLS texts are available from the main international report (Mullis et al., 2017a) and from the following website (https://timssandpirls.bc.edu/). In addition, two released texts, related questions, and sample answers are available from www.erc.ie/pirls. These include information about the percentages of pupils in Ireland and internationally who were able to answer each question correctly.

The ePIRLS projects are also available from https://timssandpirls.bc.edu/take-the-epirls-assessment/ where they can be taken as a real-life, interactive test, as experienced by pupils.

On the right is part of a PIRLS Literary text and an extract from an ePIRLS project.



That night as Macy was chasing the red hen around the yard she saw an owl gliding down on the other side of the yard fence.

The owl swooped across the grass on its huge white wings, grabbed a mouse in its claws, and soared back into the shadows. This gave Macy an idea. The next day Macy got wire and some white cloth and made two big wings on the end of a long pole. She explained her plan to her brother Sam.

That night when Macy was chasing the red hen as usual, Sam stepped into the yard with the pole that Macy had made. He made the wings swoop down towards the red hen. The hen stopped running and fluffed up her feathers, squawked furiously and beat her wings, ready to face her attacker, but the white wings kept coming closer and closer. The red hen's squawking faded away. She crouched down low on the ground, her beak pointing up ready to peck if she got the chance. Suddenly, Macy stepped in.

Macy and the Red Hen 7

Material shown is property of the IEA. Further usage and/or reproduction requires permission from IEA.

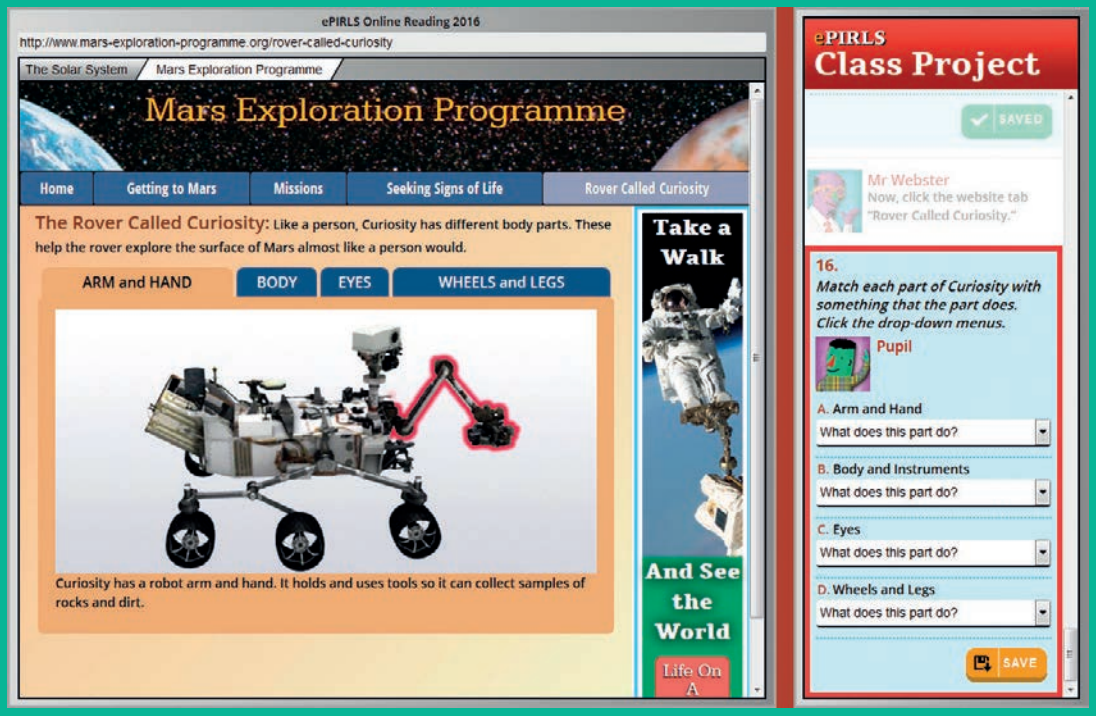
For both the paper and digital assessments, total test time was 80 minutes, with a short break between the first and second texts or projects. All assessments included a mixture of multiple choice and constructed-response items (where pupils write their answer, which can be as short as a single word, or may extend to a few sentences). All constructed-response items were scored by a trained team of coders following procedures set out by the International Study Center.

Comparison countries

As noted, a large number of countries took part in PIRLS 2016, some of which are of more interest than others to an Irish audience. Therefore, a set of "comparison countries" was chosen, based on one or more of the following criteria:

- high performance on PIRLS and/or ePIRLS.
- cultural and/or linguistic similarity to Ireland.

The selected comparison countries, in alphabetical order, are: Australia; England; Finland; Hong Kong SAR; New Zealand; Northern Ireland; Norway; Poland; Russian Federation; Singapore; and, United States. Henceforth, they are typically presented in descending order of overall mean score, or percentages attaining specified benchmarks.



Understanding scores

PIRLS Centrepoint

PIRLS compares country performance against a scale centrepoint of 500 – the mean score from the *first* time the study was conducted in 2001. It is a point of reference that remains constant from assessment to assessment, and allows countries to monitor changes in national performance over successive cycles. It is **not** the international average of the 2016 cycle (most country means were significantly above the centrepoint).

Subscale and overall scale scores

Subscale scores are developed for reading purposes and comprehension processes. Overall scales and subscales are created independently of each other, and all set to a centrepoint of 500. Because they are scaled independently, subscale and overall scale scores will be in close, but not always exact, agreement.

General statistical terms and concepts

Like most studies of pupil achievement, PIRLS uses a "representative sample" of a population to gauge performance for the population as a whole – the test scores of 4,607 Fourth Class pupils were used to give a very good, but not perfect, estimate of the reading achievement of *all* Fourth Class pupils in Ireland in 2016. As they are estimates, a margin of error is always attached to each score for each country and each subset of pupils.

When comparing any set of scores, consider three things: the size of the score gap, the standard errors and the confidence intervals. All three inform the judgement about whether a difference matters, and/or is statistically significant.

Statistical significance

A statistically *significant* difference is one that a statistical test has established is unlikely to be due to chance. As a rough guide, a difference of two or three points between mean scores is rarely statistically significant or meaningful in a real-world sense.

Standard error (SE)

Samples are used to estimate the mean score for a population, but the accuracy of the estimate can vary considerably depending on the sample (e.g., very small or biased samples will not provide good estimates). Thus, we also consider the standard error – a measure of how accurately the sample mean reflects the population mean. The smaller the standard error, the more likely the sample mean is to be close to the population mean.

Confidence Intervals

Mean scores and standard errors combine to produce confidence intervals (a score range within which the "true" population score is highly likely to fall). If their confidence interval ranges do not overlap, scores are significantly different.

For a rough calculation of a confidence interval, multiply the SE by 2 and add that amount either side of the mean. For example, if a mean score is 520 (SE=1.5), the confidence interval is 517-523. If compared against a mean score of 523 (SE=1.0), the difference is unlikely to be statistically significant as the confidence interval around the second mean is 521-525 – that is, the two ranges overlap.

Chapter 2:

PIRLS and ePIRLS in Ireland

Chapter 2:
PIRLS and ePIRLS in Ireland

The main study in 2016 was preceded by a field trial in spring 2015, and, in the case of ePIRLS, by an earlier pre-pilot in Ireland and a small number of other countries. The pre-pilot was used to generate "real-life" answers to the questions in ePIRLS, which informed the creation of a scoring guide for use in all participating countries. It was also used to trial the software supporting ePIRLS.

In Ireland, the field trial took place in March 2015 in 32 schools. As well as trialling procedures and software, it was used to determine the most appropriate model of ePIRLS administration. The main conclusion from the field trial was that using school resources was highly problematic in Ireland, contributing to lost data, technical problems, and pupils being unable to access the ePIRLS test. A second conclusion was that using rented laptops in a primary school setting was problematic. The rented models were typically large, and with a short battery life, meaning they took up a lot of space and needed to be plugged in for the duration of the test, causing trip hazards, increasing setting up and removal time, and significantly intruding into regular teaching time. A third conclusion was that most schools would require training and/or technical support on the day of ePIRLS testing, and that it was not feasible for one person to set up and support multiple simultaneous sessions within the same school.

Arising from the field trial, it was decided to administer ePIRLS to a subsample only, and only using devices supplied by the ERC. A modified sample was negotiated with the International Study Centre and Statistics Canada, whereby a random subset of PIRLS pupils within each school were selected to also complete ePIRLS. A set of 400 small laptops was bought, and later sold to participating schools at a large discount. The main advantages of this approach were: less space required; safer classroom environment due to lack of cables; much faster setting up and removal time; uniform administration and less variety in technical issues; and, offering schools a tangible incentive to participate.

Who took part in PIRLS 2016?

In 2014, the Educational Research Centre supplied Statistics Canada (the agency which draws school samples for all countries in PIRLS) with a list of all primary schools with at least one Second Class pupil. Second Class was selected as the target grade for sampling, as almost all such pupils would be expected to be in Fourth Class in 2016. From the full list of schools, random

samples of schools for the field trial and main study were drawn simultaneously, meaning that no school was selected to take part in both. For the 150 schools selected to participate in the main study, the sample was balanced by school enrolment size, language of instruction, DEIS status and gender mix.

When the selected schools were contacted in late 2015, one very small school was excluded as it had no current Fourth Class pupils, while another was excluded as it had amalgamated with a neighbouring school. All other selected schools participated in the study. Within each school, up to two Fourth Class groups were randomly chosen. In schools with two or fewer Fourth Class groups, all pupils were selected to participate in the PIRLS pencil-and-paper test. In schools with three or more Fourth Class groups, two classes were randomly selected. Within the final sample of 148 schools, 4,881 Fourth Class pupils were selected for participation in the PIRLS written test.

In schools with 22 or fewer Fourth Class pupils, all PIRLS pupils were also selected to participate in ePIRLS. In larger schools, a subset of PIRLS pupils (from across all selected classes) was randomly selected for participation in ePIRLS. Generally, a random sample of 22 was selected from the PIRLS sample but where this resulted in the exclusion of a small number of pupils, School Coordinators could opt to administer ePIRLS to all pupils. In 142 of the 148 participating schools, the number of pupils selected for ePIRLS was 22 or fewer. In total, 2,767 pupils were selected for ePIRLS.

Table 2.1 shows response rates for all aspects of PIRLS and ePIRLS. PIRLS had a weighted student participation rate of 96% (unweighted 94%). Of the original sample of 4,881 pupils, reasons for non-testing were as follows: left the school in the period between sampling and testing (30 pupils); absent on the test day (194); parental refusal (6); excluded by teacher due to intellectual disability or limited English proficiency (44). The response rates for the pupil questionnaire are almost identical to those for the PIRLS test as both were typically administered on the same day.

All 44 pupils excluded from PIRLS were automatically considered ineligible for selection for ePIRLS, thus somewhat artificially increasing the ePIRLS exclusion rate. Had random selection been applied, only some of the 44 would have been selected for ePIRLS, whereas the exclusion method applied assumes all were selected for ePIRLS. Reasons to explain the remaining non-participants in ePIRLS were: left the school between selection and test date (18 pupils); absent on the test day (148); parental permission refused (2). A further 82 pupils who completed ePIRLS were later excluded from the data as they had not also completed the PIRLS test. Thus, the ePIRLS weighted pupil participation rate for Ireland is 91% (unweighted 89%).

There were also very high response rates to all questionnaires. Considering all pupils selected to take part in PIRLS, 92% of Parent Questionnaires were returned, 99% of School Questionnaires, and 100% of Teacher Questionnaires. These very high response rates demonstrate the high level of cooperation of school staff, pupils and parents, and mean that the resulting data can be considered representative of the broader national population.

Table 2.1: Response rates to PIRLS and ePIRLS 2016 tests and questionnaires in Ireland		
Instrument	No. of sampled PIRLS pupils = 4881	
	N	% (unweighted)
PIRLS test	4607	94.4
PIRLS pupil questionnaire	4604	94.3
Parent questionnaire	4504	92.3
	No. of ePIRLS pupils =2767	
ePIRLS test	2473	89.4
	No. of classes = 219	
Teacher questionnaire	219	100
	No. of schools = 148	
School questionnaire	146	98.6

NOTE: The numbers and percentages shown related to the total sample of pupils. These data differ slightly from the international reports, which show weighted percentages for the subset of pupils who completed the PIRLS test.

Test administration

In Ireland, both PIRLS and ePIRLS assessments took place during April 2016. In each school, the PIRLS assessment preceded the ePIRLS assessment. Both assessments took place in the morning, during regular school hours, but on separate days. As ePIRLS was the first large-scale computer-based assessment at primary level in Ireland, all teachers acting as a Test Administrator (TA) received prior training on the administration of ePIRLS.

PIRLS

PIRLS was administered by a teacher from the participating school, but not the regular teacher of the selected class (e.g., if two classes were selected, the two class teachers typically switched classes to administer the assessment). After the assessment, and usually on the same day, pupils completed a questionnaire. The test was administered in two timed sessions of 40 minutes each, with a short break in the middle. The Pupil Questionnaire was untimed, but usually took approximately 20-30 minutes to complete.

ePIRLS

For almost all ePIRLS sessions, a technical support person (contracted and trained by the ERC) delivered and set up laptops in a classroom. The support person remained in the room while a teacher from the school administered the assessment, providing assistance if needed, and removing the laptops after testing (Table 2.2). In 10 schools, laptops were delivered before the test date, and telephone technical support was available during the assessment. All 10 had very few sampled pupils, and were in the vicinity of/on the way to another larger participating school. Five schools tested using their own resources. These schools had few sampled pupils, were not near any other sampled schools, and had the capacity to run the assessment on their own computers. Finally, one school was excluded from ePIRLS testing as it had no computer resources, was geographically very isolated, and had only two eligible pupils.

Table 2.2: ePIRLS main study administration mode in Ireland	
Administration Mode	**N**
Laptops and onsite technical support	132
Laptops and telephone technical support	10
School's own computer resources	5
No ePIRLS testing	1
Total	**148**

All test data were stored in encrypted format on either the hard drive of ERC laptops or on a USB. In all cases, the anonymised and encrypted data were uploaded to a secure FTP server the same day the session was completed – usually by technical support staff.

Quality Monitoring

A requirement of participation in PIRLS is that at least 10% of schools are visited by International Quality Control Monitors, employed directly by the international consortium. International monitors are complemented by National Quality Control Monitors. In Ireland, 15 schools were visited during both PIRLS and ePIRLS testing as part of the international monitoring programme, while 17 were visited as part of the national programme, conducted by members of the Department of Education and Skills Inspectorate.

The monitors observed sessions and evaluated how well the study was implemented in each country according to agreed international procedures. Feedback from the monitors in Ireland indicated very high levels of compliance with study procedures and very high levels of pupil cooperation and engagement (particularly with ePIRLS).

Other quality control measures included:

- independent verification of all national translations and adaptations of instruments.
- multiple-scoring of up to 30% of constructed-response PIRLS items (and a comparable procedure for ePIRLS) in each country, to assess within-country scorer reliability for 2016.
- scorers in Ireland also scoring responses collected in Ireland during PIRLS 2011, to assess trend scoring reliability (i.e., agreement between 2011 and 2016).
- scorers in every country also scoring a common set of responses collected during 2011, to monitor cross-country reliability in scoring (agreement between scorers in different countries in 2016).
- minimum response rates required for a country's data to be included in the international database.

Chapter 3:

Overall PIRLS and ePIRLS results

Chapter 3
Overall PIRLS and ePIRLS results

This chapter summarises overall performance on PIRLS and ePIRLS for Ireland and for all participating countries and benchmarking participants. It also provides some information about trends on overall PIRLS performance since 2011, for Ireland and selected *comparison countries* (see Chapter 1). Information about specific aspects of performance, such as the percentages of pupils reaching International Benchmarks, or performance across reading purposes and comprehension processes is provided in later chapters. Full details of the relative performance of all participating countries (including 95% confidence intervals and multiple comparison tables of performance for each country) are available in the PIRLS and ePIRLS international reports (Mullis et al., 2017a; Mullis et al., 2017b).

Comparing country mean scores

The next three tables show mean scores, standard errors and standard deviations for countries and benchmarking participants on PIRLS and ePIRLS. Each country/participant mean is compared to Ireland's, and to the PIRLS centrepoint of 500 (the anchor mean from the 34 participants in PIRLS 2001, with a corresponding standard deviation of 100). Note that differences of two or three points are rarely **statistically** significant, due to the error margin around the estimated means.

Vs Ireland:
Green shading = mean score significantly higher than Ireland.
Unshaded = mean not significantly different.
Yellow shading = mean significantly lower.

Vs Centrepoint:
↑ = mean score significantly above the centrepoint.
↔ = mean not significantly different.
↓ = mean significantly lower.

Overall PIRLS performance

The highest performing countries on PIRLS 2016 were the Russian Federation (581) and Singapore (576), whose mean scores were at least three-quarters of a standard deviation above the study centrepoint of 500 (Table 3.1). They were closely followed by Hong Kong SAR, Ireland, Finland, Poland and Northern Ireland, whose mean scores ranged between 569 and 565, and did not differ

significantly from each other. The next group of countries were Norway[1] and Chinese Taipei, with means of 559.[2]

Ireland's mean achievement score was 567. Only the Russian Federation and Singapore achieved significantly higher mean scores, while the Irish mean was significantly higher than the mean in 43 countries, including England, the United States, the Netherlands, Australia, Germany and New Zealand. Four countries (Hong Kong SAR, Finland, Poland and Northern Ireland) achieved means that did not differ significantly from Ireland's. No EU or OECD country obtained an overall mean score that was higher than Ireland's.

Table 3.1: Mean country reading scores, standard errors and standard deviations in PIRLS 2016, and position relative to the Irish mean and PIRLS centrepoint

	Mean	SE	SD	vs centrepoint		Mean	SE	SD	vs centrepoint
Russian Federation	581	2.2	66	↑	Germany	537	3.2	78	↑
Singapore	576	3.2	80	↑	Kazakhstan	536	2.5	63	↑
Hong Kong SAR	569	2.7	64	↑	Slovak Republic	535	3.1	81	↑
Ireland	**567**	2.5	74	↑	Israel	530	2.5	90	↑
Finland	566	1.8	67	↑	Portugal	528	2.3	65	↑
Poland	565	2.1	72	↑	Spain	528	1.7	65	↑
Northern Ireland	565	2.2	80	↑	Belgium (Flemish)	525	1.9	61	↑
Norway (Grade 5)	559	2.3	65	↑	New Zealand	523	2.2	91	↑
Chinese Taipei	559	2.0	64	↑	France	511	2.2	69	↑
England	559	1.9	79	↑	Belgium (French)	497	2.6	69	↔
Latvia	558	1.7	62	↑	Chile	494	2.5	79	↓
Sweden	555	2.4	67	↑	Georgia	488	2.8	79	↓
Hungary	554	2.9	75	↑	Trinidad and Tobago	479	3.3	94	↓
Bulgaria	552	4.2	85	↑	Azerbaijan	472	4.2	86	↓
United States	549	3.1	78	↑	Malta	452	1.8	90	↓
Lithuania	548	2.6	69	↑	United Arab Emirates	450	3.2	111	↓
Italy	548	2.2	65	↑	Bahrain	446	2.3	98	↓
Denmark	547	2.1	68	↑	Qatar	442	1.8	110	↓
Macao SAR	546	1.0	66	↑	Saudi Arabia	430	4.2	98	↓
Netherlands	545	1.7	60	↑	Iran	428	4.0	108	↓
Australia	544	2.5	84	↑	Oman	418	3.3	106	↓
Czech Republic	543	2.1	68	↑	Kuwait	393	4.1	105	↓
Canada	543	1.8	76	↑	Morocco	358	3.9	107	↓
Slovenia	542	2.0	72	↑	Egypt	330	5.6	124	↓
Austria	541	2.4	65	↑	South Africa	320	4.4	106	↓

1 Unlike in PIRLS 2011, Norway assessed pupils at Grade 5, as the average age of pupils in this grade (10.8) is closer to the PIRLS average than is the case at Grade 4.

2 England, Norway and Chinese Taipei have almost identical mean scores. However, the means for Norway and Chinese Taipei do not differ significantly from those of Poland and Northern Ireland, whereas England's mean is significantly lower. As such, Norway and Chinese Taipei can be viewed as a distinct third group of countries.

Table 3.2 shows the performance of PIRLS international benchmarking participants, relative to Ireland. Reflecting the Russian Federation's overall high performance, Moscow City was the highest performing benchmarking participant (mean score: 612) and the only such entity to significantly outperform Ireland. Other high-performing benchmarking participants were Madrid (549), Quebec (547) and Ontario (544), all of whom achieved means significantly lower than Ireland's. Both Norway and Denmark also took part in PIRLS at an earlier grade (equivalent to Third Class in Ireland) as a benchmarking participant. Norwegian Fourth Grade pupils achieved a mean of 517 while Danish Third Grade pupils achieved a mean score of 501.

Table 3.2: Mean reading scores, standard errors and standard deviations in PIRLS 2016, and position relative to the Irish mean and PIRLS centrepoint for benchmarking participants

	Mean	SE	SD	vs centrepoint
Moscow City, R. Fed	612	2.2	62	↑
Madrid, Spain	549	2.0	59	↑
Quebec, Canada	547	2.8	65	↑
Ontario, Canada	544	3.2	77	↑
Andalusia, Spain	525	2.1	64	↑
Norway (Grade 4)	517	2.0	70	↑
Dubai, UAE	515	1.9	98	↑
Denmark (Grade 3)	501	2.7	85	↔
Buenos Aires, Argentina	480	3.1	83	↓
Abu Dhabi, UAE	414	4.7	109	↓
Eng./Afr./Zulu–RSA (Grade 5)	406	6.0	103	↓

Overall ePIRLS performance

Table 3.3 shows the mean scores for the 14 countries and two benchmarking participants that took part in ePIRLS 2016, and their scores relative to the Irish mean. Pupils in Singapore achieved a mean ePIRLS score of 588, significantly higher than that achieved by pupils 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 in ePIRLS, including Sweden, the United States, Chinese Taipei, and Canada. Ireland's mean score was significantly higher than any other EU member state, and significantly higher than all OECD countries apart from Norway.

Comparing PIRLS and ePIRLS scores

ePIRLS results are reported on the same scale as PIRLS, with two important consequences. First, you can directly compare PIRLS and ePIRLS scores (e.g., we can say that a pupil who scores 530 on ePIRLS and 510 on PIRLS performed *better* on the digital assessment).

Second, to facilitate the direct comparison, PIRLS results shown in ePIRLS comparisons contain only those pupils who also completed ePIRLS. For example, Ireland's overall mean PIRLS score of 567 is based on the 4,607 pupils who completed PIRLS. It differs by a point from the PIRLS mean score used in ePIRLS comparisons (566), as the latter is based on the 2,473 pupils who completed both PIRLS and ePIRLS.

Table 3.3 also links ePIRLS to performance on PIRLS overall and on the PIRLS Informational subscale (as ePIRLS includes only Informational texts). Full details of performance on the PIRLS Informational scale is contained in Chapter 5. In the case of Ireland, the mean scores on all three scales are very similar. Compared to the Irish ePIRLS mean of 567, the mean PIRLS score was 566 for the subset of pupils who completed both PIRLS and ePIRLS, and the corresponding Informational mean score was 564. Generally, country mean scores across the three score types tended to be somewhat similar. In seven countries, the difference in mean scores on PIRLS and ePIRLS was eight points or fewer, while in eight countries, the difference in mean scores on the PIRLS Informational scale and ePIRLS was nine points or fewer.

Table 3.3: Mean country reading scores, standard errors and standard deviations in ePIRLS 2016, and position relative to the Irish mean

	ePIRLS Mean	SE	SD	PIRLS Mean[1]	PIRLS Informational Mean[1]
Singapore	588	3.0	78	576	579
Norway (Grade 5)	568	2.2	63	560	560
Ireland	**567**	2.5	71	566	564
Sweden	559	2.3	65	555	555
Denmark	558	2.2	66	548	544
United States	557	2.6	74	550	543
Chinese Taipei	546	2.0	66	559	569
Canada	543	3.2	74	543	540
Israel	536	2.3	82	532	530
Italy	532	2.1	62	548	549
Slovenia	525	1.9	68	543	544
Portugal	522	2.2	63	528	528
Georgia	477	3.3	73	489	487
United Arab Emirates	468	2.2	101	451	460
Benchmarking participants					
Dubai, UAE	528	1.6	89	516	524
Abu Dhabi, UAE	431	4.1	103	414	422

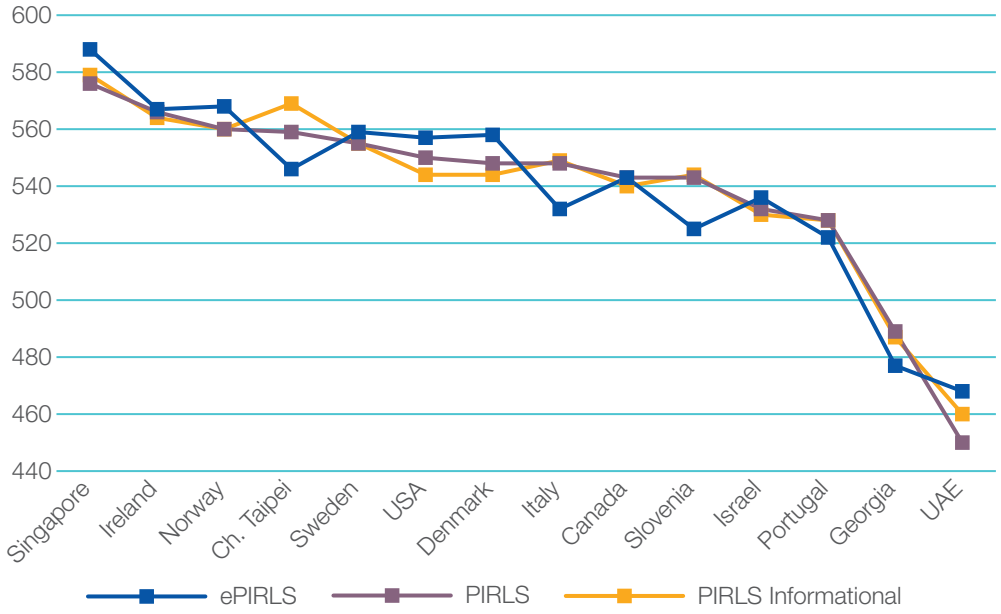
[1] Values shown are based on pupils who participated in both PIRLS and ePIRLS and may differ from means based on the full PIRLS dataset.

Countries with the largest PIRLS/ePIRLS differences were Slovenia (an 18-point advantage on the paper-based test) and the United Arab Emirates (an 18-point advantage on the digital test). Comparing national means on the PIRLS Informational subscale and on ePIRLS, the largest differences were found in Chinese Taipei, Slovenia, and Italy. In all three, mean performance on ePIRLS was considerably poorer than on PIRLS (23 points lower in Chinese Taipei, 19 points in Slovenia, and 17 points in Italy).

Figure 3.1 presents the relationship between mean scores on the three measures graphically (ordered from left to right by PIRLS mean score). There is a close match between performances on each across countries, and high achievement on one measure tends to be matched by high achievement on the others. For example, Singapore has the highest mean scores across all three

measures. Ireland and Norway are also very high-performing countries on all three. At the other extreme, Georgia and the United Arab Emirates are among the lower-performing countries on all three. Only Chinese Taipei, Italy and Slovenia show some limited variation between performance across measures.

Figure 3.1: Mean scores on ePIRLS, PIRLS, and PIRLS Informational scale, all participating ePIRLS countries



PIRLS data are based only on the subset of pupils who also completed ePIRLS.

Gender differences

In Portugal and Macao SAR there were no significant gender differences on overall PIRLS scores. In all other participating countries, there was a significant and often large difference in favour of girls. The PIRLS average gap was 19 points (Table 3.4), and ranged from as low as one point in Portugal and Macao SAR to 65 points in Saudi Arabia. In Ireland the gap was 12 points. Generally across PIRLS participants, the largest gender differences were found in Middle Eastern countries – of the nine countries where the average gender gap was 30 points or more, South Africa was the only country not located in the Middle East. However, Table 3.4 shows that large gender gaps were also found in some of the comparison countries, notably New Zealand, Australia, Finland (all 22 points) and Norway (21 points).

Table 3.4: Gender differences in mean performance on PIRLS 2016, Ireland and comparison countries, in increasing size of gender gap			
	Girls	Boys	Gender gap
United States	553	545	**8**
Hong Kong SAR	573	564	**9**
Ireland	572	561	**12**
Russian Federation	588	574	**15**
England	566	551	**15**
Singapore	585	568	**17**
Poland	574	556	**18**
Northern Ireland	574	555	**18**
PIRLS international average	*520*	*501*	***19***
Norway	570	548	**21**
Australia	555	534	**22**
Finland	577	555	**22**
New Zealand	533	512	**22**

All gender gaps shown are statistically significant. Because of rounding, some results may appear inconsistent.

In ePIRLS, gender differences on the test were slightly less pronounced, averaging a 12-point advantage for girls across all countries (Table 3.5). In Ireland, girls outperformed boys by 11 points. Again, Portugal was one of only three countries (along with Italy and Denmark) where significant gender differences were not found. In all other countries, the gender gap in performance on ePIRLS was significant, and more pronounced in the United Arab Emirates and in Singapore than in most countries.

Table 3.5: Gender differences in mean performance on ePIRLS 2016, all participating countries in increasing size of gender gap			
	Girls	Boys	Gender gap
Italy	534	531	2
Portugal	524	521	3
Denmark	560	556	4
United States	560	554	**6**
Canada	547	539	**8**
Chinese Taipei	551	541	**9**
Ireland	572	561	**11**
Israel	542	530	**11**
ePIRLS international average	*545*	*533*	***12***
Slovenia	532	518	**14**
Sweden	567	552	**15**
Georgia	485	469	**15**
Norway (Grade 5)	576	558	**18**
Singapore	599	578	**21**
United Arab Emirates	483	454	**29**

All gender gaps shown in bold are significant. Because of rounding, some results may appear inconsistent.

Trends in achievement

As well as providing cross-country comparisons of performance within any given cycle, PIRLS is specifically designed to measure trends in reading achievement. That is why the centrepoint of 500 from the first cycle is still used as the scale anchor, even though the overall mean score for the 2016 cycle of PIRLS is above 500.

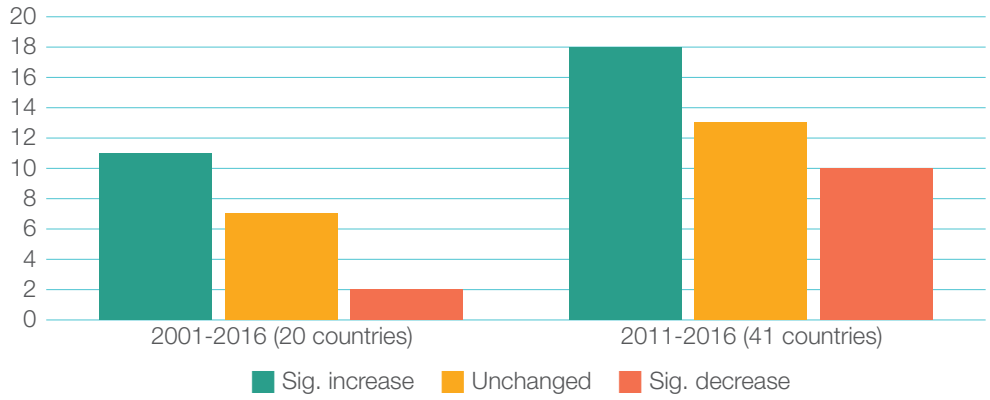
Measuring Trends

PIRLS uses a statistical technique called concurrent calibration to measure trends from cycle to cycle. Put simply, as each PIRLS cycle shares a good deal of test content with the previous cycle, common data from countries taking part in both cycles is used as part of the overall scaling process. Linear transformations are then applied, to place the results from each successive assessment on the same scale as the results from the previous assessment. Each new cycle is linked with the preceding one, (e.g., 2016 with 2011, 2011 with 2006, 2006 with 2001), making it possible to establish reliable long-term trends across all cycles of PIRLS.

This method provides a very stable measure of change over time, and has two main advantages. First, it allows us to track changes in the overall study mean, even though the countries taking part in each cycle may vary. Second, each country can track their own performance over time, even though there are different countries in each cycle. Thus, Ireland's mean in 2016 can be directly compared to that in 2011, even though different countries took part in the two cycles. Any large changes in a country's mean from one cycle to the next are more likely to be attributable to actual change than to linking or measurement error.

Twenty countries took part in both in 2001 and 2016. In these, average achievement improved in 11 countries, remained unchanged in seven, and dropped in two (Figure 3.2). Short-term trends were more mixed. Comparing the 41 countries that took part in PIRLS 2011 and 2016, mean achievement significantly improved in 18, significantly dropped in 10, and did not change significantly in 13 countries.

Figure 3.2: National trends in mean performance on PIRLS



As noted in Chapter 1, Ireland participated in PIRLS for the first time in 2011. With the exception of Poland, all our comparison countries also have data from 2011, allowing trends in achievement to be monitored. Table 3.6 shows mean reading scores in PIRLS 2011 and 2016. Of the comparison countries, the largest improvements in achievement are noted in Australia (a 17-point increase), Ireland (15 points), the Russian Federation (12 points), and Norway (10 points). In contrast, statistically significant drops in achievement were recorded in New Zealand (8-point drop) and the United States (7-point drop). Across all participating countries, Oman (28 points) and Lithuania (22 points) showed the largest increases, while Iran (29 points) and Portugal (13 points) showed the largest drops in mean scores.

Table 3.6: Mean PIRLS scores 2011 and 2016 (Ireland and comparison countries)

	2016	2011	Change
Russian Federation	581	568	**+12**
Singapore	576	567	+9
Hong Kong SAR	569	571	-2
Ireland	567	552	**+15**
Finland	566	568	-2
Northern Ireland	565	558	+6
Norway	517	507	**+10**
England	559	552	**+7**
United States	549	556	**-7**
Australia	544	527	**+17**
New Zealand	523	531	**-8**

Bold indicates that difference between the two cycles is statistically significant. Because of rounding, some results may appear inconsistent.

Generally, national-level increases or decreases in overall mean scores were reflected in similarly-sized changes in mean scores by gender. For example, the mean score obtained by Irish girls increased by 13 points since 2011, while the score for boys increased by 17 points (Table 3.7). In New Zealand, girls averaged eight points lower than in 2011, while boys averaged nine points lower. In a few countries, however, gains or drops were more pronounced for one gender. For

example, in Hong Kong, the mean score obtained by girls fell by six points while that obtained by boys increased by one point. In England, the increase in the overall national mean is largely attributable to improved performance by boys (an increase of 11 points, vs three points for girls).

Table 3.7: Mean PIRLS scores 2011 and 2016 (Ireland and comparison countries[1])

	Girls			Boys		
	2011	2016	Change	2011	2016	Change
Russian Federation	578	588	10	559	574	15
Singapore	576	585	9	559	568	9
Hong Kong SAR	579	573	-6	563	564	1
Ireland	559	572	13	544	561	17
Finland	578	577	-1	558	555	-3
Northern Ireland	567	574	7	550	555	5
England	563	566	3	540	551	11
United States	562	553	-9	551	545	-6
Australia	536	555	19	519	534	15
New Zealand	541	533	-8	521	512	-9

[1] Norway not shown as the target grade for the assessment changed between the 2011 and 2016 cycles. Changes are approximate as they are calculated from rounded mean scores.

Across PIRLS as a whole, differences in trends by gender were most pronounced in Portugal and South Africa. In the former, the mean score for girls dropped roughly 19 points between 2011 and 2016, while the mean for boys dropped by 7 points. Thus, while Portugal had a 14-point (statistically significant) gender gap in 2011, in PIRLS 2016 it was only a 2-point gap, making Portugal one of only two countries where the gender gap was not significant. In South Africa, the mean score for girls increased by about six points while the mean for boys decreased by approximately 12 points, meaning that the size of the gender gap in reading achievement in South Africa (52 points) is now only exceeded by the size of the gap in Saudi Arabia (65 points in favour of girls).

Chapter 4:
Performance at International Benchmarks

Chapter 4:
Performance at International Benchmarks

This chapter outlines the International Benchmarks used to describe the skills and strategies demonstrated by pupils at various levels of achievement. Separate descriptors are provided for PIRLS and ePIRLS. However, the two are directly comparable as the same cutpoints are used on both to categorise achievement levels as Low, Intermediate, High or Advanced. Examples of texts and questions which illustrate the difficulty of items at the various Benchmarks are provided in the main international reports (Mullis et al., 2017a, 2017b) and on the International Study Center website (http://timssandpirls.bc.edu/).

The first section of this chapter describes the PIRLS Benchmarks and the percentages of pupils in Ireland and in comparison countries reaching each of the Benchmarks. The second section provides the corresponding information for ePIRLS.

PIRLS Benchmarks

Table 4.1 describes the reading skills and strategies that pupils demonstrate at each of the International Benchmarks. As the skills required to navigate a text vary by text purpose, separate descriptors are provided for literary and informational texts. Also, there is a progression in the skills and strategies shown by pupils as Benchmarks move from Low through to Advanced, with pupils at higher levels of achievement also able to demonstrate the skills associated with lower levels.

For example, the skills of pupils at the Low International Benchmark are relatively limited – e.g., they can retrieve and reproduce a piece of explicitly stated information from a text – whereas pupils at the Intermediate Benchmark can retrieve and reproduce two or three pieces of explicitly stated information and make straightforward inferences about feelings and motivations of main characters. Pupils at the High International Benchmark demonstrate all the skills of pupils at the two lower Benchmarks, but can also identify significant actions and details embedded across the text, integrate textual and visual information, and recognise some language features (such as metaphor and tone). Pupils at the Advanced International Benchmark are able to take texts as a whole as evidence for their interpretations and explanations, can consider how the author's point of view may be expressed, and can evaluate visual and textual elements.

Table 4.1: Summary descriptions of pupils' reading skills at each International Benchmark, by text purpose		
Level	**Literary Text:** *Pupils can…*	**Informational Text:** *Pupils can…*
Advanced *(625 points)*	■ interpret story events and character actions to describe reasons, motivations, feelings, and character development with full text-based support. ■ begin to evaluate the effect on the reader of the author's language and style choices.	■ distinguish and interpret complex information from different parts of text, and provide full text-based support. ■ integrate information across a text to explain relationships and sequence activities. ■ begin to evaluate visual and textual elements to consider the author's point of view.
High *(550 points)*	■ locate and distinguish significant actions and details embedded across the text. ■ make inferences to explain relationships between intentions, actions, events, and feelings, and give text-based support. ■ interpret and integrate story events and character actions, traits, and feelings as they develop across the text. ■ recognise the use of some language features (e.g. metaphor, tone, imagery).	■ locate and distinguish relevant information within a dense text or a complex table. ■ make inferences about logical connections to provide explanations and reasons. ■ integrate textual and visual information to interpret the relationship between ideas. ■ evaluate and make generalisations about context and textual elements.
Intermediate *(475 points)*	■ independently locate, recognise, and reproduce explicitly stated actions, events, and feelings. ■ make straightforward inferences about the attributes, feelings, and motivations of main characters. ■ interpret obvious reasons and causes, 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)*	■ locate and retrieve explicitly stated information, actions, or ideas. ■ make straightforward inferences about events and reasons for actions. ■ begin to interpret story events and central ideas.	■ locate and reproduce explicitly stated information from text and other formats (e.g., charts, diagrams). ■ 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 4.2 shows the percentages of pupils reaching each of the four Benchmarks in Ireland and in the selected comparison countries, ordered by decreasing overall mean score on PIRLS. Almost all pupils in Ireland (98%) and in the comparison countries demonstrated the skills associated with the lowest Benchmark. This means that virtually all pupils in these countries had successfully mastered the most basic reading skills. Notable exceptions were New Zealand and Australia, where 10% and 6% of pupils, respectively, failed to reach the Low Benchmark. A very large majority of pupils in Ireland (89%) were also able to demonstrate the skills associated with the Intermediate Benchmark, while 62% reached the High Benchmark. These percentages are similar to the percentages in Singapore, Hong Kong, Finland, Poland and Northern Ireland, but somewhat lower than those in the Russian Federation.

Table 4.2: Percentages (SE) of pupils reaching the 2016 International Benchmarks (PIRLS), Ireland and comparison countries

		Comprehension process progression →			
		Low (400 points)	Intermediate (475 points)	High (550 points)	Advanced (625 points)
Russian Fed	%	99	94	70	26
	SE	*0.3*	*0.6*	*1.3*	*1.2*
Singapore	%	97	89	66	29
	SE	*0.5*	*1.0*	*1.6*	*1.6*
Hong Kong SAR	%	99	93	65	18
	SE	*0.3*	*0.9*	*1.8*	*1.3*
Ireland	%	98	89	62	21
	SE	*0.4*	*0.9*	*1.6*	*1.2*
Finland	%	98	91	62	18
	SE	*0.3*	*0.8*	*1.3*	*0.8*
Poland	%	98	89	61	20
	SE	*0.4*	*0.7*	*1.3*	*1.1*
Northern Ireland	%	97	87	61	22
	SE	*0.4*	*0.8*	*1.3*	*1.4*
Norway	%	99	90	58	15
	SE	*0.3*	*0.9*	*1.7*	*0.9*
England	%	97	86	57	20
	SE	*0.4*	*0.7*	*1.1*	*0.9*
United States	%	96	83	53	16
	SE	*0.5*	*1.2*	*1.6*	*1.3*
Australia	%	94	81	51	16
	SE	*0.5*	*1.0*	*1.4*	*1.0*
New Zealand	%	90	73	41	11
	SE	*0.7*	*1.0*	*1.2*	*0.6*
International median[1]		96	82	47	10

[1] Values for the international median divide countries evenly. For example, the median value for Advanced is 10%. Thus, in half of participating countries more than 10% of pupils reached this Benchmark, and in half, fewer than 10% did so.

Figure 4.1 shows countries ordered left to right by decreasing percentages reaching the Advanced International Benchmark. For each country, the diamond point shows the percentage value, while the error bars provide a 95 percent confidence interval around that percentage. If there is no overlap in two countries' intervals, it means that they differ significantly. Conversely, where intervals overlap, it usually – but not always – means that the countries do not differ significantly.[1]

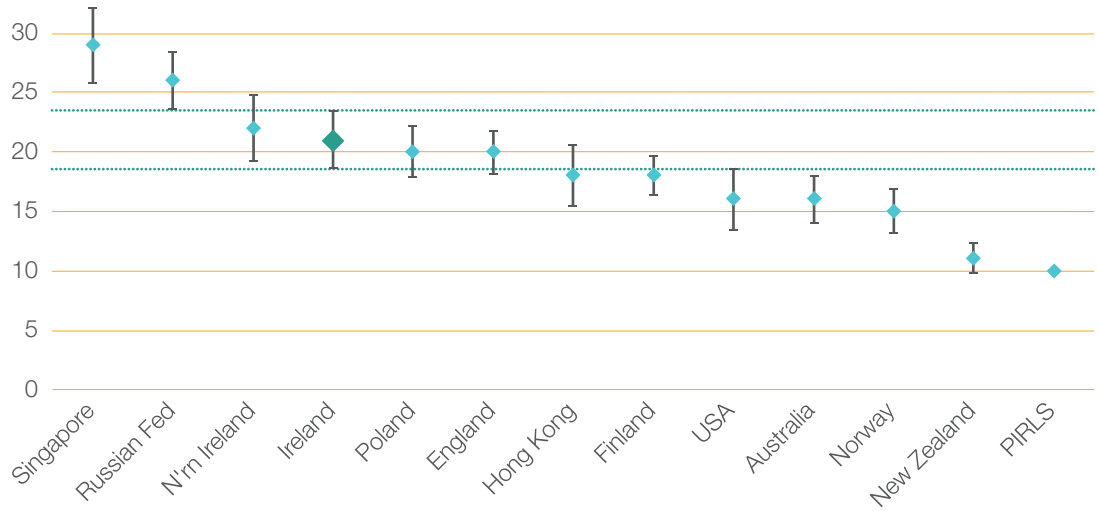
In Russia and Singapore, over one-quarter of pupils reached the Advanced Benchmark. Similar to Poland (20%), Northern Ireland (22%) and England (20%), about one-fifth of pupils in Ireland (21%) reached the Advanced Benchmark. Somewhat lower percentages of pupils reached the Advanced Benchmark in the USA, Australia, Norway, and New Zealand.

1 Confidence intervals provide a simple way to get a good indicator of statistical significance. However, tests of significance for the data in Figure 4.1 were actually determined on the basis of the square root of the sum of squares of the standard errors for the country percentages.

Some countries may have a high mean score but a comparatively lower percentage of very high achievers. For example, the mean scores of pupils in Norway and England were more or less identical (559), yet Norway had a lower percentage of pupils (15%) reaching the Advanced Benchmark than did England (20%). On the other hand, the mean score in Norway was significantly higher than that in the United States (549) and Australia (544), yet the percentages of pupils reaching the Advanced Benchmark in the three countries are very similar (15% in Norway and 16% in the USA and Australia). These examples show the need to analyse not only national mean scores but also the distribution of achievement within a country.

The pair of dotted green lines in Figure 4.1 are aligned to the 95 percent confidence interval around the percentage of pupils in Ireland reaching the Advanced Benchmark. The percentage ranges for both the Russian Federation and Singapore are above the top green line. This means that the percentages of pupils reaching the Advanced Benchmark in the Russian Federation and Singapore are significantly higher than in Ireland. The ranges for Northern Ireland, Poland, England and Hong Kong fall partially within the green lines, and the percentages reaching the Advanced Benchmark in these countries do not differ significantly from the percentage in Ireland. There is also some overlap between Finland's and Ireland's range. Nonetheless, significance tests indicate that Finland has a significantly lower percentage of pupils at the Advanced International Benchmark. For all other countries shown (e.g., USA, Australia) the percentage ranges fall on or below the lower green line, meaning the percentages reaching the Advanced Benchmark in these countries are significantly lower than the percentage of pupils in Ireland.

Figure 4.1: Percentages (and related error margins) of pupils reaching the Advanced International Benchmark (PIRLS), Ireland and comparison countries, in descending order



Benchmark Trends

Table 4.3 compares the percentages of pupils in Ireland reaching each of the PIRLS International Benchmarks in 2011 and 2016. The percentages reaching the Intermediate, High and Advanced Benchmarks in 2016 are all significantly higher than the corresponding percentages in 2011. There was also an increase in the percentage of pupils reaching the Low International Benchmark, but the increase was not statistically significant. The increases in Ireland mirrored

the overall international trend which saw increases in the percentages at the Intermediate, High and Advanced Benchmarks, but only a marginal increase in the median percentage reading the Low Benchmark (from 95% to 96%). The increase in the median percentage reaching the Intermediate Benchmark was most pronounced (from 69% to 82%), but there were also increases in the median percentages for the High (44% to 47%) and Advanced (8% to 10%) Benchmarks.

Table 4.3: Percentages (SE) of pupils in Ireland reaching each of the International Benchmarks, 2011 and 2016

	Low	Intermediate	High	Advanced
2011	97 (0.5)	85 (0.8)	53 (1.4)	16 (0.9)
2016	98 (0.4)	**89** (0.9)	**62** (1.6)	**21** (1.2)

Bold indicates that the difference between the two cycles is statistically significant.

ePIRLS Benchmarks

Benchmark descriptors are also provided for ePIRLS, outlining the skills and strategies used by pupils when reading and viewing online informational texts (Table 4.4). As with PIRLS, there is progression through the levels, with pupils at higher levels able to demonstrate the skills associated with lower levels. For example, pupils reaching only the Low International Benchmark can reproduce explicitly stated information from web pages, whereas pupils reaching the Advanced International Benchmark can also integrate information across multiple web pages, and make inferences from relatively complex information to provide support for their responses to questions. Unlike PIRLS, there are no trend data available for ePIRLS Benchmarks as 2016 was the first time ePIRLS was administered.

Table 4.4: Summary descriptions of the skills that pupils display in Online Informational Reading at each of the International Benchmarks

Level	When reading and viewing *Online Informational Texts*, pupils can…
Advanced *(625 points)*	■ make inferences from complex information to support an explanation. ■ interpret and integrate information from within and across web pages with interactive features 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)*	■ make inferences to distinguish relevant information and provide comparisons. ■ interpret and integrate information within and across web pages with interactive features to provide examples and make contrasts. ■ evaluate how graphic elements and language choices support content.
Intermediate *(475 points)*	■ locate and reproduce information presented in various forms, including independent use of navigation features. ■ make straightforward inferences to recognise reasons and actions. ■ interpret and integrate information across a web page to recognise causes, comparisons, and explanations. ■ begin to evaluate the use of interactive features to convey information.
Low *(400 points)*	■ locate and reproduce explicitly stated information from web pages 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 Mullis et al., 2017b.

In 11 of the 14 countries (including Ireland) that took part in ePIRLS, at least 95% of pupils successfully demonstrated the skills associated with the lowest ePIRLS Benchmark (Table 4.5). At least 90% of pupils in Ireland, Norway and Singapore were also able to demonstrate the skills associated with the Intermediate Benchmark (ePIRLS average: 84%). The percentages of pupils reaching the ePIRLS High Benchmark varied considerably across countries. In Georgia and the United Arab Emirates, just 16% and 22% of pupils respectively reached this Benchmark compared to 63% in Ireland and Norway and 72% in Singapore.

Table 4.5: Percentages (SE) of pupils reaching the 2016 International Benchmarks (ePIRLS), Ireland and comparison countries

		Comprehension process progression →			
		Low (400 points)	Intermediate (475 points)	High (550 points)	Advanced (625 points)
Singapore	%	98	92	72	34
	SE	*0.4*	*0.9*	*1.5*	*1.7*
Norway (Grade 5)	%	99	92	63	18
	SE	*0.3*	*0.8*	*1.6*	*1.2*
Ireland	%	98	90	63	20
	SE	*0.4*	*0.9*	*1.6*	*1.3*
Sweden	%	98	89	59	14
	SE	*0.3*	*0.8*	*1.7*	*1.0*
Denmark	%	98	89	57	15
	SE	*0.4*	*0.9*	*1.4*	*1.0*
United States	%	97	86	56	18
	SE	*0.4*	*1.0*	*1.5*	*1.2*
Chinese Taipei	%	97	86	51	10
	SE	*0.3*	*0.9*	*1.3*	*0.7*
Canada	%	96	82	49	12
	SE	*0.5*	*1.3*	*2.1*	*1.0*
Israel	%	93	78	47	13
	SE	*0.7*	*1.1*	*1.3*	*1.0*
Italy	%	98	82	41	6
	SE	*0.5*	*1.3*	*1.3*	*0.7*
Slovenia	%	95	78	39	5
	SE	*0.6*	*1.0*	*1.3*	*0.5*
Portugal	%	97	77	35	5
	SE	*0.5*	*1.3*	*1.4*	*0.6*
Georgia	%	85	54	16	1
	SE	*1.4*	*2.1*	*1.3*	*0.4*
UAE	%	75	50	22	5
	SE	*0.9*	*1.1*	*0.8*	*0.3*
International median[1]		97	84	50	12

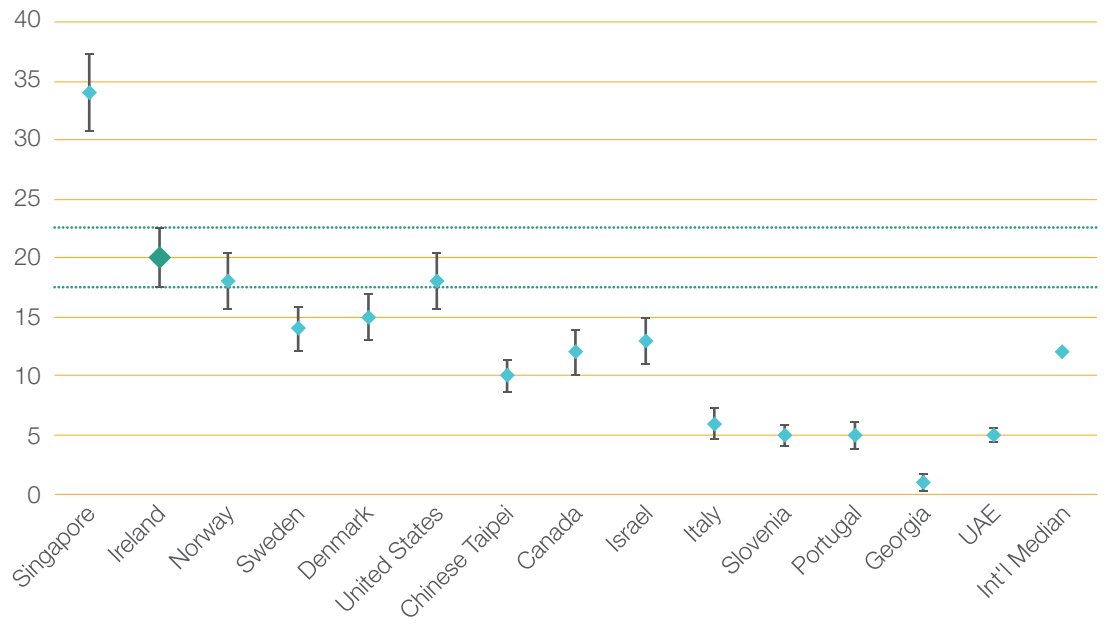
[1] Half the countries have a percentage above the median and half have a percentage below; i.e., values for the international median divide the countries evenly. For example, Advanced has a median value of 12%. Thus, in half of participating countries more than 12% of pupils reached this Benchmark, and, in the other half, fewer than 12% did so.

Figure 4.2 shows the percentages of pupils in each country reaching the Advanced Benchmark in ePIRLS, sorted left to right in descending order. At 34%, Singapore had the highest percentage of pupils demonstrating the most advanced skills, considerably higher than in the next highest

country (Ireland: 20%). The green lines around Ireland's percentage indicate the confidence interval around the percentage value. In Norway and the United States, 18% of pupils reached the Advanced Benchmark. There is considerable overlap in the confidence intervals for Ireland, Norway, and the US, and the percentages of pupils at the Advanced International Benchmark in the three countries do not differ significantly. In contrast, all remaining countries fall below the lower green line, meaning that significantly fewer pupils reached the Advanced Benchmark than in Ireland.

Five countries (Sweden, Denmark, Chinese Taipei, Canada and Israel) had 10 to 15% of pupils at the Advanced level, while four (Italy, Slovenia, Portugal and the United Arab Emirates) had approximately 5% of pupils at this level. In contrast, while Georgia performed significantly better than the United Arab Emirates on ePIRLS, overall, only 1% of pupils in Georgia reached the Advanced level, compared to 5% in the United Arab Emirates.

Figure 4.2: Percentages (and related error margins) of pupils reaching the Advanced International Benchmark (ePIRLS)



Chapter 5:

Aspects of reading performance

Chapter 5:
Aspects of reading performance

The PIRLS framework examines reading using the *purposes* for which children read and the comprehension *processes* they use when reading. The purposes for which children read are reflected in the use of Literary and Informational texts in PIRLS, and subscales reflecting these purposes. In contrast, as ePIRLS contained only Informational texts, there are no reading purpose subscales for ePIRLS. As noted in Chapter 1, the framework also identifies 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*. These processes form the basis of two reading comprehension subscales: Retrieving and Straightforward Inferencing (hereafter referred to as Retrieve/Infer); and, Interpreting, Integrating, and Evaluating (hereafter referred to as Interpret/Evaluate).

This chapter examines performance on reading purpose and reading process subscales, as well as gender differences on the subscales. Trends over time on subscales are also examined (PIRLS only).

Performance on purpose and process subscales

In this section, Irish performance on the reading purpose and comprehension process subscales is compared to that in the selected comparison countries. In the tables that follow, countries are presented in order of decreasing overall mean score.

PIRLS subscales

Subscale scores generally reflected national mean scores, but many countries showed relative strengths or weaknesses on one or more of the process or purpose subscales (Table 5.1). For example, pupils in Ireland showed a relative strength on the Literary subscale, achieving a mean score of 571, which was significantly higher than the overall Irish PIRLS mean of 567. In contrast, the mean score on the Informational subscale did not differ significantly from the overall Irish mean score on PIRLS. Other countries demonstrating a relative strength on the Literary subscale were Northern Ireland, England, the United States, Australia and New Zealand, all anglophone countries. In Northern Ireland, England, the United States and New Zealand, performance on the Informational subscale was significantly lower than the national

PIRLS mean, whereas pupils in the Russian Federation, Hong Kong, Singapore and Finland showed relative strengths on the Informational subscale.

Pupils in Ireland scored slightly higher on the Interpret/Evaluate subscale than on Retrieve/Infer, but not significantly higher than the national PIRLS mean. An advantage on the Interpret/Evaluate subscale was found in four of the five English-speaking comparison countries (England, the United States, Australia and New Zealand) and in Poland and Singapore. Finland was the only comparison country where pupils demonstrated a relative strength on the Retrieve/Infer subscale.

Table 5.1: Mean scores on PIRLS reading purpose and comprehension process subscales, Ireland and comparison countries

		Purpose		Process	
	PIRLS Mean	Literary	Informational	Retrieve/Infer	Interpret/Evaluate
Russian Federation	581	579	**584**	581	582
Singapore	576	575	**579**	**573**	**579**
Hong Kong SAR	569	**562**	**576**	568	568
Ireland	567	**571**	565	566	569
Finland	566	565	**569**	**572**	**562**
Poland	565	567	564	**560**	**570**
Northern Ireland	565	**570**	**561**	**562**	567
Norway	559	560	559	561	558
England	559	**563**	**556**	**556**	**561**
United States	549	**557**	**543**	**543**	**555**
Australia	544	**547**	543	**541**	**549**
New Zealand	523	**525**	**520**	521	**525**

Bold indicates that subscale mean is significantly different to overall PIRLS mean score.

Gender differences

In Ireland, girls significantly outperformed boys on both reading purpose subscales, although the gender gaps were smaller than the PIRLS international averages (Table 5.2). The gender gap in Ireland was smaller on the Informational than on the Literary subscale (eight points versus 17 points, respectively). This reflected a general pattern in PIRLS, as the international average gap on the Literary subscale was 23 points, compared to 16 points on the Informational subscale. Amongst the comparison countries shown in Table 5.2, girls significantly outperformed boys on the Literary subscales in all countries, and in all but two countries (Hong Kong and the United States) on the Informational subscale.

Table 5.2: Mean scores on PIRLS reading purpose subscales, by gender (Ireland and comparison countries)

	Literary			Informational		
	Girl	Boy	Gender gap	Girl	Boy	Gender gap
Russian Federation	587	572	**15**	591	578	**13**
Singapore	586	563	**23**	586	571	**15**
Hong Kong SAR	569	557	**12**	580	573	7
Ireland	580	563	**17**	569	561	**8**
Finland	576	554	**22**	579	559	**20**
Poland	577	556	**21**	573	556	**17**
Northern Ireland	582	559	**23**	569	552	**17**
Norway	571	550	**21**	568	549	**19**
England	572	553	**19**	562	551	**11**
United States	563	552	**11**	546	540	6
Australia	561	533	**28**	552	533	**19**
New Zealand	539	512	**27**	528	512	**16**
International Average	522	499	**23**	519	503	**16**

All statistically significant gender gaps are in bold. Gender gaps are approximate as they are calculated from rounded mean scores.

As was the case in most comparison countries, Irish gender differences on the two comprehension process subscales were roughly similar (Table 5.3). Again, these differences were smaller in Ireland than on average internationally. For example, whereas girls internationally outperformed boys on the Retrieve/Infer subscale by an average of 17 points, in Ireland the gender gap was 10 points.

Table 5.3: Mean scores on PIRLS reading process subscales, by gender (Ireland and comparison countries)

	Retrieve/Infer			Interpret/Evaluate		
	Girl	Boy	Gender gap	Girl	Boy	Gender gap
Russian Federation	588	575	**13**	589	575	**14**
Singapore	580	566	**14**	589	568	**21**
Hong Kong SAR	571	565	**6**	574	563	**11**
Ireland	571	561	**10**	576	562	**14**
Finland	582	562	**20**	573	552	**21**
Poland	568	551	**17**	580	559	**21**
Northern Ireland	570	553	**17**	577	558	**19**
Norway	570	553	**17**	568	548	**20**
England	563	549	**14**	569	554	**15**
United States	547	539	**8**	559	551	**8**
Australia	552	530	**22**	561	538	**23**
New Zealand	530	512	**18**	536	513	**23**
International Average	520	503	**17**	520	500	**20**

All differences are statistically significant in favour of girls. Gender gaps are approximate as they are calculated from rounded mean scores.

ePIRLS subscales

As ePIRLS was composed solely of Informational texts, only comprehension process subscales were developed for ePIRLS (Table 5.4). Pupils in Ireland did not demonstrate relative strengths or weaknesses across the comprehension process subscales. In contrast, pupils in Singapore, Chinese Taipei, Portugal, Georgia and the United Arab Emirates demonstrated a relative strength on the Retrieve/Infer subscale. The United States was the only country to demonstrate a relative strength on the Interpret/Evaluate subscale.

Table 5.4: Mean scores on ePIRLS reading comprehension process subscales

	ePIRLS Mean	Retrieve/Infer	Interpret/evaluate
Singapore	588	**594**	**585**
Norway (Grade 5)	568	567	568
Ireland	567	566	568
Sweden	559	561	559
Denmark	558	560	556
United States	557	**553**	**560**
Chinese Taipei	546	**548**	**544**
Canada	543	**541**	545
Israel	536	536	535
Italy	532	534	531
Slovenia	525	525	523
Portugal	522	**525**	**521**
Georgia	477	**485**	**466**
United Arab Emirates	468	**471**	**465**

Bold indicates that subscale mean is significantly different to overall ePIRLS mean score.

Gender differences

Although there were some gender differences on the ePIRLS comprehension process subscales, they tended to be somewhat less pronounced than was the case for PIRLS – particularly for the Interpret/Evaluate subscale (Table 5.5). For example, in five of the 14 ePIRLS countries (Denmark, the United States, Canada, Italy and Portugal) there were no significant gender differences on the Interpret/Evaluate subscale. In Ireland, girls significantly outperformed boys on both comprehension process subscales and the gender gaps on the two scales are about the same as the corresponding international averages. The advantage for girls was 13 points on the Retrieve/Infer subscale (international average: 15 points) and 10 points on the Interpret/Evaluate subscale (international average: 9 points).

Table 5.5: Mean scores on ePIRLS reading comprehension process subscales, by gender

	Retrieve/Infer			Interpret/Evaluate		
	Girl	Boy	Gender gap	Girl	Boy	Gender gap
Singapore	606	583	**23**	594	575	**19**
Norway (Grade 5)	578	557	**21**	575	560	**15**
Ireland	572	559	**13**	573	563	**10**
Sweden	570	551	**19**	565	553	**12**
Denmark	565	555	**10**	556	556	0
United States	558	548	**10**	562	557	5
Chinese Taipei	555	542	**13**	548	540	**8**
Canada	547	534	**13**	547	543	4
Israel	544	528	**16**	539	531	**8**
Italy	537	531	**6**	530	531	-1
Slovenia	535	516	**19**	528	519	**9**
Portugal	528	522	6	521	521	0
Georgia	495	475	**20**	471	461	**10**
United Arab Emirates	487	455	**32**	478	453	**25**
International Average	548	533	**15**	542	533	**9**

Differences in bold are statistically significant in favour of girls. Gender gaps are approximate as they are calculated from rounded mean scores.

Subscale trends

As noted, significant improvements in overall achievement on PIRLS since 2011 were found in Ireland, and (among comparison countries) Australia, the Russian Federation, and Norway. In contrast, statistically significant drops in achievement were recorded in New Zealand and the United States. Tables 5.6 and 5.7 examine changes on the reading purpose and process subscales over the same period.

In New Zealand there was a similar-sized decline on both reading purpose subscales, whereas the drop in performance on Interpret/Evaluate was more pronounced than for Retrieve/Infer (an 11-point drop versus a 6-point drop, respectively). In the United States, the drop in achievement since PIRLS 2011 seems to be attributable to drops in performance on Informational texts and on the Interpret/Evaluate subscale. Among countries where significant increases occurred since 2011, much of Australia's improvement can be attributed to improvements on the Literary and Interpret/Evaluate subscales. In contrast, Ireland's improved overall mean score is attributable to relatively even gains on all subscales.

Table 5.6: Mean scores on PIRLS reading purpose subscales, 2016 and 2011 (Ireland and comparison countries)

	Literary			Informational		
	2016	2011	Change	2016	2011	Change
Russian Federation	579	567	**+12**	584	570	**+15**
Singapore	575	567	+8	579	569	**+9**
Hong Kong SAR	562	565	-2	576	578	-1
Ireland	571	557	**+14**	565	549	**+16**
Finland	565	568	-4	569	568	+1
Northern Ireland	570	564	+7	561	555	+6
Norway	520	508	**+13**	514	505	**+9**
England	563	553	**+10**	556	549	**+7**
United States	557	563	-5	543	553	**-10**
Australia	547	527	**+20**	543	528	**+15**
New Zealand	525	533	**-8**	520	530	**-9**

Bold indicates that difference between the two cycles is statistically significant. Because of rounding, some results may appear inconsistent.

Table 5.7: Mean scores on PIRLS comprehension process subscales, 2016 and 2011 (Ireland and comparison countries)

	Retrieve/Infer			Interpret/Evaluate		
	2016	2011	Change	2016	2011	Change
Russian Federation	581	565	**+16**	582	571	**+11**
Singapore	573	565	+8	579	570	+9
Hong Kong SAR	568	562	+5	568	578	**-9**
Ireland	566	552	**+14**	569	553	**+16**
Finland	572	569	+3	562	567	-5
Northern Ireland	562	555	+6	567	562	+5
Norway	521	511	**+10**	513	502	**+11**
England	556	546	**+10**	561	555	+6
United States	543	549	-6	555	563	**-8**
Australia	541	527	**+14**	549	529	**+20**
New Zealand	521	527	**-6**	525	535	**-11**

Bold indicates that difference between the two cycles is statistically significant. Because of rounding, some results may appear inconsistent.

Chapter 6:
Summary

Chapter 6:
Summary

This chapter provides an overview of the performance of Irish pupils in PIRLS and ePIRLS 2016. It also outlines subsequent Irish analyses that will be published later in 2018, after the full and final international dataset is made available.

Performance on PIRLS

Ireland was one of 50 countries/regions and 11 benchmarking participants taking part in PIRLS 2016. With a mean score of 567, Irish pupils performed extremely well on PIRLS. The Irish mean was well above the study centrepoint of 500, and significantly higher than that of 43 participating countries. Just two countries (Russia and Singapore) had significantly higher mean scores than Ireland. Ireland's mean score in 2016 was 15 points higher than when PIRLS last took place (in 2011), an increase which is both sizeable and statistically significant.

In all but two participating countries (Macao and Portugal), girls had significantly higher levels of achievement than boys. In Ireland, girls had a 12-point advantage over boys, somewhat smaller than the corresponding international average of 19 points.

As well as an overall score, PIRLS provides subscale scores for reading purposes (reading for Literary or Informational purposes) and reading comprehension processes (Retrieve/Infer and Interpret/Evaluate). Irish pupils demonstrated a particular strength on Literary types of texts. This can partly be attributed to the very high performance of Irish girls on Literary texts (mean score of 580). In contrast, boys in Ireland performed well above average on both Literary and Informational types of texts but did not show a particular strength or weakness for either text type.

PIRLS data can also be analysed in relation to the percentages of pupils in a country who reach various International Benchmarks (i.e., can demonstrate a range of increasingly complex skills). Almost all Irish pupils (98%) were able to demonstrate the skills associated with the lowest level of reading achievement described in PIRLS (Low International Benchmark), compared to an international median of 96%. A large majority of Irish pupils (89%) also reached the Intermediate Benchmark (international median: 82%), and almost two-thirds (62%) reached the High Benchmark (international median: 47%). The Advanced International Benchmark was reached by 21% of pupils in Ireland, which was significantly higher than both the international median for 2016 (10%) and the corresponding figure for Irish pupils in 2011 (16%).

Performance on ePIRLS

Ireland was one of 14 countries and two benchmarking participants that took part in ePIRLS – a series of informational style "projects" with which pupils engaged via a simulated online environment. With a mean score of 567, Irish pupils performed as well on the digital ePIRLS assessment as they did on the paper-based PIRLS assessment. Only pupils in Singapore achieved a significantly higher mean score on ePIRLS than pupils in Ireland. Ireland's mean score on ePIRLS was more or less identical to the ePIRLS mean for Norway and significantly higher than that of all remaining participating countries and regions.

Girls performed significantly better than boys in 11 of the 14 participating countries, including in Ireland. However, at 12 points, the international gender gap was slightly smaller than on the paper-and-pencil PIRLS assessment (in Ireland, the gender gap on ePIRLS was 11 points). No significant gender differences on ePIRLS were found in Italy, Portugal and Denmark.

In Ireland, the percentages of pupils reaching each of the ePIRLS International Benchmarks were very similar to the percentages reaching the PIRLS International Benchmarks. Thus, almost all (98%) pupils in Ireland could demonstrate the skills associated with reaching the Low Benchmark (international median: 97%), while 90% reached the Intermediate Benchmark (international median: 84%). Almost two-thirds (63%) reached the High Benchmark (international median: 50%) and 20% demonstrated the skills associated with the Advanced Benchmark (international median: 12%).

Regarding subscales, as an assessment of online informational reading, there were no reading purpose (Literary or Informational) subscales. For the comprehension process subscales of Retrieve/Infer and Interpret/Evaluate, Irish pupils achieved broadly similar scores on each and there were no notable differences by gender.

Forthcoming reports

An important aspect of PIRLS is the large amount of contextual data collected. This includes questionnaire data from principals, teachers, parents and pupils as well as an encyclopedia which describes the education context within each participating country. While the encyclopedia is released shortly before the main international reports on PIRLS and ePIRLS are launched, the full, final international dataset (all achievement and context variables for all countries and benchmarking participants) will not be released until early in 2018. Prior to then, each participant has access to a draft international dataset. Thus, this initial report for Ireland was based on the final approved dataset for Ireland and on the draft content of the international reports.

Further in-depth analyses will be conducted once the final international dataset is released. These later analyses will examine Irish responses to questionnaire items and situate the achievement results in a national and international context. Themes that will be covered include:

- the relationship between performance on the digital and paper assessments.
- variation in the use of technology (by individual, school and country, and over the 5-year period between PIRLS 2011 and 2016).
- analyses of some of the released content from PIRLS and ePIRLS, with specific emphases on items on which Irish pupils performed atypically.
- a multilevel model of pupil achievement.
- aspects of the class and school "climate".

References

Eivers, E., & Clerkin, A. (2012). *PIRLS & TIMSS 2011: Reading, mathematics and science outcomes for Ireland*. Dublin: Educational Research Centre.

Mullis, I.V.S., & Martin, M.O. (Eds.) (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., & Drucker, K.T. (2012). *PIRLS 2011 international results in reading*. 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 Boston College, TIMSS & PIRLS International Study Center website: 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 Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/epirls2016/international-results/

Mullis, I. V. S., Martin, M. O., Goh, S., & Prendergast, C. (Eds.). (2017). *PIRLS 2016 encyclopedia: education policy and curriculum in reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/pirls2016/encyclopedia/

Educational Research Centre

Foras Taighde ar Oideachas

www.erc.ie/pirls

ISBN 978-0-900440-57-1