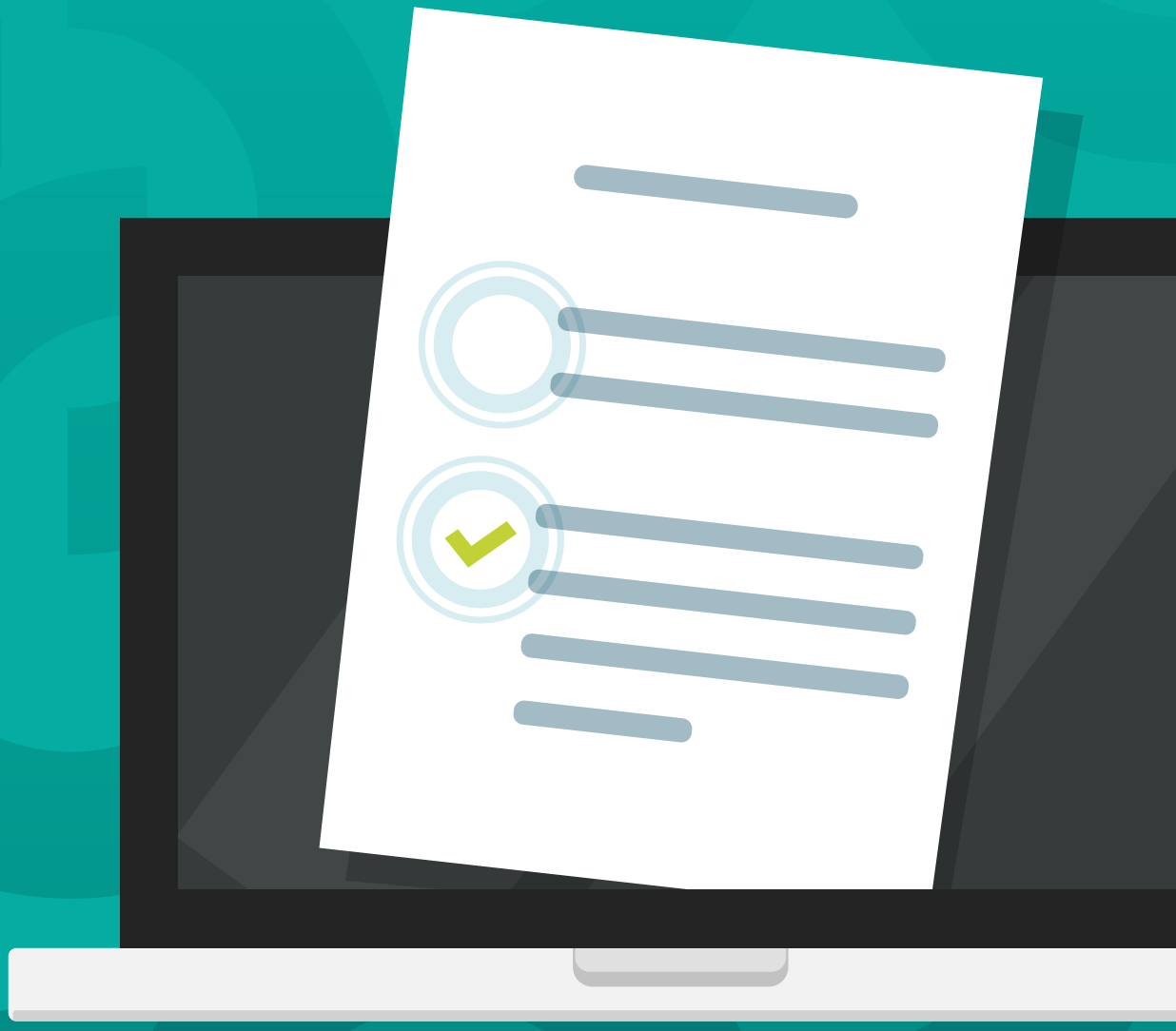
E-testing and computer-based assessment



CIDREE Yearbook **2024**

Authors:

Lieselot Vandenhoute, Bram Born, Griet Van den Eynde, Soetkin De Knijf, Dries Vrijsen, Kristóf Velkey, Fási Andrea, Emer Delaney, Adrian O'Flaherty, Rachel Perkins, Aidan Clerkin, Rachel Cunningham, Osman Buleshkaj, Fatmir Elezi, Selim Mehmeti, Dr. Philipp Sonnleitner, Steve Bernard, Dr. Sonja Ugen, Max van der Velde, Bernard Veldkamp, Jos Keuning, Max van der Velde, Bernard Veldkamp, Jos Keuning, Kevin Steinman, Katarina Aleksić, Katarina Glišić, Branislav Ranđelović, Ana Radović, Andrejka Slavec Gornik, Karin Hector-Stahre.

Editors:

Branislav Ranđelović, Elizabeta Karalić, Katarina Aleksić, Danijela Đukić

CIDREE national coordinator:

Danijela Đukić

Graphic design, layout and cover:

Institute for Education Quality and Evaluation

Published online:

https://ceo.edu.rs/wp-content/uploads/2024/11/CIDREE_Yearbook_2024.pdf

How to cite this publication:

Ranđelović B., Karalić E., Aleksić K., Đukić D. (Ed.) (2024). E-testing and computer-based assessment. CIDREE Yearbook 2024. Belgrade: Institute for Education Quality and Evaluation.

ISBN – 978-86-6072-153-4

Contents

Ireland

IRELAND: BIOGRAPHIES

Emer Delaney

Emer Delaney works as a Research Fellow at the Educational Research Centre (ERC), where she oversees the development of tests for schools in Ireland. Since 2022, her responsibilities have included aspects of the ERC Drumcondra Online Testing System (DOTS), an e-assessment platform. Previously, Emer's work has focussed particularly on literacy. From 2016-2019, she led the development of a suite of new standardised tests of reading for primary schools, normed both online and on paper. She was Ireland's National Research Coordinator for the Progress in International Reading Literacy Study (PIRLS) 2021, having also worked on the 2016 cycle. Currently, she is a member of the PIRLS 2026 Reading Development Group and contributes to an initiative of the FLIP+ international e-assessment community to develop a shared online library of test items. Her wider research interests include equity in education and the intersections of gender studies with educational research.

Adrian O'Flaherty

Adrian O'Flaherty works as a Research Associate at the Educational Research Centre (ERC). His work focusses on aspects of the computer-based testing system ERC DOTS (Drumcondra Online Testing System), mainly around test development and new system functionalities. His primary research interests are development and delivery of online assessments, and he was part of a team that oversaw the development of the ERC DOTS online testing platform, the current version of which was made available to Irish school in 2021 and was nominated for numerous eAssessment Awards. Adrian has previously worked on several international projects (PISA and TIMSS), and also on an evaluation study of a national school's support programme (DEIS – Delivering Equality of Opportunity in Schools) aimed at addressing educational disadvantage in Irish schools. He is a member of the FLIP+ international e-assessment community which is currently developing a shared online library of test items.

Rachel Perkins

Rachel Perkins is a Research Fellow at the Educational Research Centre. She works primarily on international large scale studies at post-primary level and has managed the national implementation of a number of such studies, including the Teaching and Learning International Survey (TALIS), Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS). Currently, she oversees the national implementation of PISA and is Ireland's representative on the PISA Governing Board. Her research interests include student well-being, academic resilience, educational disadvantage, and the relationship between students' self-beliefs and their academic achievement.

Aidan Clerkin

Aidan Clerkin is a Research Fellow at the Educational Research Centre (ERC). He oversees studies at both primary and post-primary levels, encompassing large-scale assessments and other research and evaluation work, including a recent evaluation of Ireland's Digital Learning Framework. He has experience of both national (NAMER) and international (TIMSS and PIRLS) large-scale assessments, delivered digitally and on paper. From 2018-2022, Aidan led the development of ERC DOTS, an online platform for delivery and reporting of standardised tests to schools. With the World Bank, he has contributed to research on a range of topics for countries including Morocco, Egypt, Saudi Arabia, Nigeria, and Tanzania. His research interests include social-emotional development, student engagement and wellbeing, and their relationships with academic achievement; intervention and programme evaluations; and longitudinal research methods.

Rachel Cunningham

Rachel Cunningham is a Research Associate at the Educational Research Centre (ERC). She works primarily on standardised, screening and diagnostic tests of mathematics. From 2016-2019, she led the development of standardised mathematics tests for primary schools in Ireland, which were normed both online and on paper. She also had a central role in the development of content for an online mathematics assessment targeted at students near the end of Grade 8. As part of this work, she was involved in the development process for the ERC's Drumcondra Online Testing System (DOTS). She has also been involved with FLIP+ in developing a shared online library of test items. Rachel has worked on international assessments (TIMSS and PISA), and more recently, on the National Assessments 2021 (a large-scale national assessment of English reading and mathematics). Her research interests include all aspects of mathematics education, particularly the diagnostic applications of assessment.

COMPUTER-BASED TESTING IN IRELAND, 2005-2024: CHALLENGES, LESSONS LEARNED, AND FUTURE POSSIBILITIES

Abstract

This chapter traces the use of computer-based testing in schools in Ireland in two contexts: international large-scale assessments (ILSAs) (since 2005) and national standardised testing (since 2016). Learning gleaned from Ireland's participation in computer-based ILSAs informed the development of a bespoke online platform for administering and scoring standardised tests (the ERC Drumcondra Online Testing System [ERC DOTS]). Conversely, knowledge about schools' use of ERC DOTS has informed Irish approaches in subsequent ILSA cycles.

Across ILSAs and standardised testing, recurrent challenges of computer-based testing are identified, including wide variation in: (i) education technology infrastructure in schools, and (ii) students' prior experiences using this for schoolwork. Although common to primary and post-primary settings, infrastructural challenges were most apparent at primary level while differences in students' experience were most pronounced below Grade 3. Adaptations and solutions trialled are discussed. Looking ahead, further possibilities of computer-based testing, including enhanced accessibility, additional item types, and adaptive testing, are considered.

Introduction

Computer-based platforms have been used in schools in Ireland to deliver international large-scale assessments (ILSAs) (since 2005) and nationally-developed standardised tests (since 2016). The Educational Research Centre (ERC), based in Drumcondra in Dublin, has responsibility for the administration of several ILSAs on behalf of Ireland's Department of Education. The ERC also has a statutory function to provide standardised tests and related materials to schools in Ireland, and in recent years this function has included the development and management of a bespoke platform to support the online delivery of standardised tests (the ERC Drumcondra Online Testing System, or ERC DOTS). Experiences of computer-based testing in Irish schools in the different contexts of ILSAs and ERC DOTS have informed and influenced one another.

This chapter begins by describing Ireland's participation in computer-based components/cycles of three ILSAs: the Programme for International Student Assessment (PISA), which assesses reading, mathematics, and science among 15-year-olds; the Progress in International Reading Literacy study (PIRLS), which assesses reading at Grade 4; and the Trends in International Mathematics and Science Study (TIMSS), which assesses mathematics and science at Grade 4 and Grade 8.

Next, the ERC's experiences of developing, using, and refining the ERC DOTS platform are examined. The rationale for creating the platform is considered, along with constraints and

unknowns that presented challenges. Specification requirements in relation to functionality and interface are discussed in relation to the particular context of students and schools in Ireland. Additionally, a test development study in which the same reading and mathematics content was normed online and on paper is explored in detail, as it provides initial insights into how mode differences may operate in Irish primary schools.

Finally, drawing on experiences with both ILSAs and ERC DOTS, we identify persistent challenges associated with computer-based testing in Ireland, important lessons learned to date, and possible avenues of future development.

Ireland's experiences of computer-based testing in ILSAs

PISA

The ERC, on behalf of the Department of Education in Ireland, has administered the national implementation of PISA since its first cycle in 2000. Ireland was among a small group of countries to be involved in the early administration of computer-based elements of the assessments. As part of the field trial in 2005, Ireland, along with 12 other countries, took part in the optional Computer-Based Assessment of Science (CBAS). The assessment, which was well-received by schools and students, was preloaded on laptops provided by the ERC and up to 20 students in each of 30 schools took part in the study as part of the field trial in Ireland. However, using externally-provided laptops was both costly and time-consuming to set up and Ireland did not participate in this assessment during the main study administration of PISA 2006 (Cosgrove & McMahon, 2005).

In PISA 2009, Ireland again took part in the optional computer-based assessment, which assessed digital reading by presenting reading literacy tasks in simulated web-based environments. In total, 19 countries, including Ireland, participated in the digital reading assessment during the PISA main study, which was carried out in addition to the print-based assessment in sampled schools. Of the 35 students in each school who were selected to take part in the 2-hour print-based assessment, 15 were randomly selected to also participate in the 40-minute digital reading assessment, which took place after the print-based assessment, usually on the same day.

The digital reading assessment was delivered via a CD-ROM, meaning that schools' own devices could be used for testing. This cycle also saw a change in Ireland's test administration procedures for the PISA assessment as a whole. While external test administrators had been brought into schools to carry out testing for the first three cycles of the study (PISA 2000 to 2006), Ireland used the school associate model in 2009, meaning that a member of the school's staff administered the test to students. This change in procedures was introduced to address falling student response rates but it also facilitated the use of schools' devices for the digital reading assessment, as each device needed to be checked before testing to ensure it met the delivery specifications. In many cases, settings on schools' devices also needed to be changed to allow the devices to boot directly from the CD-ROM, which was necessary for the test to load. In practice, not every school had the capacity to carry out the digital reading assessment using their own devices and in approximately one-third of schools, laptops were provided by the ERC.

As was the case for the PISA 2006 CBAS, the digital reading assessment was well-received by schools and there was some evidence of increased engagement with it compared to the print reading assessment (Cosgrove & Moran, 2011). However, as well as the evident wide variation in education technology infrastructure in schools, feedback from teachers who administered the tests using schools' devices indicated that the work involved in checking and changing device settings was time-consuming and unmanageable. While only a small number of technical difficulties were experienced during the main study, where these did occur, they were an added burden for school staff. For these reasons, when Ireland participated in the computer-based assessments of mathematics, digital reading, and creative problem solving in PISA 2012 (which were again carried out in addition to the print assessments in sampled schools among a subset of students), laptops were provided by the ERC to all participating schools. These laptops were configured by the ERC so that they met the requirements, and the test software was preloaded onto USBs. Ireland also reverted to using external test administrators, all of whom received training in delivering both the print and digital assessments. Technical support was provided, as needed, by the ERC.

The experience of administering the additional computer-based assessments in PISA 2006, 2009, and 2012 meant that Ireland was well-placed to understand some of the challenges associated with computer-based testing when PISA made the transition to a fully digital assessment in most participating countries in the 2015 cycle. Nevertheless, as up to 42 students per school were to receive the digital assessments, the increase in the number of devices required presented a logistical challenge. Results of a survey carried out as part of the PISA 2015 field trial in Ireland indicated that about 40% of schools would not be able to complete the assessment using their own devices, while all remaining schools would require some external laptops to supplement their school devices. Thus, external laptops were again provided by the ERC for all participating schools. These laptops were transported to each school by a technical support person, who assisted the external test administrator in setting up the devices. This additional support was required to reduce the time associated with setting up 42 laptops. Technical support personnel also assisted with any technical issues that arose during testing and uploaded students' data when the session was completed.

This approach of providing external laptops, technical support personnel, and test administrators to participating schools was also used in subsequent cycles of the study (i.e., PISA 2018 and 2022). There are a number of benefits to this model, namely that the burden of checking and setting up devices is removed from schools as no access to school devices is required and there are fewer technical issues to deal with during the testing session. Furthermore, using external laptops provided by the ERC meant that it was possible to administer the test and questionnaire directly from each laptop's hard drive, rather than a USB drive, which improved the speed at which students accessed the materials. However, providing external laptops, which are hired in each cycle, is costly and rearranging test dates if requested by schools can be logistically difficult.

The move towards online testing in PISA 2025 provides another challenge, where school's access to a reliable broadband connection must be considered. Furthermore, while schools' education technology infrastructure has likely improved in recent years, initial communication with schools suggests that there are still some schools that would have difficulty providing a

sufficient number of devices for PISA testing. Thus, it seems that any transition to using schools' own devices in PISA testing in Ireland is likely to be a gradual one, with some level of external support required at least in the short-term.

PIRLS

Ireland has participated in three cycles of PIRLS: 2011, 2016, and 2021. PIRLS was fully paper-based until the 2016 cycle, when an add-on assessment of digital literacy (called ePIRLS) was administered by 14 of the 50 participating countries, including Ireland.

The ePIRLS "projects" required students to navigate through hyperlinked informational texts in a simulated web environment and to answer questions about what they read (Mullis et al., 2015). As this content was separate to that of paper-based PIRLS, the idea was that the same students should complete both tests (Martin et al., 2015), but on different mornings, with ePIRLS second. The ePIRLS software was designed for USB delivery, but could alternately be delivered via computer hard drive.

The field trial in 2015 marked the first time that large-scale computer-based testing was piloted in primary schools in Ireland. While previous experiences in PISA had flagged the variation in education technology infrastructure at post-primary level, it became clear that this was even more pronounced at primary level. Well ahead of testing, scoping visits to schools were conducted to run a system check on any available devices and to map out the room(s) designated for testing. It was common for school devices to fail the system check, mainly due to older operating systems and/or insufficient memory. As relatively few schools had dedicated computer rooms, classrooms were typically used for testing, with multiple extension leads often required to ensure that all devices could be plugged in. Scoping visits also identified the fact that a free antivirus program widely used by schools destroyed the test software (Eivers, 2019).

Following these visits, tailored plans were developed for schools. Approaches ranged from "school equipment only", through "mix and match" (whereby school devices were supplemented with laptops and/or extension leads supplied by the ERC), to "external equipment only" (whereby all laptops were provided by the ERC). Although schools were most frequently assigned to "mix and match", this often involved just a few school devices alongside a majority of external ones. Teachers re-ran the system check closer to the test day and it was not uncommon for previously-passing devices to fail. Therefore, the number of external laptops allocated to "mix and match" schools was increased, where possible, to provide contingency. During testing, some further issues with school devices became apparent. For example, in one school with a well-appointed computer room in which most devices had passed the system check, a hardware driver due for update caused simultaneous shutdowns mid-test.

Based on the experience of the field trial, it was decided to supply laptops to all schools for the main data collection in 2016. The model was similar to that used in PISA 2015 – although, for ePIRLS, laptops were purchased rather than rented, the cost being roughly equivalent. Given the expense and the logistical challenges of setup in primary classrooms, Ireland chose the option of administering ePIRLS to a random subsample of the PIRLS students – up to 22 students in

each sampled school. Overall, the main study went smoothly and all data were successfully uploaded. Post-testing, the laptops were offered for sale at a low price to participating schools. This required some additional coordination by the ERC but was well-received by teachers.

An exploration of achievement in international studies is beyond the scope of this chapter. However, it is of interest that the ePIRLS data were placed on the PIRLS scale, enabling direct comparison between the paper-based and digital reading achievement of the same students. In Ireland, students did well on both tests compared to peers internationally, and – perhaps surprisingly, given the variability of education technology infrastructure in primary schools – their average achievement on paper-based and digital reading was virtually identical (Eivers et al., 2017).

In PIRLS 2021, countries could administer PIRLS either entirely on computer (with ePIRLS tasks included in the rotation and other "traditional" PIRLS texts transposed to a digital format) or entirely on paper (without any ePIRLS content) (Martin et al., 2019). Ireland initially opted for digital testing, proposing to provide laptops again to participating schools. Unfortunately, the field trial was interrupted by the COVID-19 pandemic and resultant nationwide school closures. For the main data collection, Ireland reverted to paper-based testing as, due to infection control measures, it was not feasible for technical support personnel and laptops to move between multiple schools.

PIRLS 2026 will be administered on computer for all countries (https://www.iea.nl/studies/iea/pirls/2026). As Ireland administered the ePIRLS hypertexts in 2016 but not 2021, the 2026 data will offer a first opportunity to examine national trends over time in primary-level digital literacy.

TIMSS

The evolution of another large-scale international assessment, TIMSS, in Ireland provides an example of the interplay between domestic and international considerations in developing and rolling out computer-based assessments. Ireland participated in the first TIMSS in 1995 at both primary and post-primary levels. After a hiatus, Ireland re-joined TIMSS in 2011 (primary only) and has participated in both the primary and post-primary components in every cycle since then (2015, 2019, 2023).

Up to and including the 2015 cycle, TIMSS was entirely a paper-based assessment. For the 2019 cycle, a digital version (known at the time as eTIMSS) was developed. eTIMSS was, to a large extent, designed to be administered as a parallel version of the paper-based assessment with many items presented in substantively the same format, although additional functionality was added relating to response options (e.g., drag-and-drop) and greater use of automated scoring. The clearest difference between the paper and digital versions was that eTIMSS included an additional interactive component designed to assess students' problem-solving skills, known as Problem-Solving and Inquiry Tasks (PSIs), which had no equivalent in the paper-based assessment. Countries participating in TIMSS 2019 could choose at the national level either to administer TIMSS on paper or to move to the digital assessment. Countries that elected to administer eTIMSS also administered the paper-based assessment to a smaller sample of

students in a bridging study, which was designed to facilitate the estimation of any mode effects (i.e., differences in students' performance related to the mode of the assessment – either paper or digital). In the end, half of the 64 countries in TIMSS 2019 transitioned to eTIMSS (Perkins & Clerkin, 2020).

In Ireland, consideration was given to transitioning to eTIMSS at this time, but ultimately the decision was made to remain with a paper-based administration in 2019 with the intention of moving instead towards digital administration for the 2023 cycle. There were two main factors behind this decision.

First, from a policy perspective, Ireland had rejoined TIMSS at post-primary level (Grade 8) in 2015 following a 20-year gap. Moving to a digital version of TIMSS in 2019, with the attendant risk of mode effects, would have introduced an element of uncertainty to the estimation of trends between 2015 and 2019. This was considered particularly undesirable in the context of the renewed focus at Grade 8 and given the value of TIMSS as a means of monitoring numeracy outcomes towards the end of the Department of Education and Skills' (DES, 2011b) flagship Literacy and Numeracy Strategy, 2011-2020.

Second, from a pragmatic perspective, the possibility of a digital TIMSS in 2019 was complicated by the development of a major new suite of standardised tests for assessing reading and mathematics by the ERC (www.tests.erc.ie). As described in the second part of this chapter, these were standardised in spring 2018 and released for schools' use in spring 2019, with online versions provided via the ERC DOTS platform. The development of both ERC DOTS and the new tests, covering multiple grade levels (including overlapping grades with TIMSS), addressed a noted need to revitalise the standardised testing options available to schools (DES, 2016) and was the culmination of several years of work involving subject experts and stakeholders from across the education system, including pilot and standardisation studies involving thousands of students, as well as significant financial and professional investment in establishing the new systems. Upon the release of the online testing platform and the new tests, concerted efforts were made from 2018-2020 to raise schools' awareness of these new resources and to encourage them to move towards using the fully updated and re-normed tests as a replacement for older, more out-of-date versions.

In this context, it was felt that introducing a separate – and at that stage, unknown – international platform for administering a test of mathematics in primary schools, at the same time as rolling out ERC's own platform for administering mathematics tests in primary schools, would have presented too great a risk for confusion among schools who would use both. At the time of making the decision to administer TIMSS 2019 on paper or digitally, there was no way of knowing how the eTIMSS platform would function in practice or the likelihood of any system problems (e.g., software crashes). As such, the risk that any problems with the eTIMSS platform could have led to reputational damage to the ERC's own online standardised tests – for example, as a result of confusion between the respective platforms/tests, or due to the development of a generalised feeling among schools that online testing is difficult or unreliable – was deemed unacceptable. For both these reasons, TIMSS 2019 proceeded in Ireland with a paper-based administration.

More recently, TIMSS 2023 has seen Ireland make the move to digital administration (no longer known as eTIMSS, as digital is now the primary modality). The main data collection in 2023 was successfully carried out following a similar model to the one described above for PISA – that is, with laptops rented by the ERC and transported to schools, and with hired technical support personnel and ERC staff deployed to set laptops up in schools and assist test administrators (generally, classroom teachers) with testing sessions. With almost 11,500 students taking part in the digital assessment across more than 300 schools at primary and post-primary levels, the procurement of these hardware and personnel resources, and the logistical arrangements required to coordinate deliveries and schedules to schools around the country, represented a significant task and made much greater demands on ERC resources and time than a corresponding paper-based study would have.

Alongside this digital main study administration, Ireland chose to implement a bridging study (required for countries moving to digital in 2019, but optional in 2023) in order to gather information on the extent of mode effects, as well as on students' test-taking behaviour across the two modalities and their views of paper-based versus digital assessment. Insights from this bridging study will be crucial to interpreting the performance of Irish students in TIMSS 2023. Initial findings from both the main study and the bridging study in Ireland will be available from www.erc.ie/TIMSS in December 2024, with secondary analyses and further findings to be published from 2025 onwards.

Development of a bespoke online testing platform (ERC DOTS) in Ireland

Overview, rationale and timeline

Initial work on standardised assessments to be administered online to students in Ireland began in 2011. However, ERC staff were conscious of the possibilities of computer-based testing for several years prior to this, as this approach was already being adopted gradually by PISA, as described above. Another factor that prompted this new approach to assessment was the proposed introduction of mandatory standardised testing in literacy and numeracy at Grade 8 as part of a broader Literacy and Numeracy Strategy (DES, 2011b). Although this particular proposal was never enacted by the Department of Education, work on developing an online platform had already been initiated by the ERC.

In 2017 a platform to facilitate online administration of standardised tests in reasoning, English reading, and mathematics was made available to post-primary schools. The platform was developed by an external software development company to the specifications of the ERC. The online tests were developed in-house at the ERC, including the creation of the digital version of the test items, tests, and all reporting and administration instruments. The ERC was also responsible for standardising the online tests with representative samples of students prior to release.

In 2019, following several years of development work, new versions of the ERC's primary-level standardised tests in English reading and mathematics were made available on the same online platform for grades 3 through 6. A second iteration of the online platform followed in 2020.

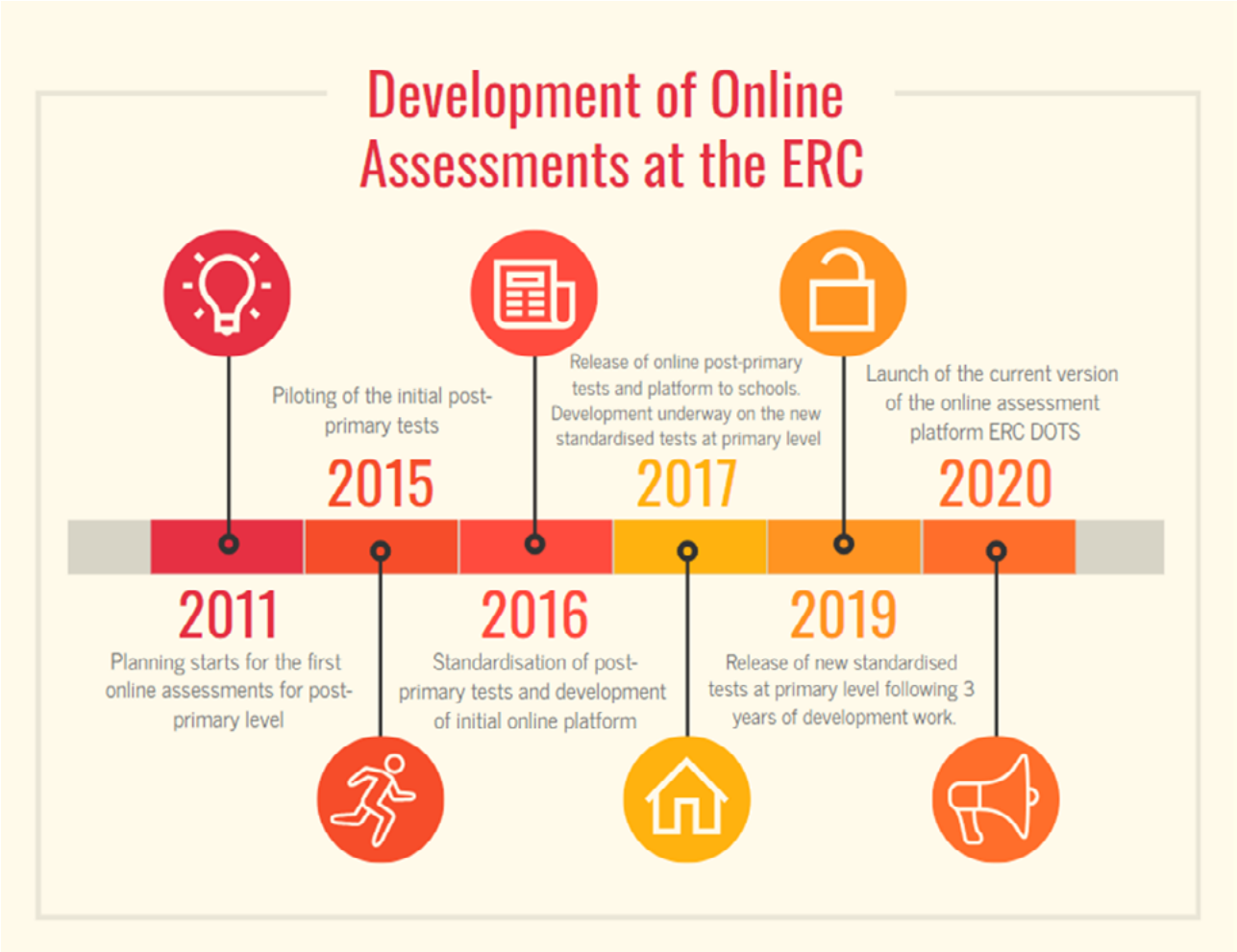


Figure 1: Timeline of the development of online assessments by the ERC

Anecdotally, there was growing demand from some schools for computer-based assessments. Online assessments were also increasingly becoming part of the large international studies, which was an additional factor behind the decision to create a computer-based option for schools for national tests.

As the technology to facilitate online assessments became more readily accessible, it was important to start exploring their obvious advantages. Among these is the reduced marking workload for teachers, with a much quicker turnaround time for reporting of results. Additionally, online assessments provide increased test security. Practising the tests in advance is not available as an option as access to the platform and to test credits for an active session is required. Replication of the online test for practice on paper would be an involved process of capturing, compiling, and printing screenshots of the content, within time restrictions (active test sessions are only available between 8am and 8pm on any given day). Online tests also allow for easier updating of content over time and more flexible piloting of new items.

On a broader level, there is increasing emphasis on the use of online resources in the classroom so assessments analogous to these pedagogical practices should be available. Digital testing platforms allow for an improved user experience with possibilities for a more engaging interface, interactive question types, adaptive testing, and specialised technology to assist test-takers where needed.

Platform development: Specifications and design considerations

The ERC has a statutory mandate to provide and support standardised testing for schools in Ireland. In developing ERC DOTS to offer digital versions of standardised tests, the aim was to provide schools and test-takers with an easy-to-use system that is reliable, secure, engaging, and meaningful in terms of the reports that it provides. The following description of requirements and system features is based on the current ERC DOTS platform.

In developing ERC DOTS, the ERC sought to create a fully integrated online assessment system in both the English and Irish languages which would permit test development (by the ERC) and test delivery (in schools) to operate efficiently in tandem. More broadly, the vision was to create a single consolidated environment for schools to engage with all ERC tests (paper and online), thereby allowing the ERC to bring its assessment offerings to schools as one cohesive set.

ERC DOTS is viewed as a dynamic entity to help the ERC respond to changes in education policies and curricula, digital technologies, and assessment and measurement technologies. It is a flexible and responsive assessment system that helps to facilitate the ERC's strategic priorities of internal capacity building (for example, in relation to the design of new, interactive item types) and research and development (for example, enabling the use of anonymised test data to study response patterns over time).

The test development functionalities of ERC DOTS include item banking and flexible assembly of test forms. At the time of writing, there are 58 discrete online tests and approximately 3,700 discrete online test items available on the platform. A variety of section and item templates are used to suit the requirements of different tests. All such content must be manually entered on the system by ERC staff. While the current tests that are available via ERC DOTS comprise multiple-choice items, partly as a result of the parallel development of paper-based forms of the same tests (described below), additional formats (e.g., open text, drag and drop) are available within ERC DOTS for future test development.

The system is available to all schools in Ireland. Teachers can register as users, purchase online test credits, prepare for test administration by uploading test taker or class details, and carry out online testing with their students (with the ability to monitor progress in real time on a monitor portal). Immediately after testing, they can access reports of student scores, expressed as standard scores and percentile ranks on the basis of pre-established norms, using an automated scoring functionality. All administration documents (administration manual, explanatory reports, technical manuals) are available for download from the teachers' side of the system.

In addition to the ability for schools to produce reports at class and individual student level, ERC DOTS also facilitates reporting on test and item characteristics (such as classical item statistics, and test form or subscale aggregate percent correct scores) to assist ERC staff with future research and test development work.

As well as its digital assessments and related features, the current functionality of ERC DOTS includes a comprehensive payment portal to allow schools to order paper tests for delivery; a backend element to streamline ordering processes for sales staff at the ERC, facilitating efficient order preparation and tracking; and an integrated scoring tool that allows teachers to easily score paper tests and produce reports with the same detail and format as those available for the online tests.

Features of the online tests and the test-taker interface

As test-takers on ERC DOTS include young children (from Grade 3 up), it was crucial that the test delivery interface be intuitive, with a minimalist design to avoid distractions on the screen while retaining all essential functionalities. It was also important to replicate (as closely as possible) the test-taking experience on paper, as parallel paper-based standardisations were occurring contemporaneously with the development of the online tests. With this in mind, online test-takers on ERC DOTS are allowed to move back and forth through the test and can review and change their answers (within the permitted time). To facilitate this, a summary review screen appears at the end of each test section. On this review screen, students can see a summary of the number of questions that they have answered, and from here they can easily navigate back to individual items and change their response if desired. The review screen also highlights the number of skipped questions and allows direct navigation to these. This feature is particularly useful for students if the allotted time is almost up.

All test instructions are available on-screen and are complemented by a read-aloud administration script. Although the option of text-to-speech is not yet available on ERC DOTS, its incorporation is planned to enhance accessibility and reduce the impact of reading load on measurement of constructs other than reading. On-screen, pre-test instructions describe the nature of each test and the user interface features needed to progress through the test. Before commencing the timed part of any test on the system, test-takers are brought through sample items of the same types they will encounter during the test. Currently, all test items on the system are in multiple choice format with four answer options, replicating the paper versions of the tests. Only one item appears on the screen at any time, allowing the test-taker to concentrate on the task at hand.

Within the test, the screen displays the test-taker's name (as entered on the system by the teacher) along with a timer at the top that counts down to show the remaining time. The title of each test section is displayed throughout, along with the item number and the number of items in the section. These features help the test-taker to orient themselves within the test and to monitor their progress. Additional test interface features to enhance engagement include child-friendly, colourful illustrations and a user-friendly format (e.g., comprehension texts in reading tests are presented across multiple tabs to minimise scrolling).

Technical specifications and operational requirements

To ensure the longevity and robustness of ERC DOTS, the system has been designed to be compatible with API (Application Programming Interface) and QTI (Question and Test Interoperability specification) standards. This allows maximal uncoupling of the platform's structure and content features to minimise risk associated with interdependencies, and to permit flexibility in future development (e.g., easier incorporation of potential future modules for scoring of text responses, provision of an offline assessment solution, or adaptive testing).

Being mindful of the variation in levels of IT infrastructure and internet access across schools (as described earlier in the context of ILSAs, and also evident from other research – see for instance Feerick et al., 2021), local device and network requirements have been kept to a minimum. An additional feature is a dummy test which allows schools to test their devices' ability to run tests prior to a test session (https://trythetest.erc.ie).[1]

Minimising the risk of server overload and maximising stable and complete response data, while taking account of changing numbers of users and fluctuations in local connectivity during testing, were prominent considerations in the development of ERC DOTS. At the same time, a key consideration, influenced by the varying quality of internet access in schools, is the ability for ERC DOTS to recover effectively from temporary local or server-level crashes or disruptions, with no loss of data and minimal impact on test-takers. For example, following a temporary drop in internet connectivity or if a student has to log out mid-test for any reason, upon logging in again the test-taker is returned to the last response recorded, without loss of data and with the correct time remaining on the test as if there had been no interruption.

Finally, it was considered essential that ERC DOTS would be accessible and intuitive for teachers to use. This requirement applied across a range of functions, including ordering tests, setting up test sessions, administering tests, accessing reports, and accessing supporting documentation. With this in mind, templates are provided within ERC DOTS for the upload of test-taker details, while report templates are downloadable in several formats (e.g. individual reports, class reports) and can then be uploaded to schools' existing content management systems (CMSs) without the need for any modifications. Since launching the platform, some work has been done towards further increasing compatibility with common CMSs (e.g., to allow for direct importation of class lists from a CMS and direct exportation of reports to a CMS).

Case study: Development of primary school tests in two modes (2016-2019)

This section describes how the ERC's online testing platform was used during a large-scale test development project.

1 Many of the features of the test-taker interface as described in this section can be viewed by readers at trythetest.erc.ie.

Background

Since 2012, it has been mandatory for primary schools in Ireland to conduct standardised testing of English reading and mathematics at grades 2, 4, and 6, to report results to parents, and to report aggregated results to the Department of Education (DES, 2011a).[2] The ERC provides tests in these subjects for grades 1–6 and, in practice, many schools opt to test at grades 1, 3, and 5 as well. The development of new versions of the primary tests was prompted in part by recognition that the norms for previous versions, standardised in 2005 and 2006, had become outdated (DES, 2016). This was probably due to a combination of schools' increasing familiarity with the test content and a genuine improvement in reading and mathematics standards in Irish primary schools, as also observed in national and international assessments (National Assessments of Mathematics and English Reading [Shiel et al., 2014], TIMSS [Clerkin et al., 2016], and PIRLS [Eivers et al., 2017]). Recognition of this issue signalled an urgent need for the tests to be redeveloped, with a new set of norms, so that students' performance could be described relative to an up-to-date population.

A key decision involved the mode(s) through which the redeveloped tests should be made available. Ireland's experience of administering ePIRLS in 2016 had highlighted the variability of infrastructure in primary schools. Moreover, while ePIRLS had not required internet access, the ERC's DOTS platform did, and it was anticipated that poor broadband might prove an additional barrier to digital testing. Therefore, continuing to provide a paper-based version of the new tests was essential to ensure accessibility for all schools. The question that remained was whether providing online versions of the tests as well would be of substantial benefit for those schools that had the infrastructure to avail of them.

In considering this question, one unknown was the extent to which primary school children in Ireland were familiar with using digital devices for schoolwork or similar purposes. If familiarity was low, or variable, this might impact on their performance on tests of reading and/or mathematics – i.e., there would be a substantial mode effect, which was not desirable. The findings from PIRLS 2016, whereby students in Ireland had performed equally well, on average, on paper-based and digital reading, were tentatively encouraging in this regard. However, in PIRLS 2016 questionnaires, many students in Ireland indicated that they had learned their computer skills mainly outside the classroom (Eivers, 2019). Furthermore, it was far from certain that the behaviours of Grade 4 PIRLS students would apply across all the grade levels targeted by the new tests.

Weighing these concerns against the advantages of online testing for schools, including the reduced burden for teachers and the potential for the interface to improve children's experiences, it was decided to pilot the new tests on both paper and online modes for grades 2-6. This would allow for data on mode effects to be collected at all these grade levels – a first in Ireland – which would inform the approach taken in the subsequent standardisation.

2 In primary schools in which Irish is the medium of instruction (about 8%), standardised tests of Irish must also be conducted and reported on at these grade levels.

Pilot (2017)

For convenience, pilot schools were selected on the basis that they had at least two classes per grade level and, per their responses to a survey, had at least 25 usable digital devices and high-speed broadband. The intention was that, at a given grade level, one class in a school would take the tests on paper and the other class would take them online. In practice, some schools that had indicated that they had sufficient infrastructure ended up needing additional laptops, and sometimes Wi-Fi routers, to be supplied, while a few schools could not participate in online testing at all. More than 2,800 students in 56 schools took part in paper-based testing, while more than 3,300 students in 52 schools took part in online testing.[3] At each grade level, multiple forms (versions) of each test were piloted.

Pilot item statistics showed that the majority of items in both reading and mathematics were easier on paper than on computer, although there were exceptions. At the level of test forms, mode differences in the difficulty of the reading tests ranged from 0-6%, while mode differences in mathematics ranged from 0-8%. At lower grade levels, mode differences tended to be larger and more consistently indicative of test content being more difficult on computer, particularly in mathematics. While the overall trend suggested that the online format was a little more challenging in both subjects, especially for younger students, the differences observed were considered small enough to allow the tests to proceed to standardisation in both modes. As part of the post-pilot review, items with very large mode differences were removed or adjusted, while minimising mode differences at form level was a consideration when items were rearranged to balance the forms. Finally, given the larger mode differences observed among younger students, it was noted that the suitability of both modes for grades 2 and 3 in particular would be reviewed post-standardisation and prior to release.

In preparation for the standardisation, significant changes were made to the Grade 2 mathematics test. In the pilot, the full test was read aloud to students, meaning that the same form of the test had to be administered within each class. However, informal feedback from schools indicated that this approach was problematic when administering the online test as copying was facilitated by the proximity and visibility of neighbouring screens. To a lesser degree, copying was also a concern for paper testing. In response to this feedback, the Grade 2 pilot forms were reworked as parallel forms for the standardisation. The first block of items was read aloud, but with minor variations in the numbers, images, and answer options used in each form, resulting in different correct answers. For the latter section of the test, the students worked through the items while reading independently, with different content in each form. This allowed multiple forms to be administered within the same class group. At more senior grade levels, although the mathematics test was not read aloud, teachers had the option of reading any words with which a student had difficulty, to reduce the impact of the reading load. This came with the proviso that mathematical terms could not be defined.

---

3 The larger number of students in fewer schools for online testing reflected the fact that schools were offered the option to test additional classes online if they wished. Illustrating some of the advantages of online testing, this created minimal additional costs for the ERC, whereas if schools had tested additional classes on paper increased costs would have been incurred for printing, postage, data entry, and shredding.

Standardisation (2018)

The standardisation brought new challenges. It was important for the tests to be standardised in both modes on a representative sample of the population – which included students in schools with and without the relevant education technology infrastructure. Therefore, unlike in the pilot, schools were sampled without regard to the availability of suitable devices or broadband. Each sampled school was asked to conduct paper-based testing at four grade levels and online testing at two grade levels.[4] Based on participating schools' reports of their own resources, a majority were provided with at least some laptops by the ERC, and about half were provided with a Wi-Fi router on the test day(s). There were proportionately more technical problems in the standardisation than had occurred in the pilot, often due to poor internet connection in schools.

Anecdotally, in schools where laptops were provided, some Grade 2 students in particular were observed struggling with practical aspects of the computer-based tests such as logging in, using a mouse to select answers, navigating between items and tabs, and zooming to enlarge content. Students at the same grade level who took the tests on their own school's devices (often tablets) appeared to have fewer such difficulties. Item statistics demonstrated that, in the representative standardisation sample, there was a substantial mode effect for students at Grade 2. As seen in the pilot, students at this grade level found the same items harder on computer than on paper; however, the differences at test form level in the standardisation were large and systematic enough in both subjects to be of concern. The overall percentage of correct responses on paper was broadly in line with what was expected and targeted. On the other hand, the lower percentage of correct responses online suggested that students who took the test on computer had had a non-comparable, and possibly demoralising, test-taking experience. This was especially the case in the reading test, where two out of three Grade 2 forms were about 8% more difficult online than on paper, with the remaining form about 5% more difficult online. On closer inspection, the relative difficulty of the online format at this level was especially noticeable when students were required to zoom to read text, and/or when content was distributed across tabs in such a way that each tab contained a semantically discrete section (for example, a specific character's perspective).

At other grade levels, form-level mode effects ranged from small to negligible. In reading, forms were slightly more difficult on computer at grades 3-5, but less difficult on computer at Grade 6. However, the patterns were somewhat different for mathematics. At Grade 3, one form was more difficult on computer, while the other was marginally less difficult. At grades 4 to 6, none of the forms were more difficult on computer, although there was some variation in the extent of the apparent advantage to taking the test on computer. It may be, then, that children's interaction with the platform was not consistent across the two subjects. This does not seem altogether surprising as the test format differed considerably by subject. For example, the reading tests included sections featuring both an item pane and a text pane, with the latter split across multiple tabs, whereas the mathematics tests featured just one pane.

4 This was considered a reasonable trade-off between the need to gather data from a sufficient number of students on each mode versus the cost and logistical challenges of supporting less-equipped schools to conduct online testing.

Having reviewed the magnitude of mode differences across grade levels, and giving due consideration to the benefits of online testing for schools that could avail of it (e.g., a reduced administration and marking burden for teachers), it was decided to release the tests for grades 3-6 in both modes but to release the Grade 2 tests in paper format only. A somewhat analagous decision had already been taken in relation to the response format for paper tests: while students at grades 3–6 marked their answers on machine-scorable answer sheets separate to the test booklets, students at Grade 2 marked their responses directly into their test booklets as the use of answer sheets was deemed likely to cause too much construct-irrelevant variance among this age group.

The paper-based and digital formats of the grade 3–6 tests were scaled separately, so that the norms now used by schools compare students' performance with that of peers in the standardisation sample who took the same test via the same mode.

Conclusion: Lessons learned and future possibilities

For schools, computer-based testing has been a paradigm shift from traditional methods. Therefore, there is an existing culture that has needed to adapt in order to take advantage of the conveniences and possibilities of digital assessment and keep pace with student engagement in an online world. This change has needed, and continues to need, careful management and system-level support to ensure that all stakeholders benefit. To facilitate this change on a wider scale, a robust ICT infrastructure (including reliable internet access) is needed across all schools in Ireland, with more work required at primary than at post-primary level. Contingent factors for an improved ICT landscape are ongoing professional development, technological knowledge, access to technological assistance, and embedding of the use of digital technology among school staff (Cosgrove et al, 2022; Donohue et al, 2024; Feerick et al, 2022). Continued learning and updating of skills are also required by ERC staff working in test development, particularly as the field of digital testing continues to evolve. Additionally, data security considerations are constantly growing and it is important for the ERC to remain up-to-date with best practice and relevant legislation, such as the European Union's General Data Protection Regulation and Ireland's Data sharing and Governance Act.

To date, the main model used to administer computer-based ILSAs in Ireland has involved the ERC renting or purchasing laptops and transporting these to schools for testing. This has facilitated Ireland's participation in important global developments in large-scale assessment and the collection of useful data regarding students' proficiency in a digital environment. However, it is a model that is very costly, as well as administratively burdensome – for schools as well as the ERC. Many of the infrastructural impediments encountered during the administration of computer-based ILSAs also presented challenges during the development of ERC DOTS. In developing the platform, it was possible to specify some design features intended to adapt to the national context – for example, the ability of the programme to run on a wide range of devices and to respond flexibly to crashes due to loss of internet connection.

Nevertheless, developing online tests for schools that have insufficient or variable education technology infrastructure inevitably involves constraints. The simultaneous development of analogous paper and digital assessments (to satisfy the needs of all schools) places heavy demands on the ERC's resources. It also requires the content and format of online tests to mirror closely those of their paper-based equivalents, restricting the ability to take advantage of interactive item formats. Additionally, mode effects may be more exacerbated in the norms than in the data resulting from use of purchased tests, as the norm group is based on a sample drawn to be representative of students in all schools while the online tests are typically purchased by a self-selecting set of schools with the necessary infrastructure.

The ERC is at the forefront of computer-based assessments in Ireland and will continue to use the knowledge and experience gained to date (through involvement in ILSAs, collaboration with national and international colleagues, and with development of ERC DOTS) to support schools in the transition to this mode of assessment. Additional possibilities of the digital format for national test development are now being explored, particularly in the areas of adaptive testing, innovative item types, and enhanced accessibility. Contributing to this, the ERC is a member of the FLIP+ international e-assessment community, a not-for-profit entity comprising researchers from many countries whose goal is to share experience and build solutions to enhance assessment globally (flip-plus.org). This includes working groups focussing on a broad range of assessment topics (e.g., process data, psychometric data, inclusion and accessibility). Much of the work of FLIP+ to date has centered on the development of an international item library which will facilitate the sharing of ideas, test items, and technical knowledge among its members.

In Ireland, there are several ongoing government strategies and frameworks in place designed to support the country's 3000+ primary schools and 700+ post-primary schools in their use of digital technologies (Cosgrove et al., 2019). In particular, the Digital Strategy for Schools to 2027 (Department of Education, 2022) – which succeeded a previous Digital Strategy running to 2020 (Department of Education and Skills, 2015) – has as one of its core pillars the improvement of education technology infrastructure in schools in Ireland. The commitments in the Strategy range from providing funding to schools for the purchase of digital technology to enhancing high-speed broadband connectivity and Wi-Fi in schools. For example, €210 million was provided to schools via an ICT Infrastructure grant between 2015 and 2020, with a further €200 million investment and an additional €13 million to improve schools' broadband connectivity promised for the period to 2027 (Donohue et al., 2024). The need to improve technical support services to schools and to streamline procurement frameworks is also acknowledged, as well as the need to provide guidance and advice to schools in areas related to digital technology. However, while improvements have been made in recent years, significant challenges in terms of connectivity and the availability and suitability of digital devices in schools are still apparent, especially among primary schools (Donohue et al., 2024).

A common principle underpinning all these supports, and another key pillar of the Digital Strategy to 2027 (Department of Education, 2022) is the promotion of the embedding of digital technologies and digital pedagogy in the classroom. Researchers at the ERC have recently completed a national, longitudinal evaluation of one such initiative, the Digital Learning Framework (Cosgrove et al., 2019; Donohue et al., 2024). Nonetheless, challenges remain also with regard to teachers' use of digital technologies, especially at primary level (Feerick et al., 2022).

As the development of digital literacy skills becomes increasingly prioritised in policy in Ireland (Department of Education, 2024), there is a growing need for assessments of digital literacy, and therefore some of the challenges noted here become more pressing. Notwithstanding these challenges, the willingness of schools to take part in computer-based ILSAs and the increasing number of schools using ERC DOTS demonstrate a positive inclination towards digital assessment. As we review progress to date and look towards the future, we recognise the importance of consultation with stakeholders – particularly teachers and students – to help guide future developments in e-assessment in Ireland.

References

Clerkin, A., Perkins, R., & Cunningham, R. (2016). *TIMSS 2015 in Ireland: Mathematics and science in primary and post-primary schools*. Educational Research Centre. https://www.erc.ie/wp-content/uploads/2016/11/TIMSS-initial-report-FINAL.pdf

Cosgrove, J., Feerick, E., Moran, E., & Perkins, R. (2022). *Digital technologies in education – Ireland in the international context: Trends and implications from PISA 2012-2018.* Educational Research Centre. https://www.erc.ie/wp-content/uploads/2022/09/DT-in-Ed-PISA-2012-to-2018-for-web.pdf

Cosgrove, J., and McMahon, S. (2005). *Report on PISA field trial of the Computer-Based Assessment of Science: Operations and outcomes in Ireland.* Educational Research Centre. https://www.erc.ie/documents/p06ftcbas_nat_report.pdf

Cosgrove, J., and Moran, G. (2011). *Taking the PISA 2009 test in Ireland: Students' response patterns on the print and digital assessments.* Educational Research Centre.

Cosgrove, J., Moran, E., Feerick, E., & Duggan, A. (2019). *Digital Learning Framework (DLF) national evaluation: Starting off – Baseline Report.* Educational Research Centre. https://www.erc.ie/wp-content/uploads/2020/01/DLF-national-evaluation-baseline-report.pdf

Department of Education and Skills. (2011a). *Circular 0056/2011. Initial steps in the implementation of the National Literacy and Numeracy Strategy.* https://assets.gov.ie/13643/037a8791893f466fa4bf439c47f6d8fa.pdf

Department of Education and Skills. (2011b). *Literacy and Numeracy for Learning and Life: The National Strategy to Improve Literacy and Numeracy among Children and Young People 2011–2020*. https://assets.gov.ie/24521/9e0e6e3887454197a1da1f9736c01557.pdf

Department of Education and Skills (2015). *Digital Strategy for Schools 2015-2020: Enhancing Teaching, Learning and Assessment.* https://www.gov.ie/pdf/?file=https://assets.gov.ie/25151/52d007db333c42f4a6ad542b5acca53a.pdf#page=null

Department of Education and Skills. (2016). *Standardised achievement tests: An analysis of the results at primary school level for 2011-12 and 2012-13.* https://www.gov.ie/en/publication/c27a52-standardised-achievement-tests-an-analysis-of-the-results-at-primary/

Department of Education (2022). *Digital Strategy for Schools to 2027*. https://www.gov.ie/pdf/?file=https://assets.gov.ie/221285/6fc98405-d345-41a3-a770-c97e1a4479d3.pdf#page=null

Department of Education (2024). *National Strategy for Literacy, Numeracy and Digital Literacy, 2024-2033*. https://www.gov.ie/pdf/?file=https://assets.gov.ie/293255/a509a8d7-a4ac-43f9-acb0-29cdc26a1327.pdf#page=null

Donohue, B., Moran, E., Clerkin, A., Millar, D., O'Flaherty, A., Piccio, G., & Dinh, T. (2024). *Digital Learning Framework (DLF) national longitudinal evaluation: Wave 2 Final Report*. Educational Research Centre. https://doi.org/10.70092/0412063.0824

Eivers, E. (2019). *Left to their own devices: Trends in ICT at primary school level*. Irish Primary Principals' Network. https://issuu.com/ippn/docs/left_to_their_own_devices_final

Eivers, E., Gilleece, L., & Delaney, E. (2017). *Reading achievement in PIRLS 2016: Initial report for Ireland*. Educational Research Centre. https://www.erc.ie/wp-content/uploads/2017/12/PIRLS-2016_inital-report-IRL.pdf

Feerick, E., Clerkin, A, & Cosgrove, J. (2022) Teachers' understanding of the concept of 'embedding' digital technology in education. *Irish Educational Studies*, 41(1), 27-39. https://doi.org/10.1080/03323315.2021.2022521

Feerick, E., Cosgrove, J., Moran, E. (2021). *Digital Learning Framework (DLF) national evaluation: One year on – Wave 1 Report*. Dublin: Educational Research Centre. https://www.erc.ie/wp-content/uploads/2021/06/DLF-W1-full-report.pdf

Martin, M.O., Mullis, I.V.S., & Foy, P. (2015). Assessment design for PIRLS, PIRLS Literacy, and ePIRLS in 2016. In I.V.S. Mullis & M.O. Martin (Eds.), *PIRLS 2016 Assessment Framework* (2nd ed.), pp. 55-69. Boston College, TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/pirls2016/downloads/P16_FW_Chap3.pdf

Martin, M.O., von Davier, M., Foy, P., & Mullis, I.V.S. (2019). PIRLS 2021 Assessment Design. In I.V.S. Mullis & M.O. Martin (Eds.), *PIRLS 2021 Assessment Frameworks*, pp. 57-74. Boston College, TIMSS & PIRLS International Study Center. https://pirls2021.org/frameworks/wp-content/uploads/sites/2/2019/04/P21_FW_Ch3_AssessDesign.pdf

Mullis, I.V.S., Martin, M.O., & Sainsbury, M. (2015). PIRLS 2016 Reading Framework. In I.V.S. Mullis & M.O. Martin (Eds.), *PIRLS 2016 Assessment Framework* (2nd ed.), pp. 11-29. Boston College, TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/pirls2016/downloads/P16_FW_Chap1.pdf

Perkins, R, & Clerkin, A. (2020). *TIMSS 2019: Ireland's results in mathematics and science.* Educational Research Centre. https://www.erc.ie/wp-content/uploads/2021/01/03-ERC-TIMSS-2019-Report_A4_Online.pdf

Shiel, G., Kavanagh, L., & Millar, D. (2014). *The 2014 National Assessments of English Reading and Mathematics. Volume 1: Performance Report*. Educational Research Centre. https://www.erc.ie/wp-content/uploads/2016/11/NA_2014_Vol1_Final-updated.pdf