On-screen exams

**Research report:
How digital assessment can
improve educational outcomes**

aqaglobal.com
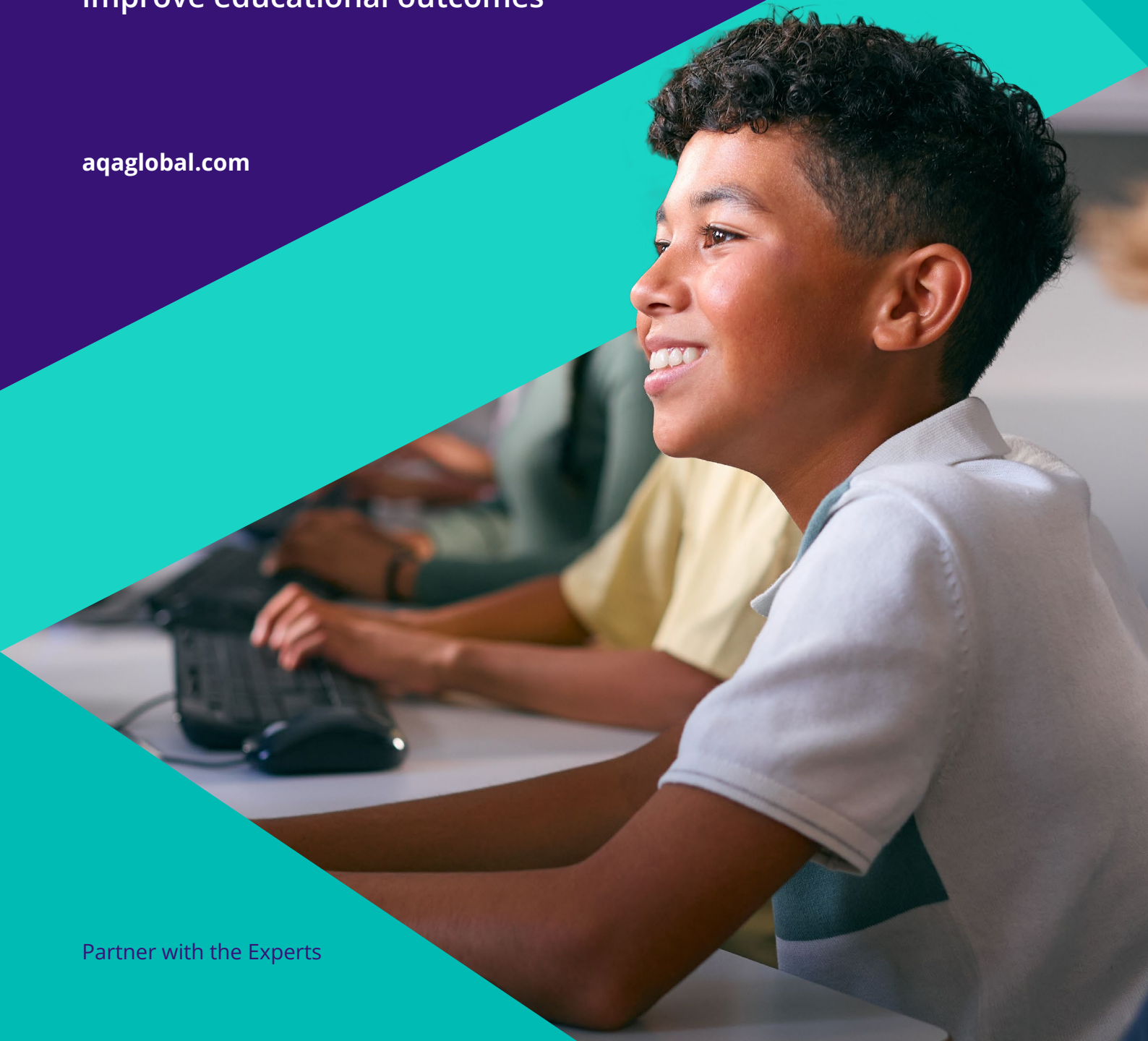
Executive Summary

Effective educational assessment is a cornerstone of successful education systems, playing a vital role in understanding student learning, guiding teaching, and informing policy decisions.

As European countries seek to improve educational outcomes new funding streams, like NextGenerationEU, present an opportunity to modernise assessment systems – particularly through digital assessments.

Assessments provide crucial data for teachers, learners, schools, employers, and governments, supporting individual learning progress, school improvement, and system-wide reforms. This report explores how formative, summative, and sample-based assessments can contribute to improved educational outcomes. It summarises the key findings from a range of educational research, including AQA, the UK's leading assessment provider.

Key assessment types discussed include:

- Formative assessments that provide ongoing feedback to improve teaching and learning.

- Summative assessments that measure learning outcomes at the end of educational stages.

- National and international large-scale assessments that offer benchmarking data and track trends over time.

The report highlights the importance of validity in assessment design – ensuring assessments accurately measure intended skills while remaining fair, reliable, accessible, and practical.

Digital assessment offers a transformative opportunity for education systems, delivering benefits such as:

- Faster feedback for students and teachers.

- Improved accessibility and support for learners with diverse needs.

- Enhanced security and efficiency in administration and marking.

- More engaging and interactive assessment experiences.

Drawing on our extensive expertise in assessment development and research, we outline how governments can leverage digital transformation to improve educational outcomes. Digital assessments not only streamline processes but also enhance data collection for evidence-based policymaking, ensuring that education systems are future-ready and aligned with international best practices.

Contents

Introduction

Educational assessment is a key part of any successful education system. It helps to measure student learning, guide teaching practices, and provide valuable data for policymakers.

As countries across Europe seek to improve their education systems new funding opportunities, such as NextGenerationEU, offer the chance to invest in modernising assessments. By leveraging these funds, governments can move towards digital assessments that better support student learning while aligning with international best practices.

AQA, the UK's leading assessment provider, has been actively researching and trialling digital assessments. This research highlights the benefits of online exams, including faster feedback, improved accessibility, and better preparation for real-world digital environments. As governments and education leaders consider these developments, it is important to understand what educational assessment is and how it has the power to provide data to drive classroom and system wide progress.

This report explores the role of assessment in education, and when and how it can be used. By considering assessment as part of an educational reform strategy, governments can use NextGenerationEU funding to create more efficient, inclusive, and future-ready education systems.

What is assessment?

Educational assessment is the process of measuring learners' knowledge, skills and abilities to understand their learning progress and achievement.

Assessments provide valuable information for students, teachers, schools, and policymakers, helping to guide instruction, improve educational outcomes, and ensure accountability within the education system (Pellegrino et al, 2001). Fulcher and Davidson (2010) identify two main types of educational assessment which are often contrasted against each other:

- **formative assessments**, which are used throughout the learning process to provide feedback, support learner development and inform teaching and learning approaches

- **summative assessments**, which are conducted at the end of a programme of study to assess whether an individual or group has been successful in meeting its aims.

Why is assessment important?

Educational assessment plays a crucial role in understanding what learners know, guiding teaching practices, and shaping education policy. Academic literature highlights several key perspectives on why assessment is essential.

- Black and Wiliam (1998) emphasise the role of assessment in improving learning through formative feedback, enabling teachers to adapt instruction to student needs.

- Sadler (1989) discusses how assessment helps students develop self-regulation and goal-setting skills, particularly when formative methods are used effectively.

- From a summative perspective, Harlen (2007) argues that assessment provides accountability, ensuring that education systems meet expected standards and that learners are prepared for further study or employment.

- Pellegrino et al (2001) highlight the role of assessment in evaluating complex cognitive skills, supporting curriculum development, and informing large-scale policy decisions.

Reasons for using assessments

Assessments are used by a wide range of people for different purposes across education, employment, and policymaking. Bachman and Palmer (1997) outline some of these.

- **Teachers** use assessments to track student progress, identify learning needs, and adjust their teaching strategies.

- **Learners** take assessments to demonstrate their knowledge and skills, receive feedback, and prepare for future education or careers.

- **Schools and universities** use assessments to determine whether a learner is ready to enter onto a course, evaluate learner (and sometimes teacher) performance, award qualifications, and ensure standards are met.

- **Employers and professional bodies** use assessments to ensure employees are competent, certify skills, ensure workforce competence, and support lifelong learning.

Governments use assessments as essential tools to monitor and improve education systems, ensuring that learners acquire the necessary knowledge and skills to succeed in further education, employment and society. Assessments provide valuable data on student performance, allowing governments to track progress over time, identify disparities in achievement, and evaluate the effectiveness of policies and teaching practices. They help in setting national benchmarks, ensuring consistency in education standards, and aligning curricula with economic and societal needs. Additionally, assessment results inform resource allocation, guiding investments in teacher training, curriculum development, and school infrastructure. International assessments also enable governments to compare their education systems with those of other countries, highlighting strengths and areas for improvement.

By using a mix of summative, formative, and large-scale sample-based assessments, governments can make data-driven decisions that enhance student learning outcomes and drive education reform. For example, In Guyana (AQA, n.d.a) the Ministry of Education has worked with schools and regional authorities to use data from assessments to inform school improvement decisions on key areas. This has supported evidence-based decisions on where in the curriculum to focus training and use the data to review its effectiveness. Using a broader approach, the Government of Bulgaria used a range of measures to help monitor and tackle disparities in attainment across schools – this was based on more information than just that derived from assessments (AlphaPlus, n.d.).

National and regional assessments

Summative assessments

These are used by governments all around the world to measure overall learner achievement at key stages and determine who is able, or eligible to move onto other stages of education such as university or vocational training. Examples include:

- GCSEs and A-levels in England[1]

- The Matura in central and eastern Europe, eg Austria[2]

- The Uganda Advanced Certificate of Education (UACE)[3]

- All India Senior School Certificate Examination (AISSCE)[4]

- The Higher School Certificate (HSC) in New South Wales[5]

Alongside their primary aim of recognising achievement, they allow governments to evaluate the effectiveness of national curricula, teaching strategies, and resource allocation. They also help identify achievement gaps between different regions, socio-economic groups, or demographic backgrounds, ensuring that targeted interventions can be implemented to promote equity.

[1] www.aqa.org.uk
[2] https://www.austriaeducation.info/tests/k-12-tests/matura-in-austria.html
[3] https://uneb.ac.ug/
[4] https://www.cbse.gov.in/cbsenew/cbse.html
[5] https://www.nsw.gov.au/education-and-training/nesa/hsc/about-hsc

Formative assessments

In contrast to summative assessments, which measure attainment of a standard, formative assessments are used by teachers and learners to review and check learning, in order to adjust lesson plans and approaches based what learners do and do not understand. Crucially, the assessment questions themselves are not 'formative'; it is the interpretation of the results, and the actions taken as a result of this that make the assessment formative.

Wiliam (2017) provides the diagram in *figure 1* which shows the three critical stages of formative assessment: knowing where a learner is in their learning, knowing where they are trying to get to, and what they need to do next. The diagram displays strategies for teachers, the learner themselves and their peers, and provides a range of strategies to advance learning.

Figure 1

	Where the learner is going	Where the learner is	How to get there
Teacher	Clarifying, sharing and understanding learning intentions	Engineering effective discussions, tasks, and activities that elicit evidence of learning	Providing feedback that moves learners forward
Peer		Activating students as learning resources for one another	
Learner		Activating students as owners of their own learning	

There is a wide range of evidence showing that this approach can have a significant positive impact on attainment. This is why it figures highly in efficacy studies, for example in England's Education Endowment Fund What Works Analysis[6].

Some countries have implemented national formative assessments which are focused on helping teachers, schools, and in some cases learners, to extract useful information about performance to boost progressions. At the same time these assessments can provide ministries of education with rich data about national cohort performance over time. For example:

- The Personalised Assessments in Wales[7]

- Scottish National Standardised Assessments[8]

Sample-based assessments

Sample-based assessments are designed to measure the educational achievements of a representative sample of students across a country or jurisdiction. Unlike large-scale standardised tests that assess every student, these assessments aim to provide data on overall educational quality and trends at the national level without testing every individual. They are cost-effective, manageable, and allow for in-depth data collection while minimising testing burdens.

---

[6] https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit

[7] https://hwb.gov.wales/curriculum-for-wales/reading-and-numeracy-assessments/personalised-assessments

[8] https://standardisedassessment.gov.scot/

Sample-based assessments are widely used around the world to monitor national educational progress, help countries track trends in student achievement, inform education policies, and provide insights into curriculum effectiveness. Some examples of the varying approaches are shown below.

Norway

National Assessments[9]

- Norwegian Directorate for Education and Training ("Udir") Division for Education and Culture Statistics runs these assessments through on-screen delivery.

- They cover literacy, numeracy and English. In addition to curricula for specific subjects, there are also separate frameworks for five cross-curricular domains, labelled as basic competencies.

- They are used by municipalities to inform school improvement.

France

CM1[10]; CP, CE1[11], CAP[12]

- These National assessments are regulated by the DEPP (the Ministry of Education's department for statistics, reports and evaluation).

- Individual results are delivered to teachers and learners' parents. Teachers have access to cohort and class data.

- Benchmarking tools have been developed to compare national and regional assessment data. Students' personal data is protected, but results according to schooling sector, social profile of the school, academic delay and gender are available on the government website for national and school academy levels.

United States

National Assessment of Educational Progress[13]

- The NAEP is the largest ongoing sample-based assessment in the United States, administered by the National Center for Education Statistics (NCES).

- It measures student performance in subjects like reading, mathematics, science, and writing for students in grades 4, 8, and 12.

- NAEP results are used to track academic progress over time, make comparisons among states, and inform educational policy at the national and state levels.

Australia

National Assessment Program: Literacy and Numeracy Sample Assessments[14]

- Australia conducts sample-based assessments in specific curriculum areas such as science, information and communication technology (ICT), and civics and citizenship.

- These assessments provide insights into students' proficiency in these domains, with results used to monitor trends and inform educational policy and curriculum design.

New Zealand

National Monitoring Study of Student Achievement[15]

- NMSSA assesses a sample of students in Years 4 and 8 on their proficiency in key areas such as reading, mathematics, science, and digital literacy.

- The assessment, conducted by the University of Otago in collaboration with the New Zealand Ministry of Education, informs national education policy and curriculum development, and supports equity initiatives by identifying performance disparities.

9 https://www.udir.no/contentassets/37ec4728a0e740cdb64e a130db355673/the-exam-system-in-norway

10 https://www.education.gouv.fr/evaluations-2023-reperes-cm1-premiers-resultats-379866

11 https://www.education.gouv.fr/evaluations-2023-reperes-cm1-premiers-resultats-379866

12 https://www.education.gouv.fr/test-de-positionnement-de-debut-de-premiere-annee-de-cap-2023-premiers-resultats-379878

13 https://nces.ed.gov/nationsreportcard/

14 https://www.nap.edu.au/naplan

15 https://curriculuminsights.otago.ac.nz/

International assessments

Many governments participate in international large-scale assessments (ILSAs). Unlike the national and regional assessments above, these assessments are undertaken in multiple jurisdictions around the world. Governments opt to take part in these assessments for several reasons:

- **Benchmarking Against Other Countries** – Results help compare a country's performance with global standards and identify strengths and weaknesses in the education system. The Council of Ministers of Education, Canada has used this approach with its Pan Canadian assessments to monitor standards across the country at a national level and also to benchmark to PISA (CMEC, n.d.).

- **Tracking Educational Progress Over Time** – By participating in multiple cycles, countries can measure improvements or declines in student achievement. For example, Australia was singled out for praise by PISA for its performance in creative thinking which has been a focus of its sample-based assessments for a number of years, leading to a targeted focus on this area (ACARA, 2024).

- **Informing Policy Decisions** – The data helps governments make evidence-based reforms, such as curriculum adjustments, teacher training improvements, or resource allocation. For example, the Welsh Government's (2023) mathematics and numeracy plan is informed by its PISA results.

- **Evaluating Equity and Inclusion** – The assessments provide insights into education gaps between socio-economic groups (Neuman, 2022), enabling targeted interventions to support disadvantaged students.

Many governments find the data from these ILSAs to be an additional and complementary set of evidence alongside their national assessments as they provide objective, comparable data that is internationally recognised. They also offer insights into teaching and learning practices in other countries, highlighting best practices from high-performing nations, that can be analysed and adapted to suit other cultures and contexts.

The three most well-known assessments – PIRLS, PISA and TIMSS – are operated by different international organisations and serve distinct but complementary purposes:

PIRLS (Progress in International Reading Literacy Study)[16]

This scheme is managed by the International Association for the Evaluation of Educational Achievement (IEA). It assesses reading literacy in Year 4 students (age 9–10) and has been undertaken every five years since 2001. It evaluates how well children are learning to read and compares reading outcomes across countries, helping governments improve early literacy strategies.

PISA (Programme for International Student Assessment)[17]

Delivered by the Organisation for Economic Co-operation and Development (OECD), this scheme assesses 15-year-olds' ability to apply reading, mathematics, and science skills to real-world problems. It has been carried out every three years since 2000. Its objective is to deliver a global comparison of how well students can use their education in practical settings, rather than just assessing knowledge of the curriculum.

> The OECD (2010) examined the relationship between learners' cognitive skills and national economic growth. They found that assessing, understanding and enhancing educational outcomes, as measured by PISA, can lead to substantial long-term economic benefits. It demonstrates that improved learning outcomes are strongly correlated with higher GDP growth rates. They estimate that if all OECD countries increase their average PISA scores by 25 points over the next 20 years this could result in an aggregate GDP gain of approximately USD 115 trillion over the lifetime of the generation born in 2010. This underscores the significant economic returns of even modest enhancements in educational performance.

TIMSS (Trends in International Mathematics and Science Study)[18]

This assessment is operated by the International Association for the Evaluation of Educational Achievement (IEA). It measures mathematics and science achievement at two stages of schooling: Year 4 (age 9–10) and Year 8 (age 13–14) and is conducted every four years since 1995. It aims to provide curriculum-based insights into how well students understand mathematical and scientific concepts as taught in schools (IEA, n.d.).

---

[16] https://timss.bc.edu/
[17] https://www.oecd.org/en/about/programmes/pisa.html
[18] https://timss.bc.edu/

Designing valid assessments

Whatever the type of assessment, it is very important that the design and approach used is valid for the intended purpose.

A valid educational assessment accurately measures what it is intended to assess and provides meaningful insights on learner or cohort performance. According to the seminal work on this by Samuel Messick (1989), validity is not just about test accuracy but also about the consequences of using assessment results. A valid assessment should be fair, reliable, and appropriate for decision-making, ensuring that results truly reflect a learner's knowledge and skills rather than external factors like test anxiety or access to resources. Later Carol Chapelle (1999) expanded on this by emphasising the importance of **construct validity** – the degree to which an assessment captures the intended learning outcomes – and the role of **authenticity** in digital assessments. Based on this work, these are some of the key factors that should be considered when designing assessments:

1 **Construct Validity** – The assessment must measure the intended skills, knowledge, or competencies without being influenced by irrelevant factors. This includes ensuring that the test design aligns with the curriculum and learning objectives.

2 **Content Validity** – The assessment should comprehensively cover the subject domain it intends to evaluate. Test items must be representative of the full range of skills and knowledge required, avoiding an overemphasis on specific topics or question types.

3 **Reliability and Consistency** – A valid assessment must be reliable, meaning that it produces consistent results under similar conditions. This involves designing clear questions, using appropriate marking schemes, and ensuring that different test administrations yield comparable outcomes.

4 **Fairness and Accessibility** – The assessment should be free from bias and should accommodate diverse learners, including those with special educational needs. This includes language accessibility, clear instructions, and options for assistive technology in digital assessments.

5 **Authenticity** – The tasks should reflect real-world applications of knowledge where appropriate. This is especially relevant for competency-based assessments and subjects requiring practical application, such as problem-solving in mathematics or analytical writing in language assessments.

6 **Consequential Validity** – The potential impacts of the assessment on teaching, learning, and decision-making should be considered. A well-designed assessment should encourage meaningful learning rather than teaching to the test, and should not disproportionately disadvantage any group of students.

7 **Practicality and Feasibility** – While maintaining validity, assessments must also be practical for implementation in terms of time, cost, and administrative resources. This includes considerations for digital delivery, marking efficiency, and security measures.

8 **Comparability and Standard Setting** – When assessments are used for large-scale purposes, such as national qualifications or international benchmarking, careful attention should be given to equating scores across different test versions, setting performance standards, and maintaining year-on-year comparability.

Digital assessment

Many of the assessments discussed in this report can be delivered in digital formats. Digital assessments can include formative and summative assessments, as well as large-scale sample-based assessments, each of which can be conducted in either fully on-screen or dual-mode formats. In a dual-mode approach, students can choose to complete assessments on paper or on-screen, ensuring flexibility while transitioning to digital delivery. This allows education systems to modernise while maintaining accessibility for all learners.

Research into on-screen assessment has identified several benefits that digital assessments can offer to students, teachers, and policymakers.

Benefits of digital assessment

- **Improved Student Experience** – Students who participated in AQA's pilot digital GCSE assessments found that typing their answers improved legibility and that digital platforms allowed for easier editing and restructuring of responses (Finch & Dunn, 2022). This is particularly useful for essay-based subjects.

- **More Interactive and Engaging Assessments** – Digital exams allow for the inclusion of interactive elements, such as video and animations, which can be beneficial for subjects like modern foreign languages (AQA, 2024).

- **Better Accessibility and Inclusion** – Assistive technologies, such as text-to-speech software, adjustable font sizes, and colour contrast settings, can help learners with disabilities complete assessments more effectively (AQA, 2023). This supports a fairer and more inclusive exam system.

- **Increased Efficiency in Marking and Administration** – Digital exams reduce administrative burdens, such as printing and distributing paper scripts. In addition, computer-based marking can improve accuracy and reduce bias in assessment outcomes (AQA, 2023).

- **Faster Feedback for Learning** – On-screen formative assessments can provide immediate feedback to learners, helping them understand their progress and target areas for improvement more effectively than traditional paper-based assessments (Finch & Dunn, 2022).

- **More Reliable and Secure Testing** – Digital assessments can incorporate security measures, such as randomised question orders and plagiarism detection, reducing opportunities for malpractice (AQA, 2023).

Which medium students say they are comfortable using for more than an hour

Computer

68%

Pen and paper

36%

Source: AQA, 2024
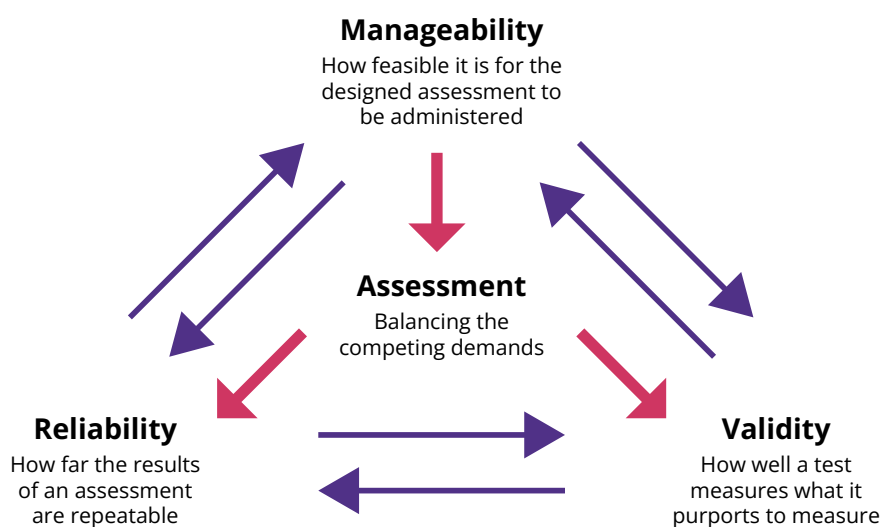
What's your vision for assessment reform?

As evidenced in this report, assessment plays a crucial role in education by providing valuable insights into learning, guiding teaching practices, and informing policy decisions.

When designed effectively, assessment not only measures achievement but also drives educational improvement by identifying gaps, supporting targeted interventions, and ensuring that education systems meet the needs of learners.

Digital assessment offers opportunities to enhance these benefits, making testing more inclusive, efficient, and data-driven.

As education systems evolve, so too must the ways in which learning is assessed. Digital assessment can modernise testing; whether you're looking to improve formative assessments, redesign summative examinations, or enhance data-driven policymaking, transitioning to digital assessments can offer significant benefits.

Assessment is all about balancing the competing demands of validity, reliability and manageability.

Manageability
How feasible it is for the designed assessment to be administered

Assessment
Balancing the competing demands

Reliability
How far the results of an assessment are repeatable

Validity
How well a test measures what it purports to measure

AQA has over 120 years of assessment expertise and knowledge, which we share with governments, ministries of education, and assessment providers to support assessment reform around the world.

We have delivered millions of high-quality assessments and have a strong track record in research and innovation. We can provide expert consultancy on assessment design, digital transformation, and policy alignment. Our work in developing and trialling digital assessments ensures that we can offer evidence-based insights into how best to implement new approaches effectively.

If you are considering assessment reform, contact us to discuss your plans and discover how we can work together to create an assessment system that meets the needs of students, teachers, and policymakers in the digital age. Email: **info@aqaglobal.com**

AQA Global Assessment Services

Learn more about digital exams and how we can support your vision for assessment reform at: **aqaglobal.com/on-screen-exams**

References

ACARA. (2024). *Australia singled out for praise in PISA 2022*. https://www.acara.edu.au/news-and-media/acara-update-archive/acara-update-archive-2024/acara-update-june-2024

AlphaPlus. (n.d.). UN Global Compact. https://alphaplus.co.uk/news/un-global-compact/

AQA. (2024). Pathways to digital exams – how teachers and students are helping us shape the future. https://www.aqa.org.uk/about-us/our-research/research-library/pathways-to-digital-exams-how-teachers-and-students-are-helping-us-shape-the-future

AQA. (2023). *Making it click: The case for digital examinations in England*. https://cdn.sanity.io/files/p28bar15/green/8857a632a5b53e844665b714950ed5c9a6f969d0.pdf

AQA. (n.d.a). *Case study: Guyana introduces a national data portal with GradeMaker Analytics*. https://aqaglobal.com/case-studies/guyana-introduces-a-national-data-portal-with-grademaker-analytics

Bachman, L. F., & Palmer, A. S. (1997). *Language Testing in Practice*. Oxford University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74. https://doi.org/10.1080/0969595980050102

CMEC. (n.d.). *Pan-Canadian Assessment Program (PCAP)*. https://www.cmec.ca/240/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-(PCAP)/Overview/index.html

Chapelle, C. A. (1999). Validity in language assessment. *Annual Review of Applied Linguistics, 19*, 254–272. doi:10.1017/S0267190599190135

Finch, K., & Dunn, T. (2021). *Student perceptions of technology in assessment*. https://www.aqa.org.uk/about-us/our-research/research-library/student-perceptions-of-technology-in-assessment

Fulcher, G., & Davidson, F. (2010). *Language Testing and Assessment*. Routledge.

Harlen, W. (2007). *Assessment of learning*. SAGE Publications.

Messick, S. (1989). Validity. In R. L. Linn (Ed.). *Educational Measurement*. MacMillan.

New South Wales Government. (n.d.). About the HSC.

Neuman M. (2022). PISA data clusters reveal student and school inequality that affects results. *PloS one, 17*(5), e0267040. https://doi.org/10.1371/journal.pone.0267040

OECD. (2010). *The High Cost of Low Educational Performance: The Long-run Economic Impact of Improving PISA Outcomes*. https://doi.org/10.1787/9789264077485-en

Pellegrino, J.W., Chudowsky, N., & Glaser R. (Eds.). (2001). Knowing what Students Know. The Science and Design of Educational Assessment. National Academy Press. https://doi.org/10.17226/10019

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144. https://doi.org/10.1007/BF00117714

Welsh Government. (2023). *Mathematics and numeracy action plan 2023*. https://www.gov.wales/mathematics-and-numeracy-plan-2023-html

Wiliam, D (Eds). (2017). *Embedded Formative Assessment: (Strategies for Classroom Assessment That Drives Student Engagement and Learning)*. Solution Tree Press