



CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT



Cambridge research papers at
AEA-Europe 2023

Thursday 2 November

11:30-12:45 Poster session

Room: Ballroom and Pre-Function Area

Accessibility considerations for digital assessments - development of a framework



Sanjay Mistry, Head of Research and Development for Digital Assessment & Evaluation, International Education

The development of classroom-based digital assessments and certified digital high-stakes assessments provides an opportunity for Cambridge International to enhance the accessibility of our digital assessment products for as many learners globally as possible. A digital assessment allows for affordances that traditional paper and pencil methods may not facilitate.

This poster summarises the thinking and associated research that has taken place within Cambridge International on accessibility considerations for digital assessments through the development of a framework. The framework has been designed under five different principles of accessibility research and is primarily for use by assessment designers in the development process. Each principle falls under one of five strands of research which will be highlighted. It will also define each accessibility principle with relevant examples. The practical element of the research to date considers a user-centred design approach and understanding the specific needs of the test taker, while adhering to appropriate digital accessibility standards and guidelines. Consideration is also given to universal design principles and the additional features given by digital assessment platforms to enhance the overall accessibility.

‘But what do we do with the results?’ A systematic approach to using assessment data holistically to improve teaching and learning



Suzanne Crocker, Assessment Products Integration Manager, International Education



Dr Irenka Suto, Head of Assessment, Cambridge Centre for Evaluation and Monitoring

Presented by Suzanne Crocker

When defining educational success, many cultures include life competencies, moral values and high wellbeing, alongside academic achievements and progression to further education or employment. More information on this can be found in the Ignite presentation on the Cambridge Learner Profile framework.

Teachers can assess or evaluate all these elements of holistic education. However, not all are confident in combining different types of data. This poster draws on academic literature and classroom experience to showcase an iterative six-step approach to collating assessment data; both quantitative and qualitative, and then showing how to use it formatively to inform resource and lesson planning. Working with 80 international schools in Malaysia, Indonesia, Thailand and Vietnam, we introduced the approach to 200 teachers in face-to-face workshops early in 2023. The attendees identified the data types they currently collected and then reviewed their own practices against the approach. Afterwards, they gave feedback through individual questionnaires and plenary discussions.

Almost all the teachers valued the approach and reported that it helped with organising and using data more effectively. However, while there was a widespread desire to adopt a holistic approach in understanding and supporting individual student needs, in practice limited resources and teacher experience could compromise its comprehensiveness.

Divergent considerations during the journey to internationalise mathematics questions in an adaptive baseline assessment



Emma Barthel, Assessment Developer,
Cambridge Centre for Evaluation and
Monitoring



Irenka Suto, Head of Assessment,
Cambridge Centre for Evaluation and
Monitoring

Presented by Emma Barthel

Many UK schools use standardised baseline assessments each academic year, valuing the year-on-year comparability of results data. We explore a tension arising in our journey to reform a long-established computer-adaptive baseline assessment. With an increasingly international user-group, we sought to improve accessibility while maximising the comparability of outcomes pre- and post-reform.

Mathematics questions were reviewed systematically by experts in mathematics, accessibility and linguistics. This included evaluating language demands against the Common European Framework of Reference and appraising the inter-cultural recognisability of images. Potential amendments were identified, and probable impacts on item difficulty and validity were hypothesised.

A subset of items was adapted, and the amended assessment was trialled internationally with children with a range of abilities. The trial data was analysed to ensure that continuity of item level difficulty had been maintained and to highlight any mismatches.

Stakeholders vary in the relative values they place upon accessibility and comparability of results. This poses a continuing challenge for assessment developers in determining what to prioritise, one which we discuss.

13:45 Comparative judgement

Room: Studio 7

Multiple marking using the Levels-only method in A Level English Literature



Emily de Groot, Statistician,
Assessment Research Division



Jo Ireland, Research Officer,
Assessment Research Division

Presented by Emily de Groot

Essays are important assessment tools allowing us to capture constructs that other assessment items cannot. To mark essays, analytical marking methods are commonly used. These require examiners to allocate marks, nested within levels of performance for different assessment objectives (AO) or features of the essay. However, this process can be time and labour intensive and thus typically only one examiner marks each script. This potentially renders the marking less reliable as it becomes subject to each examiner's individual preferences, leniencies or severities, particularly for the more subjective subjects such as English. While this can be mitigated with standardisation processes, marker monitoring and statistical scaling of results, more subjective subjects and extended writing tasks still tend to have lower marking.

One approach that can be taken to improve validity and reliability of results, and which this research explores, is to make essay marking quicker through less detailed marking methods, which could facilitate multiple marking. In multiple marking, multiple examiners independently assess the same script. The final mark is a combination of the marks.

To address this, Walland and Benton (2021) developed a less time-consuming essay marking method called Levels-only (LO) marking. With LO, the essay is assessed against each AO, retaining the original six levels of performance and accompanying descriptors. Under traditional marking,

2-3 marks would also be assigned within the levels, bringing a total mark up to 15 per question. In the LO method, these are not allocated. Walland and Benton (2021) tested the LO method on GCSE English Language. The approach involved examiners allocating each essay a level for each AO, without having to allocate marks within each level and without making annotations or summative comments. The results for double and triple LO marking yielded high reliability and predictive values while using a similar amount of examiner time as traditional marking.

In this project, we analysed how the LO marking worked with longer A Level English Literature essays. In particular, we focused on the time taken to mark, the predictive value of the method and the views and experiences of the markers using it.

Walland, E. & Benton, T. 2021. Multiple marking methods as alternatives to analytical essay marking: Comparing pairwise comparative judgement, rank ordering and Levels-only. Cambridge Assessment.

14:15 Psychometrics and test development I

Room: Studio 2A

Can examination papers always be error-free? An exploratory investigation into the conditions that can give rise to errors in assessment instruments.



Filio Constantinou, Senior Research Officer, Assessment Research Division

Successful reforms presuppose adequate understanding of the underlying problem(s). In the area of high-stakes assessment, one problem requiring attention is the occurrence of errors in examination papers (e.g. multiple-choice questions containing more than one correct answer). Errors can undermine students' performance in the examination, while also eroding public confidence in the examination system. Therefore, it is important that they are prevented. In England, although various measures are taken to this end, errors continue to occur occasionally. This suggests that the current process of developing examination papers may need to be revisited. However, before any changes to the process are attempted, the conditions giving rise to errors need to be sufficiently understood. To investigate these conditions, this study analysed interview data from 36 assessment professionals involved in the examination paper construction process. The analysis exposed the complex and often unexpected ways in which the characteristics of the paper construction process and the characteristics of the people who participate in it interact with one another, creating opportunities for error. This talk will present some 'active' and 'latent' human failures which can lead to errors, and will conclude with some reflections on whether examination papers can always be error-free.

14:15 Technical, vocational and applied assessments I

Room: Studio 8

Assessing 'competence' in education reform projects – what lessons can we learn from technical and vocational education?



Rebecca Conway, Education
Assessment Consultant



Matthew Dean, Senior Assessment
Manager, International Education

'Competence' is a concept that has become increasingly common in learning and assessment frameworks. It can be found at the heart of curricula such as the newly launched Curriculum for Wales, which is underpinned by literacy, numeracy and digital competence. In our own organisation, we are seeing significant demand from clients engaged in education reform projects that focus on competence-based curricula and assessment models.

While competence assessment is emerging in general education, particularly through its inclusion in the Organisation for Economic Co-operation and Development's (OECD), PISA (Programme for International Student Assessment) surveys, it is well established in the fields of technical and vocational education. Achieving 'occupational competence' is a desired outcome of much vocational and technical training (e.g. Scottish Vocational Qualifications) and the perceived failure to achieve it on a national scale is one of the key policy drivers for the government's sweeping reforms of the technical education system in England (Independent Panel on Technical Education, 2016). The reforms in England have been supported by research and reflection on what good 'competence' assessment looks like in high-performing jurisdictions (e.g. Germany, Switzerland).

This presentation will interrogate what is meant by 'competence', and 'competencies', and explore policies and practices from technical and vocational education to consider how these might be applied to a broader assessment context. In particular, we reflect on assessment methods, the timing of assessment, and approaches to marking and grading.

15:45 National tests and examinations I

Room: Amphitheatre

Evaluating the impact of curriculum and assessment reform in secondary education on progression to mathematics post-16



Carmen Vidal Rodeiro, Senior Research Officer, Assessment Research Division



Joanna Williamson, Senior Research Officer, Assessment Research Division

In England, GCSE (General Certificate of Secondary Education) qualifications offered to students aged 14-16 were recently reformed. For mathematics specifically, the new GCSE aimed to be more demanding, provide greater challenge for the most able students, and support progression to post-16 mathematics. However, there have been concerns that the new GCSE could deter students from further education in the subject and, to date, there has been little research on its impact on participation in and learning of mathematics post-16.

This research compared progression to and performance in post-16 mathematics and maths-related subjects pre- and post-GCSE reform and found that, contrary to fears about reduced uptake, progression generally increased following the reform. In post-16 mathematics, in particular, the increase was higher among those who achieved top grades in their Mathematics GCSE than for students with just a pass. Performance in post-16 mathematics and maths-related subjects was, in contrast to teacher expectations, generally lower post-reform.

As with any reforms, changes take time to bed in. However, in the interim, this research has raised important issues for the mathematics education community and for policy makers by increasing the understanding of how recent reforms to GCSE Mathematics have affected students, and contributing evidence on its impact on progression to post-16 study.

16:15 Assessment cultures I

Room: Studio 2C

Enacting assessment reform in Cox's Bazar Refugee Camp: a case study



Georgie Billings, Senior Assessment Advisor, International Education



Sarah Nelson, Curriculum Programmes Manager, International Education

In Cox's Bazar, Bangladesh, a network of refugee camps operates for the displaced Rohingya from Myanmar. Just under a million people live in these camps, and around half are children. Most Rohingya here have no legal identity or citizenship and are entirely reliant on humanitarian assistance.

Cambridge University Press & Assessment was asked by UNICEF to create an assessment framework for children being educated in the camps, alongside a system for monitoring and evaluating its implementation and impact. This was planned against the background of a whole-scale transition from an emergency curriculum (the Learning Competency Framework Approach (LCFA) to the Myanmar Curriculum in learning centres.

The enactment was complex, influenced by Covid-19, stakeholder engagement, and the changing geo-political landscape. Further cultural challenges, such as the engagement of girls in education, and the significant language barriers between teachers also had to be considered.

This presentation will focus on the enactment journey, with details from our experience in the refugee camps with planning, implementing and evaluating the project, updated to reflect our experiences over the next seven months of the project.

Friday 3 November

09:00 National tests and assessments II

Room: Studio 2C

How are GCSE grades used in post-16 admissions decisions?



Emma Walland, Research Officer,
Assessment Research Division



Tony Leech, Senior Researcher, OCR

Many recent policy proposals for the future of assessment in England suggest reform of GCSEs, qualifications taken by learners aged 14-16. Since GCSEs are no longer the main educational exit qualification (as learners must stay in education until 18) some writers have questioned their necessity. Some argue for the abolition of high-stakes external assessment in this phase; others for streamlining such assessment. To inform this debate, evidence of how GCSEs are used is needed. We investigate use of GCSE grades in admission processes for post-16 education, including sixth form and further education colleges. Using analysis of documents, questionnaires and in-depth interviews of teachers representing a range of school types and subjects, we explore teachers' views and current uses of GCSE grades in admissions, and what reform of GCSEs might mean for this decision-making. We found GCSEs play the central role in post-16 admissions decisions, and that external national examinations taken at age 16 were valued as selection tools by teachers, though many said they could still make selection decisions using fewer GCSEs. We explore impacts of the current system and any potential changes. We discuss findings in relation to theories of the uses and purposes, impact and consequences of assessments.

09:00 Assessment that is reactive to unforeseen circumstances III

Room: Studio 8

Covid-19-related changes to upper secondary assessments in six countries: adaptations and reactions



Nicky Rushton, Senior Research Officer, Assessment Research Division



Santi Lestari, Research Officer, Assessment Research Division

Presented by Nicky Rushton

When the Covid-19 pandemic started in 2020, many schools around the world closed and the examinations that were due to take place at the end of upper secondary education had to be cancelled, postponed or altered in some way. Covid-19 continued to affect the assessments in 2021 and 2022 in some countries. In this presentation we compare the changes to examinations, the effect on standards, and reaction to the changes in six countries (China, England, India, Italy, Spain and the USA).

Such changes could be expected to generate comment from a range of stakeholders such as teaching unions and the public. It can be difficult to find research evidence citing opinions at the time of changes, but newspapers and other media sources are able to publish such information with immediacy and often report the perspectives of important stakeholders. Therefore, we examined newspapers from the affected countries to identify opinions about the changes that occurred.

We discovered that although standards only changed in some countries, concerns were raised in every country. Some, such as increased anxiety, affected multiple countries, while others, such as technological problems, were confined to particular countries.

09:00 Assessment cultures II

Room: Amphitheatre

Exploring the longitudinal development of assessment practitioners via their participation in assessment professional development



Hannah Williams, Professional Development Projects Specialist, Cambridge Assessment Network



Simon Child, Head of Assessment Training, Cambridge Assessment Network

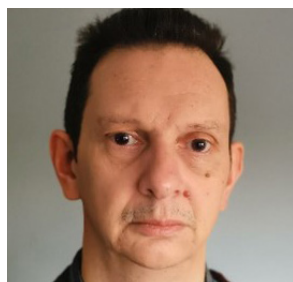
Presented by Hannah Williams

A crucial element of successful assessment reform is ensuring an increased rate of assessment literacy among education professionals, particularly for teachers and school leaders. Assessment literacy can be understood as knowledge of the principles and practice of assessment, including techniques, purposes and processes. Professional development for education professionals within the context of assessment reform and innovation is vital, yet often overlooked in research. This study examined the longitudinal impact of two assessment-related professional development courses on participants' professional lives, discussing in particular their performance, confidence, career development, and engagement with assessment as a whole. This longitudinal mixed methods study took data snapshots from three time periods: before participants undertook the courses, immediately after, and 12-24 months following course completion. We analysed questionnaires and feedback, and undertook semi-structured interviews with participants from a broad pool of professions including teachers, school leaders, exams officers and assessment managers in the UK context and internationally. Participants perceived that completing professional development directly led to career progression and that this was attributed to socio-cultural interpretations of assessment literacy including self-efficacy, confidence and motivation. These findings are discussed in relation to how assessment professional development can be embedded most effectively to support educational reform.

09:30 Technical, vocational and applied assessments III

Room: Studio 7

Lessons learned from working with partners to introduce vocational qualifications with significant project-based assessment in sub-Saharan Africa



Peter Ashton, Senior Education
Contract Manager, Cambridge
Partnership for Education

Education reforms by the Governments of Botswana and Eswatini introduced new suites of examinations focusing on technical and vocational skills as well as academic competencies (education with production). They were to be of equal value to the existing national standards (IGCSE). The reforms aimed to produce learners who were competent in both academic and vocational skills, with appropriate competencies for self-employment, employment and further training. In addition, the new syllabuses should foster entrepreneurial skills and the integration of interpersonal and application skills.

This paper discusses the lessons learned so far from working with the exam boards to develop and implement an assessment model, assessment tools and standards.

Challenges included finding enough teachers with prerequisite subject knowledge, allocating teachers to schools, providing training, encouraging take-up by schools, and securing physical and teaching resources. Schools had to decide which of the existing syllabuses to drop, find and equip space, and enrol students. Such changes risked conflict with teaching unions.

Assessment models were developed which realised the governmental vision and mitigated the challenges.

This was a challenging but worthwhile reform, requiring significant resources and co-operation to enact. Assessments need to be robust and be appropriate to the prevailing capacity, risks and limitations.

14:00 – E-assessment IV

Room: Studio 3

Evaluation of the Cambridge International Digital Mock Exams Service



Martina Kvalja, Research Division,
Digital Assessment & Evaluation,
International Education

Cambridge International now offers a Digital Mocks Service for some international GCSE and A Level exams in preparation for live exams. As part of the iterative design approach, we run a range of research activities throughout the product's lifecycle. After the first launch of digital mocks in January 2023, we were keen to evaluate the delivery of the service, and we used a particular methodological approach for this purpose. We collected assessment and user data from the test platform, validation evidence and user experience data.

The findings that are produced as a result of the data collection and analysis can then be used to demonstrate the extent to which the service's purpose was met, and to inform the continuous development and improvement of the service based on this evidence. The data collection plan required a collaborative approach with different parts of the business carrying out the proposed research activities as part of the iterative design process. This evaluation methodology aims to give the teams a structure for the evaluation so that the purpose of each data collection strand and the research questions are clear and linked to the objectives of the Digital Mocks Service.

16:00–18:00 Ignite symposium

Room: Amphitheatre

Why so many assessments? A holistic framework to help teachers to see the bigger picture – including the missing pieces



Dr Irenka Suto, Head of Assessment,
Cambridge Centre for Evaluation and
Monitoring



Suzanne Crocker, Assessment
Products Integration Manager,
International Education

Presented by Dr Irenka Suto

Data can overwhelm teachers. There are numerous types of assessment, created by government bodies, assessment organisations and schools themselves. It can be challenging to understand the purposes of different tests and evaluations, the pedagogical questions they address, and how they can fit together coherently. Moreover, teachers may need to identify gaps and unwanted overlaps within their assessment approaches.

In this talk we present a teacher-friendly, evidence-based framework of five interacting areas of teacher insight into their learners' journeys. We argue almost all assessments, evaluations and teaching resources can be understood in terms of these areas, and that our framework is a useful organising instrument for creating learner profiles. During face-to-face workshops in Malaysia, Indonesia, Thailand and Vietnam, we introduced the framework to 200 teachers in international schools. They applied it in group activities then gave feedback through questionnaires and plenary discussions. Eighty-six percent reported the framework made a lot of sense.

Creating the conditions for successful assessment reform through education system planning



Brooke Wyatt, Head of Assessment Project Design and Delivery, International Education



Marc Neesam, Head of Professional Development Solutions, Cambridge Partnership for Education

An essential factor for successful assessment reform is continuing professional development for school leaders and teachers. There are numerous cases where assessment reform has caused disruption and unintended consequences because those implementing the reform at the school level of the system are not provided with the assessment literacy required to write effective assessment policies to align with the reform.

A major issue is the lack of focus on the key principles and purposes of assessment at the point of initial teacher education, leaving those entering teaching, and those within the system, feeling unprepared for large-scale change.

This presentation focuses on a strategy to support ministries with professional development to ensure there is alignment between the intended expectations of reform and the system operations, e.g. initial teacher education needed to implement the change. This approach focuses on ministries creating the conditions for success and reducing the opportunity for unintended consequences.

Saturday 4 November

09:00 National tests and examinations III

Room: Amphitheatre

Performance in secondary mathematics topics pre- and post-reform



Joanna Williamson, Senior Research Officer, Assessment Research Division



Carmen Vidal Rodeiro, Senior Research Officer, Assessment Research Division

Presented by Joanna Williamson

A recent major programme of assessment reform in England replaced the General Certificate of Secondary Education (GCSE) qualifications that young people take aged 16. In mathematics, the stated aims of reform were ambitious: to ensure mastery of fundamental mathematics by all students, while also creating a more challenging qualification to improve preparation for further mathematical study and careers.

Research during the first few years of the reformed qualifications identified perceptions of how GCSE Mathematics reform had affected teaching, learning and attainment. To better understand the reform's impact, we decided to complement these studies by offering a quantitative analysis of student performance in GCSE Mathematics examinations. Our research analysed the performance in maths items of approximately 250,000 candidates from the final three years of pre-reform GCSE Mathematics (2014–2016) and the first three years of the post-reform GCSE (2017–2019), focusing on performance in different mathematics topics and how this changed (or did not change). We saw that candidates found post-reform GCSE assessments substantially more challenging. The proportion of marks achieved decreased more in some topics than others, but the variation across topics was not statistically significant. It may nevertheless have affected teacher and candidate experiences of different mathematics topics, and perceptions of the reformed qualification.

09:30 Perspectives of end-users and the general public on assessment II

Room: Studio 2A

Spoken practice: evaluating an external ESOL language assessment for young learners in France



Jonah Champaud, Senior Impact Evaluator, Cambridge English



Brigita Seguis, Head of Impact Evaluation, Cambridge English

As word associations go, ‘external English language assessment’ and ‘French public schooling’ might not be the most obvious pairing, given the relatively high degree of centralisation of assessment design, and public nature of assessment provision. However, since 2018, Cambridge University Press & Assessment has been delivering Cambridge English Qualifications for Young Learners in several primary schools in the state education sector.

We undertook a case study of three state sector primary schools in the Chantilly municipality, north of Paris, to study the impact of introducing our Cambridge qualifications on stakeholder attitudes and practices related to English language learning and assessment. We also aimed to capture unintended effects on the education system at the meso level.

We found that the introduction of an external English language assessment was associated with positive attitudes towards language learning and assessment, high instrumental motivation, and language-use confidence. Stakeholders highlighted the importance of equitable access to high-quality assessment experiences. Nonetheless, we also found that provision of external assessment may encourage further learner exits out of public-sector education if the assessment offer does not comprehensively cover the transition between primary and secondary. The presentation discusses the implications of this case study on future English language policy and assessment reform in France.

10:00 Other II

Room: Studio 2C

An assessment framework for education reform projects



Matthew Dean, Senior Assessment Manager, International Education



Dan Bray, Deputy Director, International Education

For the last 15 years, Cambridge Assessment International Education has collaborated on projects with various ministries, school groups and donor organisations on a wide range of assessment-related reforms.

In this presentation we will present an assessment framework that we have developed for such reform work, outlining and giving the rationale for the levels and sub-levels, and we will detail which specific activities we would expect to take place for each. We will show how the framework can be operationalised and adapted to a specific context and the requirements of the partner organisation, giving examples of specific processes, documents and deliverables that have been used or produced when working with partner organisations.

We will also discuss the key features in terms of collaboration that project evaluation has shown is essential to a successful intervention, as well as some of the challenges that we have encountered when working on different projects globally. We will outline how the framework can be presented and socialised, and consider the factors that may ultimately dictate whether the reform in question is successful.



CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT

Cambridge University Press & Assessment unlocks the potential of millions of people worldwide. Our qualifications, assessments, academic publications and original research spread knowledge, spark enquiry and aid understanding.