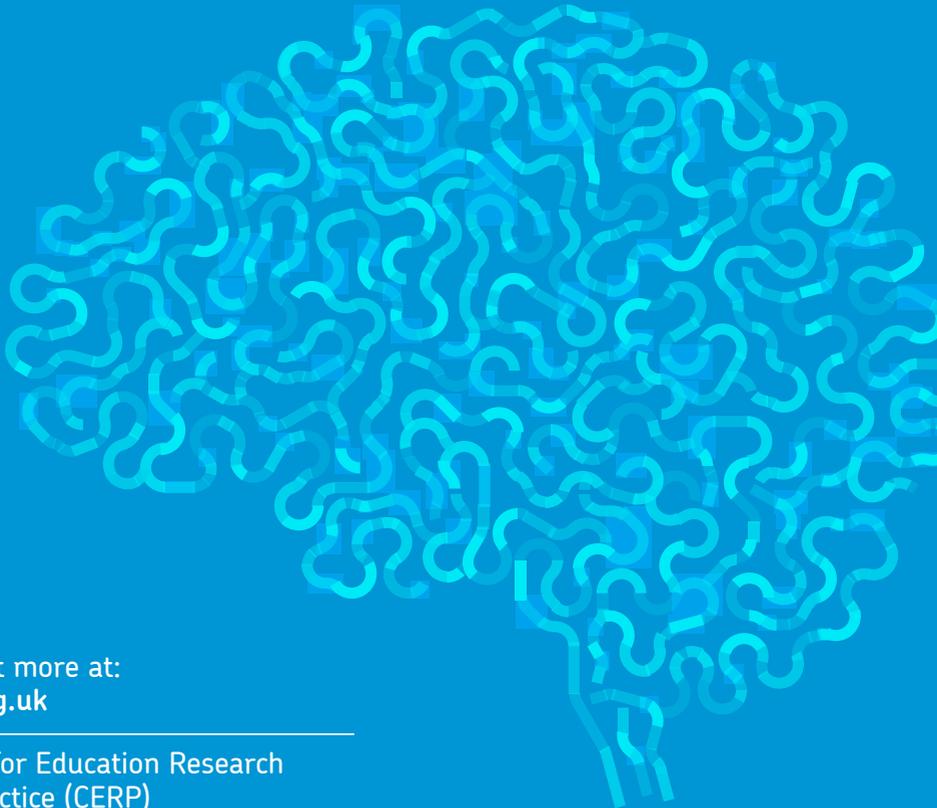


# Centre for Education Research and Practice (CERP)

Impartial and rigorous research



Find out more at:  
[cerp.org.uk](http://cerp.org.uk)

---

Centre for Education Research  
and Practice (CERP)

# About AQA's Centre for Education Research and Practice

AQA's Centre for Education Research and Practice (CERP) is a multidisciplinary unit that specialises in assessment research. Our rigorous analysis and high-quality academic research inform AQA's operational activities and assessment design, and contribute to the wider debate on examinations and education.

Much of our work focuses on improving existing methods of setting and maintaining standards. However, we also explore topics such as validity, marking reliability and the accessibility of assessments. Although our work is grounded in the practical realities of qualifications, we take an interest in theoretical and philosophical thinking about assessment and qualifications.

Our researchers publish in peer-reviewed journals and academic books, and frequently present at national and international conferences. We also regularly undertake joint projects with

partners such as universities and other research units. We are delighted to support the work of AEA-Europe by providing the secretariat for the association.

The theme of this year's AEA-Europe conference – 'Building bridges to future educational assessment' – has inspired us to explore a range of assessment matters. Abstracts for all the work that CERP will be presenting at the conference can be found in this booklet.

CERP will have an exhibition stand in the refreshment area throughout the duration of the event; please do pay us a visit.



**Alex Scharaschkin**

Director of Research and Compliance, AQA  
Executive Secretary, AEA-Europe

## A brave new world of continuous assessment

**Cesare Aloisi and Claire Whitehouse**  
Presentation

In high-stakes qualifications in England, the use of continuous assessment – defined as the aggregation of scores from many standardised and internal assessment occasions to an overall mark – has declined. This is largely due to implementation issues, which have undermined perceptions of the integrity of this type of assessment. Recent reforms have led to new qualifications that have a one-off terminal assessment, with reduced continuous assessment.

This move away from continuous assessment may compromise the validity of some qualifications, particularly in practical or applied subjects, as students no longer need to demonstrate a sustained period of effort on one piece of work. Teacher judgement is also eroded.

However, technological advances in assessment formats and methods for data collection, and storage, may present an opportunity to return to continuous assessment and restore validity.

We present a historical case study that throws light on past uses of continuous assessment in English language and literature qualifications aimed at 16 year olds. We use findings from the research literature to demonstrate how recent technological developments could overcome barriers to the implementation of continuous assessment and improve its integrity.

We go on to consider the wider implications of any potential future shift to continuous assessment, including unintended consequences for teaching and learning, and the safeguards that may be needed to reassure stakeholders about the value of such a transition.

## Building bridges between assessment and subject matter: An innovative approach to assessing GCSE Practical Science

**Yasmine El Masri, Sibel Erduran** and **Alison Cullinane**, Oxford University  
Department of Education; **Ruth Johnson**, AQA  
Presentation

Assessing practical science skills is problematic, largely because of the complex nature of the construct to be assessed. One of the key challenges is to offer students valid contexts in which they can demonstrate their practical skills and knowledge, while ensuring that the assessments are reliable and fair; at the same time, it is important to minimise possibilities for cheating and avoid a drilling style of teaching.

The tensions arising from assessing practical science under high-stakes conditions, such as GCSE examinations in England, have often contributed to narrowing the construct down to experimental work. However, researchers in science education have advocated for the inclusion of a broader set of skills, including observation, classification and argumentation.

This presentation describes a new framework for developing valid GCSE practical science assessments – based on Brandon’s (1994) matrix of scientific methods – that allows for a better representation of the construct. The approach also ensures that assessments are reliable. Task exemplars illustrating the proposed framework will be discussed.

The presentation will be relevant to researchers in educational assessment, science practitioners and examination board personnel.

*This research is carried out as part of Project Calibrate, a partnership between researchers and assessment experts from the University of Oxford and AQA. Project Calibrate is funded by the Wellcome Trust (grant number 209659/Z/17/Z).*

project  
calibrate

AQA



## The big crunch: combining sources of evidence in standard setting

**Dennis Opposs**, Ofqual; **Kristine Gorgen**, Oxford University  
Centre for Educational Assessment; **Lena Gray**, AQA;  
**Tina Isaacs**, UCL Institute of Education  
Discussion group

This discussion group draws on an international project that focused on how standards are set and maintained in national, school-leaving or university entrance examinations around the world. One of the key outputs of the project has been the conceptualisation of standard setting as a mixed methods process and the application of this approach to the practices used in the participating jurisdictions.

Given the fast-evolving landscape of educational assessment and the tremendous context specificity of education policy and practice, it is no surprise that the ways in which different sources of evidence are combined in standard-setting processes have not received more attention. However, educational assessment can take inspiration from the growing body of literature about

mixed methods research designs in the social sciences.

We propose a way to conceptualise standard-setting processes as mixed methods techniques and thus help to better explain the integration of the range of data sources used in the processes. We present this new conceptual approach and give examples of its application in Europe and around the world.

The group discussion will test and further develop the idea of conceptualising standard setting as a mixed methods process. We anticipate that participants will be able to critically reflect on and better understand how different assessment systems set standards.

This discussion will be relevant to researchers, policymakers, and practitioners interested in standard-setting processes.

## Reviewing assessments and dense data: towards effective evaluation of assessments for the non-technically minded

**Jonathan Powell**

Presentation

Examination boards have an increasing amount of rich data available to them, including detailed information on the performance of assessments. It is important that this performance data can be communicated effectively to different groups of people. A key conversation is between the question writers, who are the subject experts, and the assessment experts, who ensure that exam papers meet regulatory requirements and design principles.

This paper focuses on the review stage of the exam paper cycle. AQA has, historically, used a static Question Paper Functioning Report (QPFR) to evaluate how an exam paper has performed. The impact of these reports, and engagement with them, has been varied.

In light of qualification reform and structural changes within AQA, the process for reviewing the performance of assessments, after an examination series, has been revised. A new focus on continuous monitoring has led to the creation

of more tailored data solutions to improve understanding of issues and promote effective conversation.

This paper describes an alternative approach to how data is presented to question writers in order to encourage data-driven decisions. Case studies are used to show how the visualisation of data has been adapted and refined.

A qualitative method was employed to evaluate the effectiveness of the new approach compared to the original QPFR. The use of data visualisations within different specifications are considered, including the effect that these have on evaluating assessment performance.

## Examination malpractice analysis: defensibility, clarity and complexity

**Steve Wooding**

Poster

The number of malpractice cases handled by examination boards in England has increased rapidly over the last few years, with many cases involving social media, mobile devices, and other 21<sup>st</sup>-century technologies. This has necessitated a more systematic approach to the provision of statistical support for malpractice investigations, in order to help defend the integrity and fairness of the examinations system.

However, there is a need to balance analytical complexity and power with defensibility and clarity, given that those alleged to have been involved in malpractice need to be able to access and understand any quantitative evidence.

We present a core, two-stage approach for malpractice cases that require statistical support; the approach is based on a pairwise cross-referencing of 'same candidates, different

event' and 'different candidates, same event', and uses z-score divergences and Coarsened Exact Matching (CEM).

The application of CEM – a technique more usually employed in medical, political and social sciences – allows us to compare 'like-for-like' candidates to produce more robust detection of anomalous groups or centres. It also enables us to negate possible claims of bias against already high-performing pupils or schools, and deliver a supportably impartial analysis.



## The CERP team at AEA-Europe



### Cesare Aloisi

Cesare joined CERP in July 2017. Prior to this, he worked at the University of Reading on a project analysing student learning trajectories, critical thinking, engagement and wellbeing. He has a PhD in Education from the University of Durham and an MA in Educational Assessment from University College London. Cesare is a former language teacher and has an interest in large-scale assessments, multilevel modelling, early childhood education and social justice issues.



### Lena Gray

Lena was appointed Director of Research in May 2017, having joined CERP as Head of Research in July 2014. Lena is an Honorary Norham Fellow of the Oxford University Department of Education and is currently collaborating with members of the department, IOE and Ofqual on a major international project investigating standard-setting approaches in a range of jurisdictions around the world. Before joining AQA, Lena was Head of Service, Policy, Assessment, Statistics and Standards at the Scottish Qualifications Authority, where she worked for a number of years.



### Ruth Johnson

Ruth joined CERP in June 2015, having spent five years in the AQA English team. She has 15 years' experience as a secondary English teacher, including five years as an assistant headteacher. Ruth completed a Doctorate in Education (EdD) at the University of Manchester's Institute of Education and has a BA (Hons) in English from the University of Cambridge. She has a particular interest in the relationships between policy, assessment and school-based practices, and the assessment of skills that are difficult to measure.



### Jonathan Powell

Jonathan joined AQA's Assessment Design team in 2016. He previously worked as a management information officer at the University of Manchester, and he has six years' experience as a secondary science and mathematics teacher. Jonathan has a BSc in Chemistry and an MA in Education, both from Durham University, and is currently studying for a BSc in Psychology with the Open University. He is interested in educational misconceptions, thinking and learning, assessment design and data communication. Jonathan is currently working on projects concerned with social mobility and improving assessment design through better data communication.



### Alex Scharaschkin

Alex was appointed Executive Director of Research and Compliance in November 2015, having joined AQA as Director of CERP in July 2014. He was previously Director for Regulation, Consumers and Competition at the National Audit Office (NAO), where he led the NAO's work examining the government's use of markets in the private and public sectors. Alex has a background in assessment research: he was Principal Officer for Statistical Analysis at the Qualifications and Curriculum Authority, and held research posts at the Associated Examining Board and the University College London Institute of Education. Alex also served as a member of CERP's advisory group for four years and is currently Executive Secretary of AEA-Europe.



### Claire Whitehouse

Claire joined CERP in 2004, having spent two years in qualification management within AQA. Her current research is focused on improving the validity and reliability of assessments. Claire is working with colleagues to investigate the interactions of students, examiners and teachers with assessment materials and the assessment process, and to explore how new technologies may influence the nature of assessment. Claire gained a BSc(Eng) from University College London and a PhD from the University of Cambridge, both in chemical engineering.



### Steve Wooding

Steve joined CERP in July 2016, having spent three years as a secondary school science teacher. He holds a BSc in Physics and Materials Science and a PhD in Computational Physics. Steve has previously worked in the areas of process engineering, data science, competitive intelligence and knowledge management, psychometrics, staff development, and organisational communication and change. Steve is a keen advocate for the translation of research findings into useful and actionable information that is accessible to a wide audience. As a physicist, he has a particular interest in science education and assessment.



### Cassy Quinn

Cassy is AQA's Event Coordinator. She is responsible for organising exhibition stands and promotional materials for a range of educational conferences.



### Daryl Stevens

Daryl is a data analyst/developer within CERP. His specialist knowledge is used to improve the ways in which CERP approaches computing, including how we develop software and process data.



### Linda Wye

Linda is a senior research administrator and has worked for AQA for 16 years. As well as providing PA and administrative support for various managers, she is involved in event administration on behalf of CERP.

Find out more at:  
[cerp.org.uk](http://cerp.org.uk)  
 Centre for Education  
 Research and Practice

