

# 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)



# Hello

AQA's Centre for Education Research and Practice (CERP) provides robust evidence that informs both organisational direction and wider educational debate.

CERP is a multi-disciplinary research facility with sites in Manchester, Guildford and London. We have a record of high-calibre research that stretches back 40 years through our predecessor bodies. While we maintain exacting standards of academic rigour – our work is reviewed by a prestigious committee of national and international researchers, chaired by Jannette Elwood of Queen's University, Belfast – current research is grounded in the practical realities of assessment and qualifications.

Our team comprises statisticians, psychologists, educationalists and scientists. These varied backgrounds allow us to employ a range of qualitative and quantitative methodologies, which ensure our findings are credible and valid.

We are passionate about education, particularly high-stakes qualifications, and we share our work with the

academic community, policy makers, teachers and specialist media. We regularly publish research papers, blogs and longer articles via our website ([cerp.org.uk](http://cerp.org.uk)), which can be freely accessed.

The AEA-Europe conference is a welcome opportunity to present our research. This year's focus on social justice is of particular interest in that it allows us to showcase some of our recent work on aspects of fairness in qualifications, including Ruth Johnson's study of the language of GCSE English assessments, work by CERP's Ben Jones and Louise Banahene of the University of Leeds on the Extended Project Qualification, and research by CERP's Neil Stringer, Michael Chan and Philip Chan of Sheffield Medical School on the effects of contextual factors on attainment in medical school.

We are always keen to discuss collaborative research and look forward to many fruitful exchanges in Glasgow.

## Alex Scharaschkin

Director, Centre for Education Research and Practice

## The meaning of assessment standards in Scotland and England

**Jo-Anne Baird**, Oxford University Centre for Educational Assessment and **Lena Gray**, CERP

Fairness in assessment standards is a component of social justice. The ways in which educational standards are conceptualised and operationalised differently across nations has not been given sufficient attention. Different education systems and examinations operate in Scotland and England, and the stated value positions and processes relating to examination standards differ markedly. This paper depicts Scotland's Curriculum for Excellence qualifications as a case, discussing how Scotland defines standards for these public examinations. This will be contrasted with the view of examination standards in England.

The presentation will critically examine public positions on assessment standards in Scotland and England through the lens of recent theories of standard setting. In analysing public statements on standards, we will show that different conceptions of fairness underlie debates about standards in the two countries. We will use analysis of these cases to suggest how our conceptual framework of standard setting might be further developed. This should be of interest to researchers, policy makers and practitioners who are interested in assessment standards and how they differ across education systems.

## The Extended Project Qualification (EPQ): an equitable qualification in an unequal system?

**Louise Banahene**, University of Leeds and **Ben Jones**, CERP

The Extended Project Qualification (EPQ) is an increasingly popular option among 16-19 year olds. Its success – in theory, at least – is more dependent on the student's imagination and determination, and less on teaching and resources.

The EPQ remains optional and there is little research on the differential outcomes by type of school or socio-economic status. Conversely, a great deal of evidence exists in relation to the impact of 16-19 qualification choices, social background and degree outcomes (HEFCE, 2013;

Crozier, G and Reay, D, 2008; Crawford, 2012). Our study focused on University of Leeds entrants who had studied an EPQ. We split the data by school type, gender, broad subject area of study and socio-economic class. The intention was to develop understanding of the relationship between background, EPQ attainment and end-of-year degree outcomes.

The paper provides contextual information regarding the EPQ, including its growth in popularity and apparent beneficial impact on students' concurrent A-level grades. It also reports on initial analyses to investigate whether, and by how much, the EPQ is a more socially equitable qualification than its alternatives or equivalents, particularly A-levels.

## **GCSE English: powerful knowledge or the knowledge of the powerful?**

**Ruth Johnson**

Social inequality in the UK, and specifically in England, is increasing while social mobility is decreasing. The relationship between family background and academic success is well-documented, as is the strong correlation between academic success and future income. This paper reveals one means by which the current curriculum and assessment structures perpetuate inequitable educational outcomes. It reports on a case study that suggests that the current GCSE English assessment can position students as outsiders or insiders, depending on their cultural dispositions and experiences, and in doing so, privileges students from dominant social backgrounds.

Data used includes interviews with students, teachers and senior examiners of GCSE English, as well as analysis of assessment documents. The research engages with Young's (2008, 2011, 2013) theory of knowledge, which argues that the school curriculum ought to comprise 'powerful knowledge', and suggests that English as it is currently taught and assessed is distinctive from other subject areas in the way it privileges the context-dependent 'knowledge of the powerful'. The paper broadly suggests a way forward for producing fair and valid context-independent assessment in English that structures powerful knowledge.

## Improving the maintenance of standards in England: comparative judgement and item-facility test equating

**Kate Kelly, Charlotte Stephenson and Neil Stringer**

In English national examinations, vagaries in paper difficulty over time are ironed out through adjusting the grade boundaries. This ensures that candidates of the same ability receive the same grades regardless of when they took their exams. However, procedures have limited capacity for quantifying changes in difficulty.

A potential improvement lies in comparative judgement (CJ). CJ allows construction of a measurement scale through the iterative presentation of pairs of stimuli to a pool of judges, who must choose which of the pair most reflects a given criterion. It can thus be used to directly assess paper difficulty.

To be practicable, we need a method for relating changes in difficulty to changes in the boundary. Curcin, Black and Bramley (2009) have demonstrated how this can be done for dichotomous items, but their approach does not readily extend to polytomous items, which predominate in English national examinations. Item-facility test equating is one possible solution to this problem but is as yet untested. This poster presents both the method of item-facility test equating and a study proposed to evaluate it. The study uses A-level Chemistry past papers to compare the grade boundaries estimated using item-facility test equating against the actual grade boundaries.

## Active monitoring of on-screen marking using process capability indices

**William Pointer** and **Yaw Bimpeh**

Expert judgements are used in scoring in many domains, for example in high-stakes examinations such as GCSEs and A-levels. Candidates' examinations are marked by examiners who score each question using a set of criteria to judge how well a candidate has performed. The essential issue is that

the scores assigned by the examiners are accurate.

The increase in on-screen marking opens up possibilities for active monitoring of quality of marking, which has the potential to improve marking reliability. The marks awarded

by examiners are closely monitored throughout the marking session to ascertain whether their scoring maintains desirable levels of accuracy.

Statistical process control measures such as process capability indices have been widely used in industrial settings for well over 20 years. Process capability indices are measures used in quality control to provide information

on the extent to which a process can produce outcomes that conform to a set of specifications.

In this study we propose using a marker capability index, which can estimate markers' rate of non-conformance to the mark scheme, to actively monitor the quality of marking to ensure that candidates get accurate and fair results.

## **Spatial analysis of Key Stage 4 educational outcomes in England**

**Ben Smith and Yaw Bimpeh**

Exam results are increasingly used as school accountability measures in England. However, the contextual factors that can influence exam results are often overlooked. A grammar school that admits only very able pupils would be expected to outperform non-selective schools. There are also suggestions that the location of schools can influence attainment – for example, schools in London outperform comparable ones elsewhere in the country.

Using data from the Department for Education's school performance tables, we investigated how contextual factors related to attainment in schools with a highly disadvantaged cohort. Exploratory

models found that the variables that best predicted the 'average points score' of pupils were: the percentage of sessions recorded as absence, and the percentage of pupils for whom English is an additional language.

A geographically weighted regression was performed to determine whether the relationship between these variables varied across England. A range of regional differences emerged, with absence proving a stronger predictor of attainment in the North-East and around London than in other regions. While further work is needed to verify findings, spatial analysis has potential as a tool for analysing educational outcomes.



## Do contextual factors at the time of admission influence educational attainment on the medical course?

**Neil Stringer**, CERP; **Michael Chan** and **Philip Chan**, Sheffield Medical School

The educational attainment of young people is affected by the social and educational contexts in which it occurs. If such effects are not acknowledged, certain groups of students are likely to be disadvantaged when applying for degree courses, including medicine. It is less well established whether these factors affect performance once admitted to medical school. This study investigates the effects of socio-economic status and schooling on the academic attainment of a cohort of medical school students.

The independent variables were the: students' GCSE results; A-level results; Income Deprivation Affecting Children Index (IDACI) rank; and the percentage of A-level students at their school achieving three A-levels

at AAB or higher (in two or more facilitating subjects). The dependent variable was the sum of students' summative assessment scores over the previous four years of medical school: the basis of the medical school's submission to the UK Foundation Programme.

The results confirmed the predictive validity of prior academic attainment, documented in previous studies. There was no evidence that socio-economic background affects course performance, but students admitted to medicine from poorly performing schools achieve higher academic attainment on the course than students admitted from better-performing schools. Schooling could be taken into account for admissions purposes.

# The CERP team at AEA-Europe



## Anton Béguin

Anton became CERP's Director of Research and Innovation in 2014. He is also Director of Research at the international measurement and assessment organisation Cito. Anton has been involved in many research projects related to standard setting, equating and application of multi-dimensional item response theory. He has worked as a consultant with a range of national and international organisations, and has written many publications on examinations and measurement. He is the Chair of the Methodology and Evaluation division of the Netherlands Educational Research Association.



## Yaw Bimpeh

Yaw joined CERP as a senior researcher in September 2014. He holds a PhD in Statistics, an MSc in Mathematical Sciences and a BSc (Hons) in Mathematics. His current areas of research include marking reliability, application of the Bayesian method to standard setting, and methods for detecting exam security breaches. Yaw has experience of analysing and modelling data in a variety of fields, and is skilled in the research and application of statistical methods. He has also taught statistics and mathematics to undergraduate students.



## Lena Gray

Lena joined CERP as Head of Research in July 2014. Prior to this, Lena held a number of positions in the SQA and its predecessor organisations, latterly as Head of Policy, Assessment, Statistics and Standards. She also has experience as a secondary teacher and as a course leader and tutor at the University of Strathclyde, where she completed her PhD.



### **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 head teacher. Ruth recently completed a doctorate in education (EdD) at the University of Manchester's Institute of Education; her research project examined the relationship between policy, practice and assessment in relation to GCSE English. Ruth obtained her BA (Hons) in English from the University of Cambridge.



### **Ben Jones**

Ben joined the Joint Matriculation Board, a predecessor of AQA, in 1990 and has held a variety of research posts. As Head of Standards, Ben is responsible for overseeing the awarding process and for ensuring confidence in AQA's awards. Ben represents AQA on several inter-awarding body committees. Before joining the organisation, Ben spent time in Tonga as an educational assessment government adviser, and was a research associate in the Division of Education at Sheffield University. He has a BA in Economics and an MSc in Social Research Methods.



### **Kate Kelly**

Kate joined the team in June 2010 after completing a BSc (Hons) in Psychology at the University of Bath, which included a year spent with CERP as a placement student. Kate has been involved in a number of research projects, including looking at the validity and comparability of the Functional Skills Mathematics qualifications, examiners' job satisfaction, and the effectiveness of a reading intervention for primary school children. She is currently investigating the potential of comparative judgement for improving grading decisions.



### **William Pointer**

William joined AQA in 2010, having graduated with a Masters in Mathematics from the University of Bath. He initially worked as a Qualification Developer in the GCSE Science department before joining CERP in August 2013 to pursue a career in research. William's current research focuses on the quality of marking: how we measure and monitor marking reliability, and how we can improve it in the future. He is also interested in research on standards and comparability.



### **Alex Scharaschkin**

Alex became Director of CERP in 2014, having been a member of CERP's advisory group for four years. He was previously Director for Regulation, Consumers and Competition at the National Audit Office (NAO) in London, 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 at London University's Institute of Education.



### **Ben Smith**

Ben graduated from the University of Birmingham in 2014 with an MSci in Psychology and Psychological Research. He briefly worked as a research assistant in a West Midlands prison, before joining CERP in September 2014. Ben's research activities largely focus on performance statistics for AQA's specifications: assessing their impact, and identifying improvements. He has also been looking at the effect on standards of the shift from modular to linear exams, and at the gender attainment gap in AQA's assessments. Ben is currently working with Yaw Bimpeh on marking reliability.



### Charlotte Stephenson

Charlotte graduated from the University of Manchester in 2010 with a BSc (Hons) in Psychology. She joined CERP as a research assistant in 2014, having previously worked on the Future of Electronic Assessment (FEA) project in AQA's Operations division. Charlotte's research activities have included a study into the effects of enhanced team leader feedback on marking reliability and examiner satisfaction. She is currently investigating whether comparative judgement estimates of question difficulty can be used to set grade boundaries for GCSE and A-level examinations.



### Neil Stringer

Neil joined CERP in April 2005 from the Department of Psychology at the University of Surrey, where he was a postdoctoral Research Fellow in the field of applied human vision. As Principal Research Manager, Neil has responsibility for research in the area of assessment quality. His own current work is a mixture of operational research as well as broader research and policy work, for example on standard setting procedures and admissions to higher education.

Find out more at:

[cerp.org.uk](http://cerp.org.uk)

Centre for Education  
Research and Practice



Celebrating research conducted by AQA and predecessor bodies, on the 40th anniversary of the AQA Research Committee

