



Pearson School Assessment
and Qualifications
Research Papers at
AEA-Europe 2024
Paphos, Cyprus

Presenters.....	2
Poster.....	1
Onscreen Functional Skills: Insights from a decade of delivery of English and mathematics assessments	1
Open Papers	2
Is there any evidence of the saw-tooth effect impacting on learner performance where assessments are more skills & vocationally based? Analysing outcomes data overtime across a range of qualifications/ assessments	2
Machine Learning Modelling: Prediction of Mathematics GCSE 2023 results using 2022 Mock exam outcomes	3
Gaining insights and understanding: School and student perspectives of taking onscreen high-stakes assessments	4
Exploring the relationship between students’ use of digital technologies and their performance in digital PISA 2022 mathematics assessments.....	5
Have writing skills been left behind? Understanding current practice in teaching Writing in English Primary and Secondary Schools.....	6
Establishing modal effect in high stakes assessments: Findings and recommendations for data collection and methodology based on a trial of paper vs onscreen assessment of GCSE English.....	7
Functional Skills Qualifications: Investigating shifts in demand for onscreen and on-demand maths and English assessments in England after over a decade of delivery	8
Discussion Group.....	9
One size doesn’t fit all: How to consider the equity and fairness of access arrangements as we move to digital modes of delivery	9

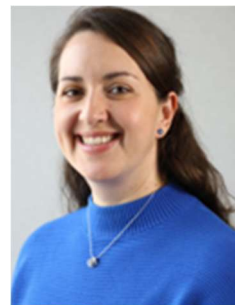
Presenters



Blake Ashworth
Assessment Research
Analyst



Emma Crampton
Learning Experience
Designer



Irene Custodio
Digital Assessment Design
Lead



Hayley Dalton
Head of Vocational and
Assessment Research



Alistair Hooper
Delivery Lead, Int. Large
Scale Assessments



Jagdeep Kaur
Head of Vocational and
Digital Standards



Kevin Mason
Senior Assessment
Researcher



Sebastian Nastuta
Data Scientist



Meredith Reeve
Head of Next Generation
Assessment

Poster

Onscreen Functional Skills: Insights from a decade of delivery of English and mathematics assessments

Hayley Dalton, Jagdeep Kaur

Functional skills qualifications in English and Mathematics (FSQs) are taken in schools, colleges, private training providers and other settings in England. Designed primarily for learners that failed to achieve a 'good' pass in English and mathematics at school, FSQs are intended to be an 'accessible and practical route for students who want to develop and improve their skills.'¹

Using data from a large awarding organisation in England with around two-thirds of the market, this poster looks at the journey from the introduction of FSQs to the present day. We reflect on the shifting demand for paper and onscreen tests across the various learner and centre characteristics.

The introduction of onscreen assessment has allowed centres and learners to personalise their assessment experience to a degree. We look at the trends and patterns of the centres and learners who are taking FSQs and whether that can provide valuable insights into the demands and needs going forward for onscreen assessments more broadly in the UK educational landscape.

¹ Dept for Education (2018). Functional Skills English and mathematics subject content: Government consultation response. <https://www.gov.uk/government/consultations/maths-and-english-functional-skills-revised-subject-content> (accessed April 2024)

Open Papers

Is there any evidence of the saw-tooth effect impacting on learner performance where assessments are more skills & vocationally based? Analysing outcomes data overtime across a range of qualifications/ assessments

Jagdeep Kaur, Blake Ashworth

Thursday 7 November, 3.30pm, Aphrodite A

There is documented evidence and research that in high stakes assessments, such as GCSEs and A levels in England, when qualifications are reformed there is generally a small dip in performance because teachers and students are less familiar with the requirements of the new content and assessments. This is termed the sawtooth effect. However, there is little research into whether a similar effect is observed in vocational qualifications which have a more blended mix of both academic, skills and vocational focus.

This piece of research explores whether vocational technical qualifications observe similar dips in learner performance during reform as other qualifications, focusing particularly on internally assessed assessments that have a stronger emphasis on the assessment of skills. The aim of the research is to start to understand the impact reforms in England in vocational qualifications may have on learner outcomes during the early years of delivery. This is to ensure that students are not disadvantaged to their peers studying A levels or T levels, as many will be competing for University places.

Both quantitative unit level outcomes data from historical years are used as well as qualitative evidence from centres/teachers on what they observe.

Machine Learning Modelling: Prediction of Mathematics GCSE 2023 results using 2022 Mock exam outcomes

Sebastian Nastuta

Thursday 7 November, 3.30pm, Aphrodite B

This paper presents a novel approach to predicting student performance in summative examinations. It evaluates the extent to which and the accuracy with which future exam results can be predicted using mock results, a common practice in the UK. This research can be crucial for schools, providing valuable assistance for high-stakes exam preparation.

Using a supervised Machine Learning approach, we leveraged item-level performance data from GCSE Mathematics Higher Tier June 2022 to predict outcomes in June 2023 for students who took the 2022 papers as mock exams. Machine learning was employed to train and develop several grade prediction models using item-level results from all candidates who sat the Mathematics Higher in June 2022.

Between November 2022 and January 2023, 633 candidates took those papers as mock exams. Their performance in this formative assessment was meticulously analysed to predict their future performance in the June 2023 exam series. By comparing the predicted grades with the actual grades for 524 matched candidates, we have robustly concluded that different machine learning algorithms (OLS regression, logistic regression, support vector machine, decision tree, K-Nearest Neighbours (KNN), etc.) provide 80% and 95% accuracy predictions.

Gaining insights and understanding: School and student perspectives of taking onscreen high-stakes assessments

Irene Custodio and Meredith Reeve

Friday 8 November; 9.00am; Zeus

Within a UK context, the potential delivery of onscreen high-stakes examinations presents the assessment community with an exciting opportunity to innovate the current educational landscape. Ensuring that the design and development of digital assessments is as valid, fair and fit-for-purpose as possible is multifaceted. One vital aspect is ensuring opportunities to listen to and respond to feedback from students in relation to their assessment experiences. Building on previous research, this paper reports on the assessment experiences of learners in multiple geographies who in 2024, have taken Onscreen International GCSEs across different subjects. Our findings are informed by surveys completed after onscreen practice (mock) and live assessments, focus groups and case studies of schools who have delivered onscreen high-stakes assessments. This supports our iterative approach to gathering data and making improvements to our onscreen assessments and building positive user experiences. We were able to gather feedback on students' interactions and experiences of the assessment as well as how students interact with the digital components of the onscreen platform. It is crucial not to underestimate these perspectives, especially of the students themselves and their test-taking experiences helping to ensure valid, reliable and fit-for-purpose assessments.

Exploring the relationship between students' use of digital technologies and their performance in digital PISA 2022 mathematics assessments

Irene Custodio and Sebastian Nastuta

Friday 8 November, 9.30am, Leda

The relationship between digital skills and achievement in digital assessment continues to be of significant interest to policy makers and educators as digital resources are increasingly used in classrooms and assessments. Previous research has yielded mixed findings regarding the impact of ICT (Information and Communication Technology) use on students' academic performance, suggesting that this relationship may be influenced by various factors including the purposes and quality of ICT use, as well as students' interests, attitudes, confidence, and competencies. This research explores the relationship between different ICT factors and student performance in PISA digital mathematics assessments.

Mathematics was chosen specifically as it was the major domain for PISA 2022 and contain items that move beyond 'paper behind glass' representations. We make use of the comprehensive PISA 2022 ICT questionnaire to investigate this relationship. Our methodological approach uses the PISA ICT questionnaire to develop 14 composite constructs that represent various aspects of ICT usage at school and home, as well as students' interest, self-efficacy, and competencies in digital technology. These scales are then utilised in subsequent correlation and multilevel analyses to determine the impact of ICT on students' performance in the digital mathematics items taken by 15-year-old students in England.

Have writing skills been left behind? Understanding current practice in teaching Writing in English Primary and Secondary Schools

Alistair Hooper, Grace Grima

Friday 8 November, 11.00am, Aphrodite B

There is a general agreement in the literature that there is less evidence about teaching and learning of Writing than about Reading. In international studies such as PISA and PIRLS, Reading is used as a proxy measure for literacy, and Writing is not included in the assessments. In England, there has been little progress since the 2012 Department for Education research report 'What is the Research Evidence on Writing?' which cited a lack of evidence as to why students perform less well in Writing in comparison to Reading and other core subjects, and a lack of understanding of the effectiveness of specific interventions with struggling writers or effective strategies for teaching spelling.

In order to investigate 12 different research questions, a mixed methods approach was adopted, collecting data via a survey of 743 practising teachers from Primary (n=391) and Secondary State schools (n=352), and 57 semi-structured interviews were carried out with survey respondents.

Key findings were the lack of training in writing for teachers, low student confidence and motivation for writing, low teacher confidence in teaching and assessing creative writing but also fundamental skills such as handwriting, spelling, punctuation and grammar, and no established interventions used for struggling writers.

Establishing modal effect in high stakes assessments: Findings and recommendations for data collection and methodology based on a trial of paper vs onscreen assessment of GCSE English

Kevin Mason, Sebastian Nastuta

Friday 8 November; 1.45pm; Zeus

In England, we are yet to see a significant transition from paper-based to digital assessments for school-based national assessments for the General Certificate of Secondary Education (GCSE). To corroborate the validity of digital assessments, a comprehensive programme of research is being undertaken to ensure confidence that the move to majority-digital assessment will not introduce new sources of unfairness into the education system. The present study is a part of this programme, examining whether the mode of assessment for GCSE English Language introduces construct-irrelevant variance in assessment outcomes.

GCSE English language consists of two assessments. 1734 students across eight schools were recruited and split into two groups. The first group took Paper 1 digitally and Paper 2 on paper; the second group took Paper 1 on paper and Paper 2 digitally. Background information relating to the students was collected. Three analysis methodologies were applied to the data: Ordinary least squares regression, multilevel modelling, and differential item functioning.

This paper describes the hurdles that needed to be overcome to obtain such a significant sample of students, and how these challenges impacted on the methodologies that were used, and the confidence in the outcomes and lessons to be learned for future work.

Functional Skills Qualifications: Investigating shifts in demand for onscreen and on- demand maths and English assessments in England after over a decade of delivery.

Jagdeep Kaur, Hayley Dalton

Friday 8 November, 2.15pm, Zeus

Using entry data obtained from a large awarding organisation, with around two-thirds of the market for Functional Skills Qualifications (FSQs) in England, we use a mixed method approach to explore how demand for these qualifications has changed over time. With a focus on onscreen and paper-based delivery, we consider how distinct types of providers and different students are accessing different assessment modes. Through this analysis, this paper sets out the reasons for the changing demand and asks if we can generalise these findings to inform future demand for onscreen assessment.

We explore the data that show a shift in demand across time from younger to older learners and from paper to onscreen and back to paper. On face value the data show a decline in demand for FSQs, but underneath that we use qualitative data to explore further how providers have evolved their offer. Themes explored will include funding, policy drivers, availability of assessment and provider resourcing. In conclusion we consider to what extent we can use these findings to inform future planning for the provision of onscreen assessments for other types of qualifications and high stakes assessments

Discussion Group

One size doesn't fit all: How to consider the equity and fairness of access arrangements as we move to digital modes of delivery

Emma Crampton, Irene Custodio, Ellen Barrow

Thursday 7 November, 1.30pm, Aphrodite

As we transition further towards digital modes of high-stakes assessments, there are opportunities to consider current provision and to think about the different types of access arrangements that may be required to support learner's needs.? Interesting questions emerge about equity and fairness and whether access arrangements level the 'playing field' or may present an unfair advantage to some students.

In this discussion group, we explore questions around current access arrangements in the UK and how they may need to change as we move towards more digital modes of delivery:

- How do we work together as an assessment community to create clear guidance, that is equitable and fair for all students across formats?
- As technology develops, how assessment organisations work together to agree a common approach to how technology updates affect what is available to students as part of normal ways of working and to inform access arrangements?
- Should all students have access to the same accessibility tools?

We hope to build participants' awareness of the different aspects of equity and fairness that need to be considered when trying to address and balance regulatory requirements, access arrangements and students needs in digital high-stakes assessments.