Pearson

GCSE Mathematics Qualification — UK

Regulated Qualification Efficacy Report

March 2019

Contents

Introduction

Pearson's mission is to help people make progress in their lives through learning. But helping people to achieve the learning outcomes that matter most in life, like new knowledge and skills to support progression into further or higher education, or in a career, isn't something that happens by accident. It happens by *design*.

When we first start our learning journey, the choices of parents and educators often drive decisions about learning. As we grow, we take over from parents and educators, becoming the designers of our own lifelong learning journey. We identify the outcomes we want to achieve the most, and select learning experiences that suit how we want to access and engage with learning.

At Pearson, we are committed to supporting you to achieve the outcomes that matter most to *you.* That's why we design products focused on supporting the achievement of those outcomes, why we underpin the design and implementation support with evidence about what works to improve teaching and learning, and why we measure the impact of use of our products on outcomes. We use what we learn to continuously improve how our products and services are designed and used.

This is what our commitment to efficacy means, and why it is at the core of our mission as a company. It brings rigour to how we break down big learning goals into manageable steps, and focus and alignment to how we combine our world leading content, assessments and qualifications, and technology capabilities into learning experiences.

We also publicly report on the efficacy of use of our products. Our reports help us, and the wider education community, build a better understanding of not just what works, but how, why, and in what context — helping us learn, not guess, about how the design and use of products relates to the achievement of outcomes that matter most to you.

The 2019 Product Efficacy Reports include three audited, standards-based efficacy research studies on: Revel for *Psychology*, 1st edition by Marin and Hock in North America, MyPedia in India and Sistema COC in Brazil. We are simultaneously publishing non-audited efficacy reports on two of our most frequently used assessment and qualifications products — Pearson Test of English Academic and the UK regulated GCSE Maths qualification.

We remain committed to continuously improving how we are applying efficacy in education, all with a focus on helping more people make progress in their lives through learning. This sense of purpose gives us a reason to keep on fighting, nothing spared, to improve how we do things in education.

Kate Edwards
Senior Vice President,
Efficacy and Learning Research, Pearson
March 2019

Special thanks

We want to thank all the educators, students, research institutions and organisations who have collaborated with us. Of course, our work isn't possible without partnerships within Pearson and the support of our leadership.

If you are interested in partnering with us on future efficacy research, have feedback or suggestions for how we can improve, or want to discuss your approach to using or researching our products, we would love to hear from you at **efficacy@pearson.com**. If we, as a sector, tackle this together, we will help more learners, learn more.

About efficacy reporting at Pearson

To be as open and transparent as possible about how we design, develop, and evaluate the impact of use of our products on learning, we produce efficacy reports. We have two main types of efficacy report: audited and non-audited reports.

About audited product efficacy reports

To find out more about our audited Product Efficacy Reports, go to **pearson.com**.

About non-audited qualifications efficacy reports

The reports on our assessment products are not externally audited. This is because the auditing framework is organised around learner outcomes, and we do not typically expect our assessment products to have a direct effect on learner outcomes. Instead, we focus on ensuring that our assessments are valid, reliable, and fair. To find out more about our non-audited assessment efficacy reports, go to **pearson.com**.

We have some products and services that do not fit into this categorisation and are not currently subject to independent assurance as described in our **Efficacy Reporting Framework**. Our UK qualifications are one example of that. This particular report focuses on one of those qualifications: Pearson Edexcel Level 1/Level 2 GCSE (9–1) in Mathematics (1MA1). This is an example of a General Certificate of Secondary Education (GCSE) qualification.

About regulated qualifications efficacy reports

UK qualifications encompass both subject content that must be taught and assessments aligned to that subject matter. A qualification is, therefore, more than a stand alone assessment.

Pearson offers regulated GCSE (9–1) qualifications as one of **five Awarding Organisations (AOs)**, operating under the supervision of the Office of Qualifications and Examinations Regulation (Ofqual), which reports to the UK Parliament.[1] AOs design, develop, deliver, and award the recognition of learning outcomes (knowledge, skills, and/or competencies) of an individual. They follow an assessment and certification process that is valued by employers, learners, or stakeholders.

Alongside the assessments and associated materials, Pearson seeks to support teachers to deliver the qualification effectively, and to provide their students with the best possible learning experience at this level. We do this by creating comprehensive and holistic teaching and learning resources that are well aligned to the learning aims of the qualification.

The process for the design and development of qualifications in England is highly regulated by Ofqual. This is to ensure that our approach to qualifications has the necessary integrity. That said, given the composition of a qualification, as both defined subject matter and aligned assessment, we believe that there is also value in undertaking research on the direct effect of a qualification on learner outcomes. As a result, we are also partnering with world leading education researchers, such as those at University College London (UCL) Institute of Education, to conduct research on the learner outcomes.

This research has focused, initially, on the qualification implementation experience and progression from the qualification. We have further research planned to explore the impact of the qualification on specific knowledge and skills learners have gained as a result of taking the qualification.

The research on the impact of our qualifications is undertaken to meet the standards expected for publication in peer-reviewed academic journals. We are exploring how we would approach adding a further level of assurance in the form of independent auditing by PwC and independent verification and peer review by SRI Education. We are doing this because we believe there is huge value to our customers and to learners in using our efficacy reports to share openly and transparently how we design, develop, and continuously improve our qualifications based on research and feedback, and how the design of our qualifications enables learners to achieve some of the outcomes that matter most to them.

[1] Pearson Education Ltd is also regulated by CCEA Regulation in Northern Ireland, Qualification Wales in Wales, and the Scottish Qualifications Authority in Scotland.

When we compile an efficacy report on our qualifications, we are reporting about:

- **Impact on learner outcomes:** Teachers' and learners' experiences of the qualification, the breadth of knowledge, skills, and/or competencies the learner develops, and learners' progression after the qualification
- **Fitness of assessment design:** Whether the scores and other diagnostic information that form the basis of the assessment component of the qualification provide an accurate reflection of what the learner knows and can do

In other words, the efficacy of a qualification is both the opportunity it provides for a wide range of learners to acquire sufficient depth and breadth of knowledge, skills, and/or competencies, and the extent to which the accompanying assessment measures the extent to which they have developed these.

Impact on learner outcomes

In terms of assessing impact of the qualification on learner outcomes, we have undertaken research to explore qualification implementation experience and learner progression from the qualification. In particular, we looked at impact on:

Category	Learner outcomes	Research questions
Learner access and experience	• Learners across the ability range have opportunities to develop their mathematical abilities • Through their teachers, learners are able to access the free qualification support to ensure equity and access for all • Learners are supported to achieve their potential and demonstrate their abilities by the design of the assessment materials	• What is the opportunity in a qualification for the range of students to have a fulfilling learning experience? • What is the consequent breadth and depth of the experienced curriculum? • How are teachers and students perceiving, responding to, and using, a given qualification's assessments and related support materials? • What is the perceived impact of those on teachers, and on students' knowledge, skills and affect?
Learner progression	• The qualification is effective in allowing learners to progress to further study, training, or work	• What best supports them in achieving their potential in a qualification and what might have acted as barriers? • What is students' experience during and after sitting a qualification and how are they supported during their transition to post-16 study? • How does the study of a given qualification prepare students for, and support them in, their further study?

Fitness of assessment design

In the Standards for Educational and Psychological Testing, Ofqual (Newton, 2017), the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) (AERA, 2014) have defined validity, reliability, and fairness as follows:

- **Validity** is "the degree to which it is possible to measure whatever that qualification needs to measure by implementing its assessment procedure (Newton, p16). Validity requires evidence that test scores can be interpreted as they are intended, and can be appropriately used for a specific, defined purpose.
- **Reliability** is "the consistency of scores across replications of a testing procedure" (p33). Reliability requires evidence of the consistency of scores over time, across multiple forms of the assessment, and/or over multiple scorers.
- **Fairness** suggests that "scores have the same meaning for all individuals in the intended population" (p50). Fairness requires evidence that when assessments are administered as intended, items are not systematically biased against any particular group of test-takers and students are not hindered in demonstrating their skills by irrelevant barriers in the test administration procedures.

Given the longstanding role of these standards as a source of guidance on best practices in the development and evaluation of tests, and the role they play in the legal defensibility of assessment, Pearson has adopted these three attributes as the Assessment Quality Indicators on which we publicly report evidence underlying our assessment products.

Pearson's qualifications products are designed, developed, and maintained over time by teams of subject matter and assessment experts. These teams regularly collect evidence of validity, reliability, and fairness of the assessment element of the qualification. A representative group of mathematics education academics and professionals, including teachers, is consulted throughout the lifetime of the qualification every time the assessment is taken.

Qualifications efficacy report structure

Our regulated qualifications efficacy reports are made up of two parts.

Qualifications Technical Research Reports include the technical manual or technical report associated with each element of thequalification including research on its design and implementation.

Qualifications Efficacy Reports include information about the research that informed the qualification's design, and any guidance about how the qualification, is impact on learner outcomes like experience and progression. They also summarise all the relevant technical reports associated with the assessment component of the qualification and any impact evaluation research related to the impact the qualification has on learner outcomes.

History and reach of the GCSE Mathematics qualification

GCSE (General Certificate of Secondary Education) Mathematics was introduced in summer 1988. The GCSE assessed attainment of the full range of students at age 16. It was based, at the time, on a mathematics curriculum a newly introduced National Curriculum for mathematics.

In the intervening years, there have been a number of changes to the National Curriculum in mathematics, as well as to the structures of GCSE Mathematics assessments. The intention of these changes has been to support all students to demonstrate their mathematical capabilities.

The version of GCSE Mathematics described in this report assesses a National Curriculum mathematics programme of study introduced for first teaching from September 2015, with first examinations in June 2017. This reform to the National Curriculum in mathematics was accompanied by a change from letter grades to numerical grades; students taking GCSE Mathematics are now awarded grades from 9 (for the highest achieving students) to 1 (for the lowest achieving students). For clarity, the reformed Pearson Edexcel Level 1/Level 2 GCSE (9–1) in qualification is sometimes referred to as "GCSE (9–1) Mathematics". The specific qualification Pearson offers is the Pearson Edexcel Level 1/Level 2 GCSE (9–1) in Mathematics (1MA1) qualification, or for brevity, the Pearson Edexcel GCSE Mathematics qualification. (Edexcel is a qualification brand for academic and generation qualifications from Pearson in England.)

The reformed curriculum features a renewed emphasis on deep conceptual fluency, mathematical reasoning, and problem solving, as well as wider and deeper content coverage. The impact of the written curriculum alone is not enough, however — it must be accompanied by coherent teaching and learning strategies which enable children to meet their aspirations (Schmidt and Prawat, 2006). This is known to be particularly true where the intended curriculum is novel or demanding — or both, as it is in this case. Pearson's provision of assessments and supporting resources, in support of this reformed curriculum, is therefore central to what is enacted in the classroom.

In England, mathematics is a compulsory subject for all students up to age 16. GCSE Mathematics is one of the core subjects of the English Baccalaureate. This is a key performance measure for schools pre-16, based on the number of students that take GCSEs in English, mathematics, history or geography, two sciences, and a language, and on their students' achievement in these subjects. GCSE Mathematics also contributes to measures of student progress, such as **Progress 8**.

Post-16, students can opt for academic ("A-level") or vocational courses, work-related apprenticeships or other employment — or a mixture of those. An important feature of GCSE Mathematics is that it is a key gatekeeper qualification for progression. Post-16, students can opt for academic ("A-level") or vocational courses, work-related apprenticeships or other employment — or a mixture of those. Under the current education system in England, students are required to continue studying toward a grade 4 if they fail to achieve this at age 16.

The GCSE (9–1) *Qualification level conditions and requirements*, published by the UK **Office of Qualifications and Examinations Regulation (Ofqual)**, sets out the rules and regulations for GCSEs (9–1) in all subjects, including mathematics. This is supported by *Qualification level guidance*, which explains how to comply with these rules.

Pearson offers these qualifications as one of five **Awarding Organisations (AOs)**, operating under the regulation of Ofqual. AOs design, develop, deliver, and award the recognition of learning outcomes (knowledge, skills, and/or competencies) of an individual following an assessment and quality assurance process that is valued by employers, learners, or stakeholders. Each AO intending to offer a GCSE qualification in mathematics must submit a specification, sample assessment materials (SAMs), and an assessment strategy to Ofqual for accreditation.

The Pearson Edexcel GCSE Mathematics qualification was accredited for first teaching from September 2015 and first assessment in June 2017. Since the qualification was first introduced to schools, Pearson has successfully delivered it to learners in June 2017, November 2017, June 2018, and November 2018.

To date, Pearson has issued 872,235 GCSE (9–1) Mathematics grades to learners in the UK. In August 2018, Pearson issued 59% of all GCSE Mathematics awards in the UK. Approximately 3,400 schools and other accredited institutions currently offer the Pearson Edexcel GCSE Mathematics qualification.

What are the elements of the GCSE Mathematics qualification, and what informs their design?

Content and regulation

The broad design and content of the GCSE Mathematics qualification is set out by the English Department for Education (DfE) in the document **Mathematics: GCSE subject content and assessment objectives**, and reflects the National Curriculum in mathematics. This document provides the basis of a national standard to which all AOs must adhere when designing the elements of their qualification. It outlines the subject aims and learning outcomes, the subject content, assessment objectives, and mathematical formulae that students must memorise.

DfE appointed a Mathematics Expert Group, composed of a small number of experienced members of the mathematics and mathematics education communities in England, to advise on the age-appropriateness, details, and pathways to implementation of the reformed curriculum for students aged 5–16 (culminating in GCSE assessments at age 16, in students' year 11).

The Mathematics Expert Group consulted with representatives from the wider mathematics and mathematics education communities, including representatives from AOs. As part of **this consultation**, Pearson provided DfE with feedback on the design of the subject content, from its development from draft to final version for publication. This feedback was underpinned by research and supported by Pearson's External Subject Advisory Group (ESAG) for mathematics. The seven members of the ESAG were mathematics teachers, experienced examiners, university lecturers, or experts associated with national subject or teaching associations. This group was consulted throughout the development process.

The content of the reformed mathematics curriculum (and by association, the reformed GCSE Mathematics) is much wider and deeper than in the previous GCSE, and incorporates an increased focus on problem solving and reasoning. Problem solving and reasoning account for between 50% and 60% of marks. The Office for Standards in Education, Children's Services and Skills, Ofsted (2012), found that mathematics teachers in England had little experience teaching and assessing problem solving and reasoning across the range of students, so there are significant new demands for teachers in this curriculum.

In comparison to the previous GCSE, students also need to memorise more formulae.

The ***Subject level conditions and requirements for mathematics*** and the ***Subject level guidance*** for mathematics stipulate the regulatory requirements for the subject as set out by Ofqual. These requirements include:

- calculator use and weighting
- tiering requirements
- weightings and interpretation of assessment objectives
- content domain area weightings

Before these requirements were established, there were several formal public consultations. Pearson participated in these consultations to inform the broad design of the qualifications. This design must be adhered to by all AOs wishing to offer a GCSE Mathematics qualification.

To give all students the opportunity to demonstrate their mathematical capabilities, the *Subject level conditions and requirements* document states that there should be two "tiers" for GCSE Mathematics, Foundation and Higher, each offering a different, overlapping grade range.

Students must sit all components of the qualification in the same tier. Foundation tier candidates are awarded a grade from 5 to 1. Higher tier candidates are awarded a grade from 9 to 4. Grade 3 is an "allowed" grade at the Higher tier; that is, it is possible for a student taking the Higher tier assessments to be awarded a grade 3, but these assessments are designed to be taken by candidates who will achieve grades 4 and above. This is to avoid the situation where students who narrowly fail to achieve grade 4 performance receive no qualification after two years of study.

Qualification design

In 2013, when the reforms to the National Curriculum and qualifications system in England were proposed, Pearson created an expert panel to advise on them. Panel members were chosen either because of their expertise in the UK education system, or because of their experience in reforming qualifications in other systems, in markets as diverse as Singapore, Hong Kong, Australia, and Europe.

The panel guided Pearson through a rigorous qualification development process that included:

- extensively comparing subject content against the highest performing jurisdictions in the world
- benchmarking assessments[2] against UK and overseas providers to ensure that they are at the right level of demand
- establishing External Subject Advisory Groups (ESAGs), drawing on independent subject specific expertise to challenge and validate our qualifications
- subjecting the final qualifications to scrutiny against the Department for Education (DfE) content and Ofqual accreditation criteria[3] in advance of submission to Ofqual for accreditation

Pearson's World Class Qualification Principles ensure that our qualifications are:

- demanding, through internationally benchmarked standards, encouraging deep learning and measuring higher order skills
- rigorous, through setting and maintaining standards over time, developing reliable and valid assessment tasks and processes, and giving end users confidence in the knowledge, skills and competencies of certified students
- inclusive, through conceptualising learning as continuous, recognising that students develop at different rates and have different learning needs, and focusing on progression
- empowering, through promoting the development of transferable skills

All Pearson qualifications are designed to be accessible to the full range of learners.

The Pearson Edexcel GCSE Mathematics specification sets out the purpose of the qualification, building on the requirements set out in the *Qualification level conditions and requirements* and making them relevant to GCSE Mathematics.

- Provide evidence of students' achievements against demanding and fulfilling content, to give students the confidence that the mathematical skills, knowledge and understanding that they will have acquired during the course of their study are as good as that of the highest performing jurisdictions in the world
- Provide a strong foundation for further academic and vocational study and for employment, to give students the appropriate mathematical skills, knowledge, and understanding to help them progress to a full range of courses in further and higher education. This includes Level 3 mathematics courses as well as Level 3 and undergraduate courses in other disciplines such as biology, geography, and psychology, where the understanding and application of mathematics is crucial
- Provide a basis for schools and colleges to be held accountable for the performance of all of their students (Pearson Edexcel, 2017: Rationale)

The specification also states that the aims and objectives of GCSE Mathematics are to enable students to:

- develop fluent knowledge, skills and understanding of mathematical methods and concepts
- acquire, select, and apply mathematical techniques to solve problems
- reason mathematically, make deductions and inferences, and draw conclusions
- comprehend, interpret, and communicate mathematical information in a variety of forms appropriate to the information and context (DfE, 2013: p3)

[2] See Pearson (2013) **International Comparative Analysis of Science and Maths Assessments** and Pearson (2013) **The evidence base for tiering in GCSE Exams.**

[3] These criteria were defined by Ofqual for GCSE qualifications in their GCSE (9–1) Qualification Level Conditions and Requirements document, first published in April 2014 and updated annually.

Post-examination review, June 2017

Following the first sitting of the GCSE (9–1) Mathematics assessments in June 2017, we commissioned a number of internal and external reports, reviews and analyses, including an impact evaluation report, the Summer 2017 Assessment Monitoring Technical Report, and detailed collated feedback from key stakeholders, such as mathematics teachers. We identified three main areas we should focus on to improve the accessibility of our assessments and therefore enhance the examination experience for all students:

- **Language:** Ensure questions are designed to test maths skills, not language skills. Use simple and straightforward language, and ensure questions set in real-life contexts use language that is purposeful, clear, and relevant to all.
- **Level of difficulty and complexity:** Provide the least able students with greater opportunities to demonstrate what they can do by ensuring the early questions on Foundation tier papers are accessible to all. Additionally, improving the accessibility of the questions targeted at grades 4 and 5 (including common questions), and the overall complexity of the approach to assessing problem solving, were felt to merit further attention.
- **Ordering of questions within assessments:** Arranging questions so they are progressively more difficult was a valued design feature in the legacy assessments. A smooth increase in difficulty within the assessment helped to build students' confidence and led to an overall more positive examination experience.

The Writers' Guide and Assessment Design Principles were amended to address these areas, to help create assessments at both tiers that provide a better examination experience for students at all levels. These principles were embedded for assessments being taken by students from June 2018 onwards.

Assessment

The Pearson Edexcel GCSE Mathematics qualification is assessed by three equally weighted written examination papers at either Foundation or Higher tier. Each paper is 1 hour and 30 minutes long (for a total of 4 hours and 30 minutes of assessment time) and has 80 marks (for a total of 240 marks). One of the three papers at each tier is a non-calculator assessment. The content outlined for each tier is assessed across all three papers and the assessment objective weightings are met within each paper. Each paper includes a range of question types and questions gradually increase in difficulty, with questions targeting the highest grades appearing toward the end of the papers.

Pearson has adopted validity, reliability, and fairness as Assessment Quality Indicators (AQIs). We use these indicators when reporting on the evidence underlying our general qualifications.

AQI1: Validity
GCSE Mathematics grades can be interpreted as a measure of candidates' comprehension of course content

A key goal of Pearson Edexcel GCSE Mathematics is to enable qualification users to make sound interpretations of candidates' capability. It is therefore important that the content described for Pearson Edexcel GCSE Mathematics is reflected in test specifications and test development. It is also important that a single construct underlies the responses to the items on a test, in order to validate the single grade reported.

Table 1: Evidence related to AQI1

Subject Assessment Strategy for GCSE (9–1) Mathematics	Explains how the General Qualifications Assessment Strategy and various other regulatory documents have been applied in the design of the qualification
Writers' Guide and Assessment Design Principles	Ensure that each test created is consistent with the sample assessment materials (SAMs)
Scrutineers' reports	The scrutineer is a subject expert with input towards the end of the assessment writing process, before papers are finalised
Reports for awarding by senior examiners and the awarding officer	Completed at the end of the marking process to inform and explain grade-setting decisions

Subject Assessment Strategy

The **Subject Assessment Strategy** describes research conducted at the time of development. This included benchmarking of assessments against equivalent international qualifications, small-scale trialling, and stakeholder engagement with teachers, higher education lecturers, learned societies, and employers (reflecting the range of progression route for this qualification).

This supported a three-paper model in which each paper has an assessment time of 1 hour 30 minutes. When compared with a two-paper model, the three-paper model provides learners with manageable examination time (particularly at Foundation tier), minimises examination fatigue and provides a greater number of total marks. This means more content areas and skills can be assessed, leading to a more valid assessment of mathematics at this level. The greater proportion of content to be assessed is crucial, given the wider and deeper body of content at both tiers.

This three-paper model also allows a non-calculator weighting of 33.33% (within the allowable range of 33–50%), which was also supported by the ESAG for mathematics. The first paper is the non-calculator paper, a decision which was endorsed by the ESAG for mathematics, as well as by learners and teachers in Pearson customer research, and which reflected an appropriate balance in terms of time devoted to these skills through teaching and learning.

Apart from the use of calculators, all other aspects of the design of the papers, such as content and assessment objective weightings, are uniform across the papers. This supports the holistic nature of the subject and the design intention, which is to create a consistent examination experience with no disproportionate emphasis on a particular area or skill in one paper when compared to another. It also means that performance on each component is likely to be similar to other components, and representative of overall performance. It also ensures there is a balance in emphasis placed on all qualification content and skills and that the teaching and learning experience is coherent and balanced.

A total of 80 marks is available in each paper. This design decision is based primarily on trialling and an internally developed formula for calculating assessment time. This formula was adapted to factor in outcomes from trials, which demanded increased thinking time for problem solving questions.

Pearson qualifications use a range of question structures and complexities, and a range of levels of structuring, so as to assess multiple processes while maintaining reasonable accessibility across the range of students. Question types are chosen by consideration of the skill being assessed.

Questions used in Pearson Edexcel GCSE Mathematics are designed to be similar to the previous GCSE Mathematics qualification where this is possible, because they are familiar to students and there is a wealth of performance data to support how well they function. Where this is not possible, for example when assessing content not in the previous qualification, in addition to deeper problem solving and reasoning skills, new questions were developed and tested in the small-scale trials. Many question types in the reformed GCSE Mathematics have been developed to be less procedural than in the previous curriculum. Instead, the focus is now on demonstrating understanding, conjecture, proof, and non-routine problem solving.

The mark schemes for the assessment continue a change introduced with the previous GCSE qualification: they tend to be more generic and open, rather than attempting to capture all possible approaches to answering the question. This approach proved successful with the previous qualification, leading to improved marking quality. This also helps markers when assessing problem solving. This style of mark scheme is intended to make it possible to more accurately reward appropriate use of mathematics, including correct methods, processes, communication, and accuracy.

The assessments are designed to be accessible to the full range of learners, while continuing to stretch the most able students. The papers are intended to gradually increase the demand placed on the examinee, and are designed to help students approach the exams with confidence.

Writers' Guide and Assessment Design Principles

The Writers' Guide and Assessment Design Principles ensure that standards are maintained over time, that each assessment targets the appropriate ability profile, and that candidates' experience of assessments is consistent.

This ensures that assessment writers have a common and consistent understanding of how to create tests, with detailed information about the number of marks to target at each grade, the weightings of content areas, and which content statements are suitable for assessment at each grade range. It makes clear requirements relating to Assessment Objective weighting, and how to apply them in the light of technical interpretations. Finally, it gives clear instructions about the construction and creation of consistent mark schemes for markers.

Grades 4 and 5, which are targeted in both Foundation and Higher Tier, are targeted by a set of common questions. These appear in the same way in both Foundation and Higher Tier papers and help to ensure that these grades are assessed in the same way for all candidates.

Scrutineers' reports

After assessments have been written, they are given to two "scrutineers". These scrutineers are subject specialists. They first complete the assessment themselves, to make sure that all the questions can be answered, and that the assessment can be completed in the allocated time. Next they review the mark scheme alongside the question paper, to ensure that the mark scheme takes into consideration the range of different ways students may approach the questions and appropriately rewards students for what they are demonstrating. Pearson Edexcel GCSE Mathematics has two scrutineers, repeating the same procedure.

Reports for awarding

At the end of the marking process, a range of reports are written. Some are written by senior examiners in charge of marking each paper. These report whether questions performed as intended: that is, whether the questions were understood and answered in the way they were intended, both in terms of the skills used by candidates, and also whether they targeted the intended candidate ability.

This draws the attention of colleagues less familiar with the paper to which questions discriminated well, and which are likely to contain evidence of borderline performance at grade thresholds. The reports also help senior examiners review candidate work as part of the grade-setting process.

After grades have been set, the Pearson awarding officer also writes a report, detailing and justifying the decisions made. This report is detailed and wide-ranging, covering:

- potential risks identified prior to the assessment being taken, and the mitigations put in place for them
- an overview of complaints received about papers and questions
- statistical descriptors of cohorts
- mark distributions
- statistical recommendations for final grade distributions, made in conjunction with Ofqual and the other GCSE AOs

AQI2: Reliability
A candidate's outcomes are stable, in that the grades are consistent, both over time and on multiple test administrations

Another important goal of the Pearson Edexcel GCSE Mathematics qualification is to minimise errors in judgement and decision making by providing outcomes that are consistent over different assessment occasions and administrations. This AQI is measured in two ways: inter-marker agreements (the extent to which different markers marking the same item agree with each other) and internal reliability (the extent to which all questions in the assessment test the same thing).

Table 2: Evidence related to AQI2

Scrutineers' reports	The scrutineer is a subject expert with input towards the end of the assessment writing process, before papers are finalised
Standardisation	Standardisation is the process by which all markers learn how to apply the mark scheme in as consistent a way as possible
Marking quality reports	Information relating to marker performance, for monitoring accuracy during marking
Data from post-results	After results are issued, schools and other institutions can request that a senior examiner reviews the marking of individual papers
Assessment monitoring technical reports	A report containing summary statistics to help describe the functioning of questions and assessments

Scrutineers' reports

Ensuring marks are reliable starts with the creation of the assessments. One of the roles of the scrutineers, when they report on the papers before they are finalised, is to review the mark scheme alongside the question paper to ensure that marks are being appropriately awarded, and that it is clear what a candidate needs to do in order for a mark to be awarded.

Standardisation

Every marker for GCSE Mathematics must go through standardisation. This process is conducted online. They mark some answers shortly after the assessment in order to get used to applying the mark scheme (these marks are later reset). They then get in contact with their supervising examiner to address any questions or concerns.

Finally, they attempt two sets of items for each question: a practice set, which is annotated with guidance for application of the mark scheme, and a qualification set, where their marks are compared to a set of marks agreed by a senior examiner. If their marking is accurate enough in this qualification set, then they are cleared to mark. If it isn't, then they talk through their understanding with their team leader and get a second qualification set. Markers who do not meet the accuracy requirements on their second qualification set will not be able to mark that question.

Marking quality reports

Online marking uses two measures of marking quality during the marking period: backreading and validity items.

Backreading is when a senior examiner reviews the marks from a more junior marker. They see the original mark awarded, as well as any notations. If they feel the mark awarded represents a correct application of the mark scheme, then they confirm the score. Otherwise, they amend it, and the mark of the more senior examiner is taken.

Validity items are items marked by senior examiners at the start of the marking process. These items then appear as part of an examiner's marking (without the senior examiner's mark). The scores given by the senior examiner and the examiner are then compared. This is normally reported as "agreement within a tolerance", or "adjacency".

In June 2018, more than 200,000 quality checks were carried out on Pearson Edexcel GCSE Mathematics. Final agreement figures ranged from 96% to 98% for all papers, for both backreading and validity measures.

Data from post-results

From the day of results, schools can view their candidates' completed assessments online, without charge. These transcripts include the marks awarded for each item, alongside the finalised mark scheme and grade boundaries. If schools have concerns about results, they can ask for a senior examiner to review the marking of individual transcripts.

After the June 2018 series, nearly 20,000 requests to review marking for Pearson Edexcel GCSE Mathematics were completed. Fewer than 0.8% of the total entry had their grades changed following review. Out of more than 43,000 examination transcripts reviewed, fewer than 100 resulted in a mark change greater than 4 marks (or 5% of the total paper marks).

Assessment monitoring technical reports

As well as inter-marker reliability, Pearson also monitors internal reliability using Cronbach's Alpha. Internal reliability is a measure of whether the items in the test produce similar scores. Pearson uses this measure in two ways.

First, we apply the measure to complete tests. A value above 0.8 for this is considered to be good.
The Cronbach's Alpha values for Pearson Edexcel GCSE Mathematics in June 2018 ranged from 0.876 to 0.907.

Second, we calculate Cronbach's Alpha for each item, if that item is removed from the paper. If the Cronbach's Alpha value with that item removed from the paper is higher than the value for the complete test, then it suggests that the item is not assessing the general construct of the test in the same way as the other items. This is rare in Pearson Edexcel GCSE Mathematics, and where it does occur, it is usually at the very beginning or very end of the paper. These questions target the extremes of the ability profile, and therefore this is to be expected.

AQI3: Fairness

Qualification grades can be interpreted the same way for candidates of different subgroups

Pearson strives to provide grades that can be interpreted in the same way for all candidates, regardless of demographic characteristics like gender or ethnicity. Fairness implies that when the assessments are administered as intended, items are not systematically biased against any particular group of test-takers, and candidates are not hindered in demonstrating their skills by irrelevant barriers in the test administration procedures.

Table 3: Evidence related to AQI3

Writers' Guide and Assessment Design Principles	Ensure that each test created is consistent with the sample assessment materials (SAMs)
Scrutineers' reports	The scrutineer is a subject expert with input towards the end of the assessment writing process, before papers are finalised
Assessment monitoring technical reports	A report containing summary statistics to help describe the functioning of questions and assessments

As well as information about content weighting and targeting, the Writers' Guide and Assessment Design Principles instruct assessment writers how to use language and frame questions to avoid unintended bias for or against any subgroups taking the test.

This is also something that is considered by the scrutineers when they report back on the assessments. Assessment monitoring technical reports give breakdowns of grade achievement by gender for Pearson Edexcel GCSE Mathematics. Across all AOs in the UK, GCSE Mathematics shows gender discrepancies at the higher grades — more than 53% of candidates achieving grades 8 and 9 were male (50.6% of Pearson candidates are female). Pearson's results reflected this national trend — a larger proportion of candidates achieving higher grades were male.

However, group differences are not, in and of themselves, evidence of bias. If group differences remain after controlling for candidates' ability, then items should be further analysed for potential sources of bias (Zieky, 2016). This trend is not restricted to Pearson Edexcel GCSE Mathematics and research and analysis into the causes of this difference, as well as exploring potential mitigations, is ongoing.

Qualification support

Alongside the design and development of the qualification, Pearson developed both free and paid-for resources intended to support the implementation of the Pearson Edexcel GCSE Mathematics qualification in schools and other institutions. It is crucial that these resources are aligned to the qualification content and assessment, and that they fully support the aims and objectives of the qualification.

Pearson is well placed to offer a coherent curriculum, incorporating the design of the curriculum itself, free support, and paid-for published resources for teachers and learners. With a coherent curriculum for all, students can become more confident mathematics learners and be better prepared to use mathematics in their lives, whatever their next steps are.

Targeted support was developed to help teachers deliver new content and embed fluency, reasoning and problem solving: areas where we have found many teachers can particularly appreciate additional support. GCSE resources were developed around a sound pedagogical evidence base and an innovative mastery approach. While these resources are mostly paid-for, Pearson is in a position to provide some components as part of the free support: for example, lesson plans and aligned schemes of work.

The development of the free and paid-for qualification support involves:

- consulting with teachers and with educationalists, and analysing historical trends to identify which elements of the specification are likely to need the most support
- reviewing pedagogies and approaches internationally to identify which strategies provide evidence of improving confidence, fluency, problem solving, and reasoning skills
- creating an author team drawn from mathematics teachers and educationalists to apply the principles to our support materials
- developing a progression map, identifying the underpinning knowledge and skills and the relationships between different concepts. This underpins the design of the support materials: teaching plans and content, and related assessments
- regularly testing and trialling content through development with teachers

Table 4: Summary of free qualification support resources offered alongside Pearson Edexcel GCSE Mathematics

	Free qualification support resources
Guidance/ training	**Getting started guide:** provides an overview of the reformed specification; the changes to the content, assessment and support; and the implications for teachers and learners **Content mapping:** detailed content mapping from the content in the new specification to content in the previous specification **Getting Ready to Teach events:** face-to-face and online training events to provide teachers with all the information and support they need to start teaching the reformed specification **Parent and student guide:** a brief downloadable guide targeted at parents and students, explaining the reforms to the GCSE and how students will be assessed **Tiering and assessment guidance:** a package of materials intended to support understanding of the assessment design and structure, as well as the tiering model, to help inform tiering decisions
Planning	**Content guidance:** provides commentary from the senior examiner's team on the individual content statements within the specification to provide more detail on how the specification will be assessed. Also supplies generic examples, relevant question references, and specific guidance on command words **Teaching guidance:** breaks down each content statement into learning objectives and provides sample questions **Schemes of work (1, 2, 3, and 5-year):** offers a way of delivering the course in individual lessons. **Interactive schemes of work:** allow teachers to re-organise content and lessons to suit their students
Teaching	**Transition from Key Stage 3 teacher resources:** includes baseline tests, transition scheme of work and end-of-term tests and markbooks **Lesson plans:** with links to resources and aligned to the scheme of work **New content resources:** produced by teachers for teachers, with guidance on how to teach content new to Foundation and/or Higher tier **Formulae posters:** classroom posters to help students memorise all the required formulae **Access to Foundation tier resources:** classroom resources targeted at helping low ability students to build the foundations required for GCSE

Assessing	**Specimen papers:** published before the first set of live papers, to give teachers an understanding of what the assessments will look like and potentially to use as practice for their students
	Problem solving practice papers (gold/silver/bronze): papers that allow teachers to adapt problem solving questions, increasing or decreasing the amount of scaffolding
	Common questions papers: papers collating all the questions common to both Foundation and Higher tier papers, targeted at grade 4 and 5 students
	Problem solving questions and topic tests: A collation of additional problem solving questions and topic tests aligned to the schemes of work
	Practice papers and themed papers: collections of questions from previous papers, re-packaged and re-organised as additional assessment support
	Mock papers: original papers with unseen questions written by the same team of senior examiners following the same process as for live papers. Intended for use by teachers as mock exams ahead of the real exams being sat by their students
Exam materials	**Question papers:** the live examination papers students have taken
	Mark schemes: the documents used by examiners to mark the question papers
	Examiner reports: written for teachers by the senior examiner, giving question by question feedback on students' performance, common misconceptions and areas for improvement
	Grade boundaries and statistics: the thresholds for each grade and the national and question-level performance statistics
	Exemplars: real student responses to some questions, showing common misconceptions and how the mark scheme has been applied
	Feedback on examinations: detailed feedback on the examinations from Pearson, either pre-recorded or as a live event

Note: This summary is not exhaustive. The free qualification support is continually added to and enhanced throughout the qualification lifecycle.

Alongside the resources, a range of other free and paid-for services and resources have also been established that sit alongside the GCSE.

Table 5: Services and paid-for resources offered alongside the Pearson Edexcel GCSE Mathematics qualification

Free services	Paid-for services and resources
Qualification-specific website: a quick, easy way for teachers to access the support materials **Mathematics Emporium:** a website offering a vast number of past papers, examination-related materials, and bespoke resources, along with an email update service for mathematics teachers **ResultsPlus:** an online results analysis tool that gives teachers a detailed breakdown of their students' performance, so they can identify topics and skills where students could benefit from further learning **Access to Scripts:** an online portal where teachers can view and download candidates' marked exam transcripts **ExamWizard:** an exam preparation tool containing a bank of past Edexcel exam questions, mark schemes and examiners' reports	**Post-results services:** allow students to request a review of the marking of their exam transcript **Continuing professional development:** including training on marking mock exams **ActiveLearn:** an online service that includes tools for teachers and students. Tools for teachers include a planning tool, teaching tools (including printable and digital versions of whole class and interactive activities), assessment and tracking tools, tests, and reporting tools. Tools for students include practice games, problem solving activities, and supplementary targeted practice materials **Published resources:** including textbooks for each tier, additional problem solving practice books, and other print materials including support for low attainers, all specific to GCSE Mathematics

What does the impact evaluation research say about the effectiveness of the design and implementation of each element of the Pearson Edexcel GCSE Mathematics qualification?

We have carried out a longitudinal programme of research examining the implementation and effectiveness of the Pearson Edexcel GCSE Mathematics qualification, including its free qualification support, over the first two cycles of its enactment for year 11 students (aged 15–16) and into those students' year 12.

The classroom focused study *Key Stage 3 Maths Progress and GCSE 9–1 Mathematics* ran in parallel with this one. That study focused on the use of paid-for resources and provided information about enactment and use of resources. The use of paid-for resources is out of scope for this current report.

There is little systematic central evaluation of the implementation, risks and impact of the reformed and aspirational mathematics curriculum. The largely qualitative programme of research described here therefore has an important contribution to make to our national understanding of the longitudinal enactment of the reformed GCSE, the impact of its introduction, and the implications of similar aspirations in other markets.

Programme of research into Pearson Edexcel GCSE Mathematics

Research questions examined in this programme of research	Metrics used to investigate the questions
Experience (phases 1 and 3): • How are teachers and students perceiving, responding to, and using, GCSE assessments and related support materials? • What is the perceived impact of those on teachers, and on students' knowledge, skills and affect? • What is the consequent breadth and depth of the experienced curriculum? • What is the opportunity in GCSE assessments for the range of students to demonstrate their mathematical capabilities?	• Interviews with Heads of Mathematics • Interviews with Mathematics teachers • Focus groups with year 11 students • Surveys of year 11 students
Progression (phases 2 and 4): • What best supports them in achieving their Mathematics GCSE potential and what might have acted as barriers? • What is students' experience during and after sitting Mathematics GCSE and how are they supported during their transition to post-16 study? • How does the study of Mathematics GCSE prepare students for and support them in their further study?	• Reflective telephone interviews with Heads of Mathematics • Interviews with Mathematics teachers in post-16 settings • Focus groups with year 12 students

Methods

How did we undertake the programme of research?

This largely qualitative programme of research took place in four phases over the first two cycles of the reformed GCSE Mathematics, preceded by a pilot study in Summer 2016. We worked collaboratively with a mathematics education specialist researcher at UCL Institute of Education[4] to add rigour to the development and validation of the research, and largely used subject-specialist (internal and external) researchers to complete fieldwork, so as to ensure sensitivity to subject-specific issues either raised or implicit.

Phases 1 and 3 took place in the Spring 2017 and Spring 2018 terms respectively. These phases probed the perceptions, outcomes, and impact associated with enactment in year 11 both teachers and students.

Phases 2 and 4 took place in the Autumn 2017 and Autumn 2018 terms respectively. These phases explored the early impact of students' experiences, learning, and GCSE attainment on their subsequent pathways. This involved sampling teachers and year 12 (post-GCSE) students who were engaging with various post-GCSE options: a calculus-intensive pre-university course (A-level Mathematics), a contextualised problem-solving course (Core Mathematics), a GCSE Mathematics resit course, mathematics-intense pre-university courses (such as science, engineering, or computing), and other mathematics-using academic or vocational courses.

The programme of research involved 74 year 11 classes in 28 schools, and their teachers and Heads of Mathematics, over two years. The sample was reasonably representative in terms of key school, teacher and student characteristics known to affect teaching and learning in mathematics. These include school performance measures, inspection rating, catchment area (for example, rural, inner city, or urban), school type (for example, independent, mixed, or selective), and students' socio-economic status (by proxy measure). Each year 11 included students and teachers of both the Higher and Foundation tiers.

All interviews were transcribed, and all data analysed in relation to the research questions and open coded emergent grounded sub-themes (Charmaz, 2006). Interpretations were validated across researchers and sometimes across teacher participants.

A sample of this size might not be fully representative of the range of students and institutions engaging with Pearson Edexcel GCSE Mathematics, and enactment will almost certainly continue to develop as teachers' knowledge of the reformed GCSE matures. However, the study is indicative of early responses, and of Pearson's proactive engagement with the inevitable limitations of early assessments and support materials. Any qualitative study is dependent on interpretation, even one as systematic as this one, where validity was enhanced by the representativeness of samples, as well as by cross-researcher and cross-participant validation.

[4] UCL Institute of Education has been #1 for Education in the QS world rankings from 2014–2019 inclusive.

Table 6: Data collection

Phase	Data collection
Phase 1: Spring 2017	19 Head of Mathematics interviews 38 year 11 class teacher interviews 38 year 11 focus groups (n~140 students) Year 11 student survey (n=1,627 students)
Phase 2: Autumn 2017	19 Head of Mathematics reflective telephone interviews Post-16 progression work in 10 post-16 settings: • 26 teacher interviews • 26 year 12 focus groups (n~78 students)
Phase 3: Spring 2018	19 Head of Mathematics interviews 38 year 11 class teacher interviews 38 year 11 focus groups (n~140 students)
Phase 4: Autumn 2018	19 Head of Mathematics reflective telephone interviews Post-16 progression work in 18 post-16 settings • 56 teacher interviews • 56 year 12 focus groups (n~170 students)

Findings

What were the research findings?

Content and specification

In the first year of enactment, given a limited lead time from specification accreditation to first teaching, teachers usually engaged with the national curriculum or specification indirectly, relying initially on school "schemes of work" (usually based on available curriculum resources) to interpret the curriculum, and then increasingly on emergent assessment materials, which dominated their planning in the second year.

Teachers and students generally supported the principles behind the reformed curriculum, especially in relation to problem solving, but concerns remained about the demands on the weakest students at both tiers. These were focused on two issues:

- the language used, especially around problem-solving contexts, was felt to be a barrier to weaker readers and to those with English as an additional language
- it was widely felt that weaker students entered at each tier had limited opportunities to demonstrate their mathematical capabilities

Although teachers did not articulate it as a key issue for them, the research revealed that teachers needed more support to develop systematic approaches to problem solving and reasoning. Many also had mixed conceptualisations of these elements. Teachers therefore made significant use of the emergent assessment-related materials to interpret and support them in those areas.

In the first year of enactment, teachers reported that some of the most mathematically able students felt over-challenged by the new specification, at least in relation to the challenge and expectations of their other subjects. This was less apparent when talking to those students.

By the second year of enactment, this concern appeared to have largely dissipated. Importantly, students preparing for papers on which they were expected to perform highly — that is, the stronger students at either tier — appeared to be thriving on the new specifications, and to be enjoying the challenge they perceived in emergent assessments, recognising the satisfaction of aiming high and reaching potential. The stronger students at both tiers reported feeling confident and equipped to build on their GCSE Mathematics qualification in their future pathways.

Assessment materials

Teachers were satisfied that Pearson's assessments were a valid reflection of the intentions of the curriculum. But in the first year of enactment, there was apprehension about the depth of demand on students who had only recently come to this curriculum. They felt that Pearson's specimen assessments were clearly presented, but concerns remained about the language demands for many students. As assessment materials emerged, they became highly influential in driving interpretation and enactment of the curriculum.

In the first year of enactment, tier entry decisions caused significant anxiety and were often made as late as possible. In practice, and in retrospect, most were happy with both their decision and the student grade outcomes. Students' responses to tiering decisions and attainment appeared to have been influenced by teachers' perceptions and communicated attitudes. To address this, Pearson is now providing enhanced tools and guidance to support teachers to make tiering decisions.

By the second year, most teachers were much more confident about such issues. The high-stakes nature of GCSE Mathematics meant that middle-attaining classes were often taught selected Higher tier topics so that students could target the small number of marks needed for a Higher tier grade 4 or 5. This happened in at least 22 of 39 such classes:

"our experiences of the students who are put in for Foundation are very unlikely to get a 5, and not many of them at our school seem to get a 4. So we'll be trying to enter at Higher... Because that's the best chance that they can have of getting a 4"
— Higher teacher 21, Spring term 2018

"[the allowed grade 3 gives] a little bit of a comfort blanket... that they're not going to fall off the end"
— Higher teacher 14, Spring term 2018

Other teachers made more mathematics-focused decisions:

"The temptation is there for this teaching to the test... but that absolutely serves them no good whatsoever because they're not learning maths that could be useful in the future, they're just going to be passing a test. That's not what we're about"
— Head of Maths 13, Spring term 2018

The former approach sometimes equated to a mathematically rather incoherent experience, and in interviews, students from these classes lacked confidence in their mathematical functioning, or indeed inclination to build on their GCSE Mathematics qualification:

"I think it's just like doing what you can, like even if it's just a little bit, write it down and you can get a mark from it"
— Higher tier student 27B, target grade 4

"Normally I'll skim over the question first to see if there are any numbers in the question and then I'll look over it again and try and see what they are actually asking because as I said, it confuses me quite a lot."
— Higher tier students 5D, target grade 5

In contrast, similar-attaining students whose teaching prepared them for Foundation tier entry reported thriving on their experience, and exuded mathematical confidence and interest:

"Yes it's good, it's hard but I'm getting there when I really try, I get quite a sense of satisfaction"
— Foundation student, March 2018

Free qualification support

Mathematics Emporium was regarded as a highly supportive collection of materials and a unique differentiator for Pearson. Teachers reported that its bank of resources and email update service were "highly supportive", and it was very influential in schools' choice of the Pearson specification:

"I think you probably provide the best resources of any exam board and I think that is key for us really. It's great to have. You know, a couple of times a year we get an email through, going have you noticed all these extra things we've put up. And I think that's really good"
— Head of Maths 9, Autumn term 2018

"Well all the support, I mean for example even having this discussion today is the support, it's the listening...It feels collaborative if that makes sense"
— Head of Maths 12, Spring term 2018

"I value the support. I value the Emporium. I value the fact that I can email people direct and get an answer. I've always had a positive experience with Edexcel"
— Head of Maths 22, Autumn term 2018

"I think the Emporium's always been the big thing for me, and as a head of department I think it's always been structured in a way that's allowed me, with all the changes that we've had, to put together schemes of work, to put together packages for teaching in terms of the assessment and the like"
— Head of Maths 11, Spring term 2018

In school, Heads of Mathematics typically prioritised and filtered the use of free qualification support materials through years 10 and 11, so that students' access was controlled and progressive. Students and teachers particularly valued practice question papers and targeted sets of questions, such as the problem solving papers:

"The bronze, silver, gold problem-solving papers, they're really good because they build up your confidence"
— Foundation tier student, Spring term 2018

"If you take the practice papers, and the problem-solving papers, together, you work with your teacher and your friends to build up what you can do, but then you gradually come to believe you can crack them by yourselves and that's a pretty good feeling"
— Higher tier student, Spring term 2018

In the second year of enactment, year 11 teaching was often highly supported by, if not entirely dependent on, the free support, with much less use of paid-for resources. The free support available was widely felt to offer high quality support for the range of students, including the weakest, although even with increased accessibility, emerging live papers were still felt to offer limited opportunity to the weakest students at each tier.

Such support was used in a variety of ways, from a focus on building up experience with assessment questions, to a very obvious development of reported practice:

"Well we've gone through them quite a lot in class...maybe like start from the end and then work back to the beginning of the question"
— Student 28A, March 2018

"Recently we've been on the crossover questions in class, we've had lots of that and Miss has been, she's been quite harsh to make the point about the reasoning questions that you have to write sentences... It's just getting used to using words rather than numbers, because loads of us forget"
— Student 7A, March 2018

"What I've told people is don't just accept an answer. You need the student to explain where the answer's come from, their thinking, how they worked it out, what their reasons for getting that answer were. We're spending more time questioning students about their thinking"
— Head of Maths 4, Spring term 2018

Teachers, and especially Heads of Maths, also valued enhanced opportunities to view and download their students' marked transcripts for free (through Pearson's Access to Scripts service) for formative purposes, and teacher support events:

"We get all the papers back, we look at them. We've got pretty much all of the papers back from the students...we've looked at their particular questions ...and why that happened and what the misconceptions are....it's really very useful"
— Head of Maths 9, Autumn term 2018

"We have downloaded the papers and I do think that's a great thing that Pearson do... I do think that is fab....It's a lot easier in regards to re-marks or targeting weak areas for Year 12 resits. It's a great service that is."
— Head of Maths 8, Autumn term 2018

"Yes, booked on it [feedback event]. I do like those, they are very good"
— Head of Maths 1, Autumn term 2017

Progression

In interviews in the progression phases (phases 2 and 4), year 12 teachers reported sound and improving impact on the progression of the more strongly attaining students in each tier in the second year, including in terms of attitude, persistence, and problem solving (though possibly less algebraic fluency and a few small concerns about basic mathematical fluency, such as with percentages):

"They are better prepared ...because they are deeper thinking"
— Maths A-level teacher 6, Autumn term 2017

"They are a lot more willing, independent, confident to give it a go"
— Maths-related teacher 11, Autumn term 2018

"I find that a lot more of them are better at attempting the problem than perhaps previous years would have been even to start it"
— Maths A-level teacher 9, Autumn term 2018

"They can hit the ground running in a much better way. They're much more used to unfamiliar contexts... they're braver in a way"
— Maths teacher 16, Autumn term 2018

"Because of the way the GCSE is set up, they can achieve a 6 without really having a grasp of the algebra"
— Maths A-level teacher 10, Autumn term 2017

Such positive impact was not always obvious because of variability among successive cohorts of students at the institution level, but by the second year of enactment, it was widely reported for many students of A-level Mathematics or of other similar level mathematics-intensive academic or vocational courses. Teachers also reported emerging confidence in the GCSE to support participation in post-16 mathematics, whether in A-level or as a user of mathematics:

"Numbers [at A-level] have grown a bit but the quality of the students that we're getting is stronger" —
— Head of Maths 8, Autumn term 2018

"There's more [mathematical] rigour in the new A Levels so psychology and some of the sciences... were ...trying to put their benchmark up in terms of the maths, ...but actually these kids have tended to get on the course and succeed there even though they've been a grade down: the new GCSE is giving them the skills and confidence they need"
— Head of Maths 15, Autumn term 2018

However, the study showed that the attitudes and experiences of low attainers at either tier need further attention. Pearson had been proactive about responding to early concerns about the very weakest students, and that proactivity was recognised and valued by teachers:

"They were easier to access to start, the Foundation [papers]. So the children built their confidence and then could apply their mathematics further on into the exam papers"
— Head of Maths 6, Autumn term 2018

"[making the language easier] was important because you know we had found that students hadn't been able to access them before and they were put off too quickly, couldn't show what they could do. But they're still prohibitively hard for the weakest, once you get past the first few questions"
— Foundation and GCSE resit teacher 21, Spring term 2018

Of note was data related to students of Core Mathematics, a contextual mathematics problem solving course aimed at students who want to maintain mathematical functioning without studying for a calculus-rich pre-university course. Students of Core Mathematics showed mixed attitudes to their GCSE Mathematics experiences (associated with their tiering experiences), but recognised the GCSE qualification as a suitable foundation for the study of their current more realistic and embedded mathematical activity, appropriate to their age and current pathway:

"GCSE maths was OK, you need to know that stuff, but now we're really getting to grips with where we need to use it, and that's really interesting, and makes you much more confident you could work with maths-y things later on"

— Core Maths student, Autumn term 2018

Discussion

In response to issues emerging during the first year of the study, Pearson was able to further develop the free qualification support to provide even better support for student access, equity of opportunity, and empowerment. We were also able to improve future assessments within the confines of Ofqual regulations. These actions were both noticed and appreciated by many of the teachers who participated in the study. Further improvements are being made to the free qualification support as a result of the second year of the study.

Questions of improving accessibility are challenging. Arguably, the assessment (and learning) experiences of the weakest students entered for Higher tier papers could be improved by preparing them instead for Foundation papers, where they could expect to be able to address most of the papers. Given the spread of grades targeted by Higher tier papers, reliable assessment requires that grade thresholds for grades 3, 4, and 5 will inevitably reflect an overly challenging assessment experience for many lower-attaining students at that tier.

However, GCSE Mathematics is a high-stakes examination for both students and teachers, so perceptions of "easier routes" to a grade, albeit in the face of careful work by Pearson to ensure cross-tier comparability, may sometimes lead to teachers making grade-focused decisions about approaching the subject content. The study shows that such approaches can result in mathematically incoherent learning experiences for students, for example through the selection of content which is then treated at a superficial level rather than incorporating the range of mathematical processes intended in the curriculum.

Such grade outcome-focused approaches were reported to be supported by the allowability of a grade 3 on Higher tier: the removal of an allowed grade 3 might support more mathematically productive experiences for students. In light of this research evidence, this is something that Pearson will discuss with DfE and Ofqual. In response to concerns about possibly overly ambitious entry at Higher tier, Pearson has produced tiering entry guidelines for teachers.

Provision for the weakest students at the Foundation tier is a different issue. Again, however, the study shows teachers taking a variety of approaches: either addressing a limited range of content, but a full range of mathematical processes, such as elementary reasoning and problem solving; or addressing a wide range of content, all at a superficial level. The former approach appeared to be aligned with more positive student experiences, as evidenced in focus groups, but teachers appear to have mixed views about the implications for students' grade outcomes.

Teachers were, however, positive about improvements to the accessibility of early parts of Pearson examination papers made after the first full cycle. This was a genuine enhancement of opportunity to demonstrate mathematical functioning, as evidenced by increased grade boundaries:

"I think the improved ramping is good because it builds on their confidence a little bit and they tend to have a little bit more time towards the end. ...And I think they're quite used to that as it gets harder, their brain starts to click in"

— Head of Maths 24, Spring term 2018

"[making the language easier] was important because you know we had found that students hadn't been able to access them before and they were put off too quickly, couldn't show what they could do"

— Foundation Teacher 21, Spring term 2018

"I think this year was more fair all round, to be honest. But still tough, but we didn't come out of it going 'that's horrible'. We came out of it going, 'yeah, there's some challenging questions and there needs to be and that's okay'. Just seeing much more at the right level, in our opinion, though for the weakest students, it's still very, very hard"
— Head of Maths, Autumn term 2018

But at the end of the second cycle it remained the case that the weakest students in the cohort were finding that the reformed GCSE Mathematics was unhelpfully demanding, and that they had limited opportunities to demonstrate prowess. While these constraints largely lie within the mandated curriculum and assessment criteria. Pearson will consider how this evidence can be used to improve the accessibility of the qualification for the weakest students, and inform the design of future qualifications at this level.

For more highly attaining students, at both tiers, the progression phases of the study show good (and improving) effectiveness of the reformed GCSE Mathematics over the first two enactments. There is no reason to suppose that effectiveness will not continue to improve. Possible areas that merit revisiting for effectiveness include the profiles for assessment of algebraic functioning for the most highly attaining students, and the continued assessment in context of core "functional" skills such as dealing with percentages in a meaningful way. Some such concerns might be addressed as assessment content sampling develops over time. If they persist, they might well be suitable candidates for national qualification criteria scrutiny and further development.

Within the confines of the criteria for GCSE Mathematics, then, this programme of research shows that Pearson has reason to be confident about many aspects of the effectiveness of the new specification. The research has, moreover, offered an opportunity to further enhance the support offered to promote access, equity, and empowerment among young users of mathematics. Actions taken so far include:

- Developing and implementing assessment principles from the June 2018 assessments onwards, to address the concerns around language and accessibility for the least able students
  - Detailed analysis and enhancement of the accessibility of questions, so as to privilege minimal linguistic demands consistent with valid assessment of the relevant mathematics
  - A more accessible start to all papers, but especially those at Foundation tier, resulting in enhanced facility of those early questions, and more confident teachers and students at Foundation tier
- Enhancing the range and structure of free qualification support available, especially in relation to problem solving and reasoning, and targeted at both teachers and students
- Developing freely available podcasts to address the teaching and learning of the more challenging areas of content and process
- Enhancing the frequency and geographical spread of both face-to-face teacher networks and drop-in sessions with expert teachers — again, supporting teaching in a way that is coherent with curriculum intentions
- Enhancing tiering support provided as part of the free qualification support, to provide teachers with confidence in tier entry decisions
  - Includes a package of diagnostic tests and exemplification of performance for grade 4 and grade 5
- Further improving the communication, ease of use, reach, and navigation surrounding Mathematics Emporium, given its centrality to classroom experiences in year 11, so that it becomes more easily useable for the range of teachers of GCSE Mathematics
- Enhancing student performance analysis tools to support better formative assessment of newer areas of the curriculum

Finally, it should be noted that the research itself was seen as indicative of Pearson's commitment to improving the quality of education associated with its qualifications:

"I like the fact that having been involved in the research, they listened, which was another positive"
— Head of Maths 8, Autumn term 2018

"They've added value to my teaching, as a non-mathematician"
— Head of Maths 24, Autumn term 2018

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Washington, DC: American Educational Research Association

Charmaz, K. (2006). Constructing grounded theory. A practical guide through qualitative analysis. London: Sage.

DfE (2013) Mathematics: GCSE subject content and assessment objectives available online at **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/254441/GCSE_mathematics_subject_content_and_assessment_objectives.pdf** last accessed 15th March 2019

Edexcel (2015) GCSE (9–1) Mathematics available online at **https://qualifications.pearson.com/content/dam/pdf/GCSE/mathematics/2015/specification-and-sample-assesment/gcse-maths-2015-specification.pdf** last accessed 18th March 2019

Mullis, I.V.S. and Martin, M.O. (2015) TIMSS Advanced 2015 Assessment Frameworks available online at **https://files.eric.ed.gov/fulltext/ED545251.pdf** last accessed 18th March 2019

Newton, P.E (2017). An Approach to Understanding Validation Arguments. Ofqual/17/6293

Ofqual (2017) GCSE (9 to 1) Qualification level conditions and requirements available online at **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/644333/gcse-9-to-1-qualification-_level-conditions-and-requirements.pdf** last accessed 18th March 2019

Ofsted (2012) Mathematics made to measure. HMSO, London

Remillard, J.T. (2005) Examining Key Concepts in Research on Teachers' Use of Mathematics Curricula *Review of Educational Research* Vol 75, 2: pp.211-246

Schmidt, W.H. and Prawat, R.S. (2006) Curriculum coherence and national control of education: issue or non-issue? Journal of curriculum studies Vol 38: 6, pp.641-658

Pearson