Jennie Golding (UCL IOE); Benjamin Redmond (Pearson UK); Grace Grima (Pearson UK)

***Theme*: summative assessments**

(Note overarching theme of *assessment reforms as journeys delineated by intentions and enactment).*

High stakes assessment that supports mathematical problem solving: a journey of realistic aspiration or of chimera?

From the first national curriculum in England in 1988, successive mathematics curricula for ages 5-16 have attempted to promote robust mathematical problem-solving, reasoning and communication for all. This aspiration, although enjoying broad support, has remained largely unrealised. The curriculum introduced from September 2014 again included a renewed focus on these key mathematical processes - this time within an increasingly high-stakes assessment system. National assessments in England are developed within a marketized system; they include GCSE Mathematics, taken by nearly all students at age 16, and high stakes for them as individuals, for their teachers, and for their schools.

We draw on three secondary school curriculum enactment studies of this curriculum and its assessment, two predating the covid pandemic and one 'New Normal' study probing emerging practices and learning post-pandemic. These all harness classroom observations and related teacher and student voice. We uncover a story of repeated attempts to support development of these processes via resources and assessment reforms initially well-aligned with intentions. We analyse the interdependent challenges of doing so – for assessment organisations, producers of curriculum resources, students, schools, teachers, and policymakers. We argue that systemic changes are needed if mathematically laudable aspirations are to be realised.

High stakes assessment that supports mathematical problem solving: a journey of realistic aspiration or of chimera?

We focus on attempts to reform assessment to support mathematical problem-solving within GCSE Mathematics in England. Successive national mathematics curricula, from 1988, have

attempted to promote mathematical problem-solving, reasoning and communication. This aspiration, although enjoying broad support, has remained largely unrealised. The current (2014) curriculum included a renewed focus on these key mathematical processes – this time within an increasingly high-stakes assessment system.

National assessments in England are regulated by the central 'Ofqual', but developed and delivered within a marketized system. GCSE Mathematics, offered by four independent Awarding Organisations including the research funder, is taken by nearly all 16-year-olds and is high stakes for students (as gatekeeper to many future pathways), but also for their teachers and schools, via a variety of school accountability measures introduced from 2010. The reformed GCSE, with a nominal weighting of 20-25% for problem-solving, was first examined in 2017, and is now entirely assessed by a set of written examinations (except in 2020 and 2021, when pandemic disruption to schooling catalysed alternative, school-led assessments).

The funder is well-placed to support curriculum coherence, since it provides GCSE Mathematics assessment, teacher-educative curriculum resources and teacher development; the studies drawn on here were funded in an effort to understand the use and impact of that offer, and potentially to enhance it. Studies were led by an independent academic (the first author), with classroom-close research undertaken by subject-expert, funder-external researchers.

Within the mathematics education literature, 'problem-solving' is generally taken to comprise those tasks for which the student has no known routine method of solution – so 'problem' is relative to the student and is likely to include novelty, limited structure, and complexity/multiple-steps. It requires well-organised and flexibly accessible domain-specific knowledge, heuristic methods for problem analysis and transformation, positive mathematics-rated affect, and related metacognition (Törner et al. 2007); it therefore draws on both mathematical reasoning and mathematical communication, and is complex to teach. In England, teacher education now requires little attention to subject-specific knowledge and pedagogy.

There is little historical experience of formally assessing mathematical problem solving at GCSE level. Internationally, we note that its valid and reliable assessment in short formal

tests has proved intractable (Mullis et al., 2021); additionally, its student-dependent nature adds challenge within assessments targeted at a broad range of students.

Funder curriculum resources and assessments were developed to meet curriculum and GCSE criteria, with support from an international subject-expert Advisory Board. Two challenges emerged: first, the absence of a definition of 'problem solving' in criteria led to different interpretations across Awarding Organisations, and secondly, early unstructured, novel and multi-step tasks, while supporting valued outcomes, proved overly-demanding in timed written examinations. Ofqual intervention subsequently resulted in more consistent, but less aspirational, enactment in papers across Awarding Organisations.

We explored classroom enactment, and the impact on learning of the funder's GCSE curriculum and assessment resources, in three studies spanning 2016-2022. Each drew heavily on teacher and student voice, and on full lesson observations. The first two were large-scale and longitudinal, the third a 'deep dive' snapshot of the 'new normal' post-pandemic in Autumn 2022:

A Secondary: age 11-16 (years 7-11)	**Classes followed for 2 years; termly interactions October 2016-September 2018**	**15+ schools/Heads of Mathematics (HoM), 32+ classes and teachers**
B GCSE Mathematics and progression from that: age 15-17 (years 11-12)	2 cohorts: October 2016-November 2018; termly interactions through students' year 11 into year 12	15+ schools/HoMs, 30+ GCSE classes and teachers, 32+ post-16 student groups, 32+ post-16 teachers
C The New Normal: age 11-16 (years 7-11)	Nov-Dec 22: 2-day intensive visit	5 schools/HoMs, 20 classes observed and teachers/students interviewed

Figure 1: Mathematics studies drawn on

Almost all participants were supportive of intended curriculum aspirations, but enacting them proved very demanding; misinterpretation of 'problem-solving' as 'worded questions' persisted through to 2022 in some classrooms. Emerging examination papers proved

hegemonic in interpretations of curriculum, with curriculum resources a secondary proxy. The high stakes nature of assessments meant examination changes catalysed prompt adjustment in teacher aspirations, and the marginalisation of more demanding problem-solving and reasoning tasks from resources and practice. Enactment of problem-solving in 2022 remained fragile and usually unambitious, with a prevalent discourse of 'basics and fluency before problem-solving'. Perceptions of curriculum overload and limited assessment demand for problem-solving, crowded out more aspirational processes.

This presentation highlights systemic variables that combine to undermine mathematically-desirable aspirations: high-stakes, marketised assessments, a content-heavy curriculum, and limited teacher education for such aims. We argue systemic changes are needed if mathematically laudable aspirations are to be realised.

References:

Mullis, I. V. S., Martin, M. O., Fishbein, B., Foy, P., & Moncaleano, S. (2021). *Findings from the TIMSS 2019 Problem Solving and Inquiry Tasks*. https://timssandpirls.bc.edu/timss2019/psi/

Törner, G., Schoenfeld, A.H. & Reiss, K.M. (2007). Problem solving around the world: summing up the state of the art. *ZDM Mathematics Education* 39, 353 (2007).