ICMI Study 24
SCHOOL MATHEMATICS CURRICULUM REFORMS: CHALLENGES, CHANGES AND OPPORTUNITIEs
Tsukuba, 26-30 November 2018

WHAT PRICE COHERENCE? CHALLENGES OF EMBEDDING A COHERENT CURRICULUM IN A MARKET-DRIVEN AND HIGH-STAKES ASSESSMENT REGIME

Jennie Golding

University College London Institute of Education

Curriculum coherence is widely valued as underpinning the enactment of a curriculum consistent with intentions. Such coherence depends on good alignment between, among other things, the written curriculum, available resources, the range of assessments, and teachers' capacity, including their knowledge, skills and affect. We report on a group of longitudinal studies mapping the early enactment of an aspirational curriculum in England which initially appeared to be supported by a coherent system, and show how the maintenance of that coherence was threatened by competing teacher beliefs in a market-driven and high stakes accountability regime.

BACKGROUND

Mathematics curriculum intentions in England and elsewhere (and what we know)

Intended mathematics curricula (Mullis and Martin, 2015) are being reconceptualised globally in an effort to meet the perceived needs of students and of society in the twenty-first century. Priorities for that are contested (Gravemeijer et al, 2017), but in England have focused on an aspirational deep fluency, accompanied by a renewed focus on mathematical reasoning and problem solving, and effective communication of those. Curriculum change in England is driven by education ministers in consultation with subject and other experts, and the political timescale of that model means attempts at curriculum change must be accomplished in a short timescale. England's first national curriculum of 1989 and subsequent high profile 'National Numeracy Strategy' were followed by significant changes to relative profiles of content and process in the national curricula of 1999 and 2007, in response to both perceived shortcomings in enactment and to changing political, economic and philosophical perceptions, including international attainment comparison studies and an influential review of student and employer needs (ACME 2011a, b).

The approach adopted has been to publish a new curriculum 5-16 (DfE, 2014) which distinguishes three levels of demand at 14-16, tested at two overlapping tiers; mandate assessments at 16 ('GCSE') consistent with that, although offered in a market of three competing Awarding Organisations (AOs); and encourage but not mandate the production and use of curriculum-compliant resources. At age 11, national attainment tests are produced centrally. Additionally, ministers have funded a semi-autonomous organisation to support mathematics teacher development (see www.ncetm.org.uk). Changes at A-level, the post-16 calculus-based mathematics route, have followed, with similar aspirations and first large-scale assessment in Summer 2019. Since GCSE Mathematics is high-stakes for schools and teachers, as well as for individual students, ministers expect their adopted approach to result in a valid enactment of their intentions, despite similar attempts historically, and elsewhere, having proved intractably challenging at scale (e.g. Eurydice, 2011).

THE STUDIES

This paper draws on a set of longitudinal studies undertaken by a team of 9 researchers led by me, and focussing on the enactment of this new curriculum at a variety of student stages. The studies asked how teachers and students were experiencing the curriculum, particularly in relation to the renewed foci, and how the resources and assessments produced by a major publisher supported or otherwise impacted that experience and the student outcomes. They also asked in what ways teacher capacity affected the answers to those questions. The first two studies focused on the teacher and student experience and impact of resource schemes that included digital teaching and learning packages as well as printed textbooks and workbooks etc, the GCSE study focused on the impact of summative assessments at 16 and the 'free surround' to those, and the A-level study, still in progress, is focusing on both a similar range of publisher resources and A-level assessments. The scale and scope of the studies and the data collected are summarized in table 1.

Focus	**Study**	**Size**	**Data**
Primary: 5-11 (y1-6)	2 years Oct 2016-Sept 18 (y1-2, 5-6)	9 schools and mathematics coordinators, 18 classes and teachers	18 pre- and post-class assessment data. *Yearly:* 25 Autumn, 18 Spring, 25 Summer/Autumn teacher interview transcripts, 18 lesson observation notes, student focus group transcripts
Secondary: 11-16 (y7-11)	2 years Oct 2016-Sept18 (y7-8 or 8-9 and 10-11)	15+ schools and Heads of Mathematics (HoM), 32+ classes and teachers	32+ pre- and post-class progression data. *Yearly:* 35+ Autumn, 32+ Spring, 35+ Summer/Autumn teacher interview transcripts, 32+ lesson observations, 32+ student focus group transcripts, 32+ whole class surveys (>800 students)
GCSE Mathematics and progression: 15-16 (y11)	2+ years: Oct 2016-Nov 18	15+ schools and HoMs, 30+ GCSE classes and teachers, 16+ post-16 groups of students and teachers	*Yearly*: 30+ Autumn, 30+ Spring, 15+ Autumn GCSE teacher/HoM interview transcripts, 30+ student focus group transcripts, 30+ whole class surveys, 16+ post-16 focus group transcripts, 16+ post-16 teacher transcripts, 30+ class GCSE results.
A-level Mathematics/ Further Mathematics: 16-18 (y12-13)	3+ years: Sept 2017-Oct 2020	12+ schools and HoMs, 24+ A-level classes and teachers	*Yearly:* 24+ class progression data, 24+ Autumn, 24+ Spring, 24+ Summer/Autumn interview or survey transcripts, 24+ lesson observations, 24+ student focus group transcripts, 24+ whole class student surveys

Table 1: Curriculum 2014 impact studies

All samples achieved reasonable representativeness over a number of variables known to influence enactments; nevertheless, scale was such that generalizability of outcomes cannot be claimed with confidence. Tools were developed iteratively, and data were analysed using a grounded approach (Charmaz, 2014). Coding was validated by at least one other researcher, and final interpretations and reports offered to teacher participants for validation.

THEORETICAL APPROACHES

Coherence across the education systems.

'Coherence' appears in the education literature in various guises, sometimes meaning little more than alignment of intentions of various learning-related actors of influence (Schmidt and Prawat, 2006). Better-defined constructs are developed in e.g. Newmann et al (2001), who both argue for and demonstrate the importance of 'instructional program coherence', within classes and longitudinally, for supporting improvement in student learning outcomes. Here, I use Schmidt and Prawat's (2006) definition of curriculum coherence as alignment of all elements of a curriculum system (intended curriculum and related documentation, assessments and accountability systems, teacher knowledge and skills, related resources, support of a range of informed or powerful stakeholders, high proportion of intended curriculum actually studied,...) together with underlying age- and stage-appropriate sequencing and progression. Without that, curriculum enactment undergoes 'iterative refraction' (Spillane, 2004) both horizontally and vertically.

The reported studies are predicated on a claim that significant efforts to achieve curriculum coherence were attempted by central government, supported in these cases by a large publisher and examining body. Importantly, the 5-16 curriculum for mathematics (DfE 2014), and the subsequent A-level Mathematics criteria, were widely perceived to be consistent with recommendations in ACME (2011a,b) and enjoyed a high level of support among those involved in English mathematics education, though felt by some to be overly-aspirational. They were perceived to be highly challenging for both teachers and students (Golding and Grima, 2018), particularly in relation to mathematical problem solving and reasoning, with considerable demands being made on both depth of subject knowledge (SK) and on subject pedagogical knowledge (SPK), a contruct developed for mathematics by Ball, Thames and Phelps (2008). In England, even highly mathematically-qualified and experienced teachers have often had fairly limited experience of developing these key renewed-focus mathematical processes in the classroom when using previous curricula, and that has often been attributed to summative assessments, though not intended curricula, that demanded little in these strands (e.g. Golding 2017b).

Teacher change for valid enactment

Clarke and Hollingsworth (2002) model teacher development as having an initial stimulus and then moving amongst different domains: 'the personal domain (teacher knowledge, beliefs and attitudes), the domain of practice (all professional actions, together with the professional context), the domain of consequence (perceived salient outcomes), and the external domain (sources of information, stimulus or support)' (p 949). Change within one of these domains can impact change in another by enaction, and by reflection. Thus, teachers learn and change through their professional activity, and within institutional systems that include choices of resources, teacher interactions, and approaches to external assessments. In the case of this curriculum change, the external domain includes curriculum documents, available resources, and external examinations; examination results are among the 'salient outcomes' in a high stakes assessment regime.

In England, most primary (5-11) teachers teach across the curriculum and few are specialists in mathematics. Secondary teachers are more likely to have studied a mathematics-intense discipline at university, but many have not, and those latter are typically allocated to younger students, or those who are less confident mathematically. Teaching preparation courses are typically of a year's postgraduate duration, and many have very limited subject-specific input. The intended curriculum therefore makes considerable demands on teacher development. Golding (2017a) conceptualises the teacher-level characteristics needed as 'teacher capacity for change', and that capacity sits within Clarke and Hollingsworth's personal domain, incorporating a range of knowledge, skills and affect. Fan, Liu and Miao (2013) show how textbooks can both support and limit teacher development for a new curriculum, and similarly Madaus and Russell (2011) address and evidence the impacts, positive and negative, high stakes summative assessments can have on classroom practices adopted.

FINDINGS AND DISCUSSION

I address four aspects of curriculum coherence: the communication of curriculum intentions, especially with respect to renewed curriculum foci of mathematical reasoning and problem solving, the impact and use of resources, development of teacher SK and SPK for the new curriculum, and the role of emergent sample assessments and examinations. In each case I draw from a subset of the studies, but also give a longitudinal account of development over the early years of enactment.

Communication of curriculum intentions

Almost all teachers across studies used proxies for curriculum documents: in primary and secondary schools those were often textbooks or (scheme-provided or school-developed) 'schemes of work'. Importantly, the publisher concerned used a team of internationally-respected mathematics education experts to support the production of both their resources and their initial examinations, so as to translate the stated curriculum intentions, themselves the product of comparative studies of international apparently successful practice, into high quality evidence-based artefacts. Teachers of students aged 14+ often used emergent sample or early live examination papers to triangulate textbook interpretations of curriculum, with official examination questions always 'trumping' those in textbooks or other resources: 'The textbook stresses mental methods as a first recourse, which is what I value, but the tests don't seem to want much of that, so we've changed our approach there' *(Year 6 teacher, Spring 2018).* Swan (2014) argues that greater exemplification ('beautiful examples', p 628) in curriculum specifications would pre-empt reductive interpretations in either resources or assessments – at classroom level or in production. Within such curriculum 'translations', a majority of teachers across the studies began with fairly conservative use of reasoning and problem-solving opportunities provided in resources, or the more challenging aspects of fluency ('mastery') that underpin the intended functioning with problem solving and reasoning, and often limited such wider, deeper experiences to their higher attainers. Probing these enactments in interviews, it became apparent there was also a variety of grasp of teacher interpretations of these fundamental processes.

However, as sample and early external summative assessments emerged, teachers began to talk about the challenges of teaching the weakest students in each tier of entry for reasoning and problem-solving questions: 'we've got to get them all so they can solve problems, which is a challenge, though the problem solving examples in the books really help, and the hints they give' *(year 8 teacher, Spring 2018),* or about perceived misalignments of textbooks with later papers: 'what we're seeing now, in these latest papers, is easier: we don't need to go as far as the textbooks do' *(Head of Maths, Spring 2018).* At A-level, though, whereas teachers still talked about the very real challenges of changing their teaching, and building up their professional knowledge, to accommodate genuine problem solving and reasoning, even early interviews, which followed on from first enactment of the new GCSE, commented on the learning already achieved by both themselves and their students in relation to these areas: 'Actually, even a year ago, I was thinking I can't do this, but it's amazing what the students have brought from the new GCSE, and for me too, that experience of teaching the new GCSE is underpinning my confidence to go for this' *(experienced year 12 teacher, Autumn 2017*). Coherence across year groups clearly bears fruit. These (largely specialist and mathematically well-qualified) A-level teachers unanimously valued the new expectations, while often simultaneously worrying about whether their weakest A-level students would be able to access papers: 'I don't know how I'm going to get them all to where they need to be, I simply don't know how to do it, but I do know that how I used to teach won't achieve that' *(year 12 teacher).*

Use and impact of resources

All participants in the first two studies, and many in the last two, used resource schemes produced by a leading publisher. All classes in the last two studies sat summative examinations from this publisher, one of three major GCSE and A-level AOs in England; many of those also used the publisher's resources ('the integrated offer'). Examination class teachers and students valued the coherence offered by adopting such an integrated offer, which eclipsed written curriculum specifications for almost all participants, though as perceived divergences among materials appeared, it was without exception the examination materials that prevailed.

In terms of pedagogy, use of the resources for most classes, and throughout the studies, was conservative, with more novel or creative features of printed and e-resources poorly harnessed. Limited use of e-resources was particularly conspicuous, with many teachers making no use of interactive features at all, for example (Evers et al, 2018). Teachers explained this with reference to the considerable demands of beginning to teach a new curriculum, feeling first, that there was insufficient time to explore new resources more fully, and second, that external examinations were such high stakes that experimenting with new approaches was in general too risky if it could be avoided. The range of participant teachers, even those with outstanding subject knowledge, talked about feeling insecure about teaching new topics, or preparing students for very different assessments. Such responses point to the considerable cost (in time, energy, money – and sometimes student learning) of coming to enact a new curriculum, especially one with associated high stakes assessment, even if an improved quality of education eventually emerges.

For non-specialist teachers at all levels, though, published resources supported confidence with planning for a new curriculum, and to a large extent, confidence with delivering that. In general, teacher positive affect (confidence, self-efficacy, enjoyment, motivation, willingness to take risks, persistence, etc) appeared to breed student positive affect (Barrow et al, 2018). Such confidence was not always well-placed: there was a range of quality of learning achieved in the lessons observed, with the strongest lessons, well-aligned with curriculum intentions and building deeply effective mathematical functioning in students, observed when teachers appeared to have robust capacity for change (Golding, 2017), underpinned by strong knowledge of the whole curriculum system, of appropriate mathematics and its pedagogy, and of their students. Such teachers, more than most, appeared able to tolerate perturbantions in emergent assessments or hiccoughs in student attainment; they supported the development of reasoning and problem solving through allowing student time to mull, to explore, to make mistakes, to harness multiple representations and resources, and supported them in learning to talk about their own thinking and listen to others'. Such pedagogical expertise takes time to develop, and necessarily draws on deep subject knowledge.

Teacher development for the new curriculum

Teachers using such resources can access both included teacher e-development units and additionally paid-for face to face support, but very few participants had taken advantage of that – or indeed, acknowledged their existence. Yet by international standards, these were not in general teachers with strong subject-specific knowledge or skills. Resource planning support also offers good subject development, e.g. in identifying prior knowledge needed, likely misconceptions emerging, and suggesting ways to address those, but many teachers did not recognize a need to use that support (Evers et al, 2018). They often worked with school colleagues to develop teaching for the new curriculum, though in some cases this was within a community of mathematically poorly-equipped teachers.

This was true even of changes to the A-level curriculum, where, for example, the introduction of work with (analysis/interpretation of) a 'large data set' – that is, one not susceptible to analysis via the use of a hand-held calculator – was a source of considerable concern to teachers, who typically had little experience themselves with authentic use of such data sets, often lacked the IT skills they wanted students to learn, and of course needed to develop the pedagogical skills appropriate to such technology-supported mathematics (Golding et al, 2018a). Different AOs have operationalised the requirement in different ways, usually drawing on the use of 'pre-release' material published in advance of final written summative examinations. It has to be remembered that for AOs also, significant curriculum changes can also bring challenges of demand for new types of assessment, as for the large data set – but still within a system that is high stakes for students, for teachers - and for AOs, representing large entry, high status examinations.

System development for such changes as the introduction of the large data set brings with it, then, a need for new teacher knowledge and skills. Resource writers, assessors and those providing a variety of support for teachers, whether as examination preparation support provided by the AO, resource-linked development, or otherwise channeled, all need to develop new skills and capacities. Without incentives for systemic investment in high quality professional development, it would appear that the validity of curriculum enactment for such an aspirational curriculum will inevitably be limited.

Role of emerging examination papers and sample assessments

Especially important in a high stakes assessment system is the quality of the assessment system and its alignment with curriculum intentions: I have described how influential emerging assessment materials were for teachers at all stages of the mathematics education system. As well as harnessing high quality external expertise to operationalise the curriculum via resources or initial examinations, the publisher also developed extensive 'free surround' support for teachers preparing students for mathematics GCSE or A-levels, and that was very much valued by teachers. Additionally, the national assessment body, Ofqual, invested, and continues to invest, in significant work, e.g. attempting to specify the nature of mathematics problem solving in ways operationalizable by AOs, and analysing the nature of student difficulties in tackling such problems (Ofqual, 2017): some aspects of central commitment to curriculum coherence remain high. The earliest examinations studied here, supported by external expertise, also showed high fidelity to curriculum intentions, but with resulting high demand on students and teachers (Golding and Grima, 2019).

However, the examination system in England at 16 and 18 is not only high-stakes but market-driven, and initial experiences of GCSE live papers suggested teachers perceived one AO to be offering papers much more accessible than others'. Ofqual responded with in-depth reviews and renewed criteria, but the result for the publisher studied, was the emergence of papers of lesser demand, but also arguably less coherent with curriculum intentions. Over time, that cycle has repeated, with a consequent growing gap perceived between the publisher's GCSE textbooks, for example, and their GCSE papers, as evidenced above.

Some participant schools have changed the AO used, for example for weaker students, since they perceive their papers to be more accessible, or more rewarding in terms of outcome grades: 'So now we use (AO x) for our two weakest sets: 'I like working with (AO y), I think they're mathematically better and they give teachers fantastic support, but at the end of the day these are more accessible papers and our kids feel they have done a better job. It doesn't make them better mathematicians, but it gets them better grades.' *(Head of Maths, Spring 2018).*

What we see, then, is a system in which, four years into a rolling introduction of a broadly-espoused new curriculum, considerable efforts by both central authorities and publishers to support curriculum coherence are being undermined by both the very high stakes nature of related assessments and a market-driven system of GCSE and A-level assessments which is leading to a downward pressure on aspirations. Teachers are challenged to respond ethically in such a situation: these studies offer evidence that many teachers, while limited in their subject-specific expertise, have been working hard to enact curriculum intentions while the system remains coherent, but they experience hierarchical and not always consistent beliefs relating to tensions between fundamental principles of mathematics education and the importance of external assessment outcomes. Any systemic change of the scale intended is expensive in many ways: if there is to be a net gain for student learning, it is important that ways are found to address such challenges to the coherence achieved.

Acknowledgement

The studies on which this work is based were funded by the Pearson UK Research and Efficacy team.

References

ACME (2011a). *Mathematical Needs: Mathematics in the workplace and in Higher Education*. Royal Society, London at http://www.acme-uk.org/media/7624/acme_theme_a_final%20(2).pdf.

ACME (2011b). *Mathematical Needs: The Mathematical Needs of Learners*. Royal Society, London at http://www.acme-uk.org/media/7627/acme_theme_b_final.pdf.

Ball, D. L., Thames, M. and Phelps, G. (2008). Content knowledge for teaching: what makes it special? *Journal of Teacher Education* no. 59(5), 389-407.

Barrow, E., Golding, J., Redmond, B. and Grima. G. (2018). "I get better and better all the time": Impact of resources on pupil and teacher confidence. In Golding, J., Bretscher, N., Crisan, C., Geraniou, E., Hodgen J. and C. Morgan (Eds). *Research Proc. of the 9th British Congress on Mathematics Education*.

Charmaz, K. (2012). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. Thousand Oaks: Sage Publications.

Clarke, D., & H. Hollingsworh (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education* 18, 947-967.

Department for Education (2014). *The 2014 national curriculum in England: Key stages 1-4 framework document* London: HMSO

Eurydice (2011). *Mathematics Education in Europe: common challenges and national policies*. EACEA.

Evers, K., Golding, J. and Grima, G. (2018). (Missed) opportunities for teaching and learning with digital resources: what and why? In Golding, J., Bretscher, N., Crisan, C., Geraniou, E., Hodgen J. and C. Morgan (Eds). *Research Proc. of the 9th British Congress on Mathematics Education*.

Fan, L., Liu, Y. and Miao, Z. (2013). Textbook research in mathematics education: development status and direction *ZDM 45*, 633-646

Golding, J. and Grima, G. (2019). Education for the twenty-first century: assessment to support the development of problem solving goals in mathematics curricula 5 to 16 *submitted*

Golding, J., Krishnaswamy, V., Redmond, B. and Grima, G. (2018a). Assessing mathematics education for intelligent use of big data *in press*

Golding, J., Green, C. and Grima, G. (2018b). *From 'jumping off a cliff' to 'it was pretty good': tier entry decisions in GCSE Mathematics* Presentation given to Ofqual Assessment conference, April 2018.

Golding, J. (2017a). Mathematics teacher capacity for change *Oxford Review of Education 43(2)*, 502- 517

Golding, J. (2017b). Policy critics and policy survivors: who are they and how do they contribute to a department policy role typology? *Discourse: Studies in the Cultural Politics of Education 38(4)*, 923-936

Gravemeijer, K., Stephan, M., Julie, C., Lin, F-L. and Ohtani, M. (2017). What mathematics education may prepare students for the society of the future? *Int J of Sci and Math Educ 15*, S105-123.

Madaus, G. and Russell, M. (2011). Paradoxes of high stakes assessment. *Journal of Education 190*, 21-30.

Mullis, I and Martin. M. (eds) (2015). *TIMSS 2015 Assessment Framework* Boston, Ma: TIMSS and PIRLS International Study Centre.

Newmann, F.M., Smith, B., Allensworth, E. and Bryk, A. (2001). Instructional Program Coherence: What It Is and Why It Should Guide School Improvement Policy? *Educational Evaluation and Policy Analysis* 23(4), 297 - 321

Oates, T. (2014). *Why textbooks count* Cambridge Assessment, Cambridge, England.

Ofqual (2017). *An evaluation of the difficulty of the assessments and the characteristics of the problem-solving (AO3) items* at https://assets.publishing.service.gov.uk/government/uploads /system/uploads /attachment datafile /668587/Summer_2017_GCSE_Maths_assessments_review.pdf

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

Spillane, J. P. (2004). *Standards deviation: How schools misunderstand educational policy*. Cambridge, Mass.: Harvard University Press.

Swan, M. (2014). Improving the alignment between values, principles and classroom realities in Li, Y. and Lappan, G. (eds.) *Mathematics curriculum in school education:* Springer, Dordrecht, 621-636