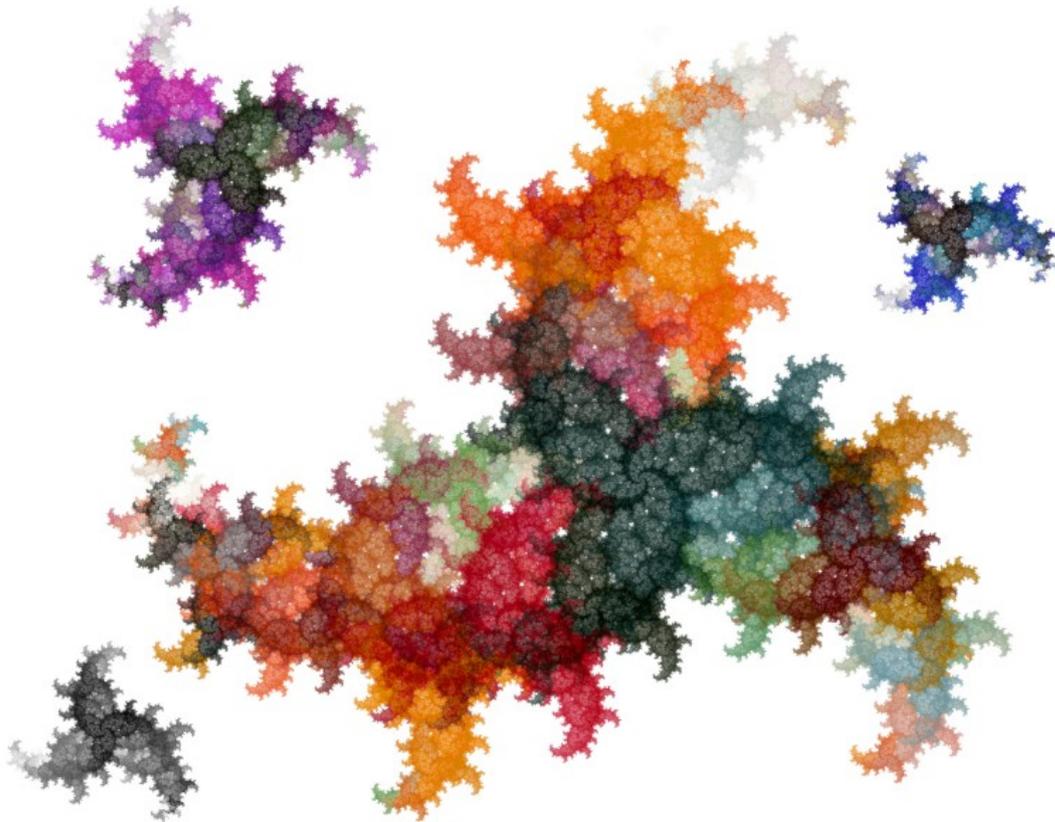# Re-envisioning STEM education: curriculum, assessment and integrated, interdisciplinary studies

Howes, A., Kaneva, D., Swanson, D., and Williams, J.
The University of Manchester

Re-envisioning STEM education: curriculum, assessment and integrated, interdisciplinary studies

Executive summary

The school curriculum and its assessment are ultimately determined by politics, but the debate can sometimes be informed by educational research, and evidence-based argument. This report draws on a literature review to inform a vision for reform of the school science and mathematics curriculum, one that would likely take 20 years.

The key, central strategy is to engage or re-engage children and youth in science and mathematics in ways that are (i) authentic and (ii) interesting and meaningful to the learners themselves. The literature shows above all that the system is failing in this regard today: the curriculum and its assessment are dull, encourage teaching narrowly to low-level fact-based tests, fail to encourage creativity, and switch the majority of learners off. Research has repeatedly shown, however, that learners and teachers can be excited by science and mathematics activity involving creativity, problem solving, modelling, and interest-led projects. Any vision for reform needs to find ways to implement such activities across the system.

Systemic reform requires attention on many fronts, beginning with assessment. Our review suggests that significant improvements in pedagogy and student engagement must shift the emphasis to 'formative assessment' in the classroom itself, and away from graded, exam-based 'summative assessment'. This is required in particular for higher level learning outcomes such as deep conceptual understanding and problem solving strategies, but is also a key to encouraging learners to take control of and 'regulate' their own learning. It can also allow the curriculum and teaching to develop around the learners' own interests, as is the aim of "science in society", "real problem solving in mathematics", and "projects" in technology and across STEM. For some that need to be re-engaged, arts-led, or social science-led projects involving STEM might be the way, such as those promoted by STE(A)M and other integrated or interdisciplinary projects.

However, this reform in turn requires teachers to develop and to get involved in assessment, and considerable effort will be needed to re-establish trust in the profession. Our review of international comparisons suggests that both these are required to improve the quality of classroom learning. All this requires R&D – but here R&D that centrally engages teachers themselves in improving the quality of learning and teaching, of the sort Japanese lesson study seems to promote.

We therefore argue for a reform of the curriculum, assessment, teaching and learning to re-focus on learners' interests, needs, and attitudes, rather than their grades, marks and league table positions. We suggest, for instance, that every student should be entitled to leave school with a portfolio of evidence of activity in STEM that demonstrates who they are, what they are interested in, and what they can do.

Re-envisioning STEM education: curriculum, assessment and integrated/ interdisciplinary studies

Howes, A., Kaneva, D., Swanson, D., and Williams, J., The University of Manchester

Introduction

This paper seeks to inform a strategy for the next 20 years development of science in schools, focusing on issues relating to curriculum and assessment: it is comprised of (i) a literature review, and (ii) a provisional vision statement.

(i) Literature review

Curriculum and assessment are topics for research and philosophy but also politics, and they are strikingly influenced by the outcome of the political debate. The associated tensions cannot be resolved by evidence or research, but they can be informed by them. This review focuses on literature concerned with STEM for school age learners (5-18) and in particular the curriculum, and its assessment. Within this, integrated and interdisciplinary working is of particular interest. Crucial for our literature search and synthesis, however, is the notion of 'vision' for STEM: what in today's literature might provide the seeds of a strategy for the next 20 years of development of curriculum and assessment in school science and mathematics? We conducted a database-driven survey of the last ten years of educational research to identify the state of play across STEM education, and constructed case studies to learn from significant processes, obstacles and innovative possibilities. In addition, we focussed a search on the apparent effects of different provision in other countries. In this first part of the report we review this literature, identifying some seeds for a vision. In the second part we suggest how a strategy might build on these seeds, while acknowledging areas of tension, conflict and challenge.

First, some key concepts in curriculum and assessment. A distinction in the literature is made between the written curriculum in national documents and exam syllabuses, and the 'intended' curriculum which various stakeholders intend should be taught (Remillard, 2005; Millar, 2011). Again this differs from the 'enacted' and 'experienced' curriculum, i.e. that enacted in practice in the classroom or lab, etc., and that which is experienced or perceived by learners. Finally, these are distinguished from the curriculum 'attained' as learning outcomes, which in turn may only be approximated by the results of assessments in various ways (as some outcomes may not even be assessed). In addition, we should bear in mind that the 'hidden' curriculum might shape much of what is learnt, for instance in attitudes to STEM. It is very difficult for research to establish an evidence trial that connects *all* these aspects of curriculum and assessment: most research deals with parts of the chain and the big picture requires visionary inference and synthesis. Here therefore we must warn against facile leaps – e.g. causally linking the written curricula (at one end of the chain) and the nationally aggregated results of attainment assessed by international tests (at the other) without due care for the intermediary variables.

In addition, all this work has to be read in the context of the politics and philosophy of curriculum: a key issue concerns the 'why' of curriculum (Black and Wiliam, 2005). Do we want the curriculum to shape learning for work, the economy and the future competitiveness of UK plc? Or do we want to prepare citizens to be active and critical in debates and decision making at every level? If yes to both, are these contradictory (Millar, 2011)? We could add others to this dichotomy, and note contending emphases on STEM 'for all' versus for the future elite, or for learning 'facts and procedures' versus 'problem solving and attitudes'. These are much discussed in philosophical, political and grey literatures, and the historical

case study in appendix B2 speaks to this: it was essentially local, cultural and political differences rather than evidence or research that was decisive in the case of the downfall of assessment of problem solving in Victoria as against its continuation in Queensland, Australia.

Much of the educational research literature we review here is empirical but assumes one or another such perspective. In this paper we will assume readers might mix interests in some measure, and indicate where there might be contradictions or resolutions. We also note that defining the 'content' of science or mathematics curricula has not been the focus of this research, beyond the general empirical finding supporting the position of STEM in the core of a balanced curriculum.

Figure 1: from Stein et al. (2007), Fig 8.1, showing the relations between curricula and assessment

Figure 8.1   Temporal phases of curriculum use.
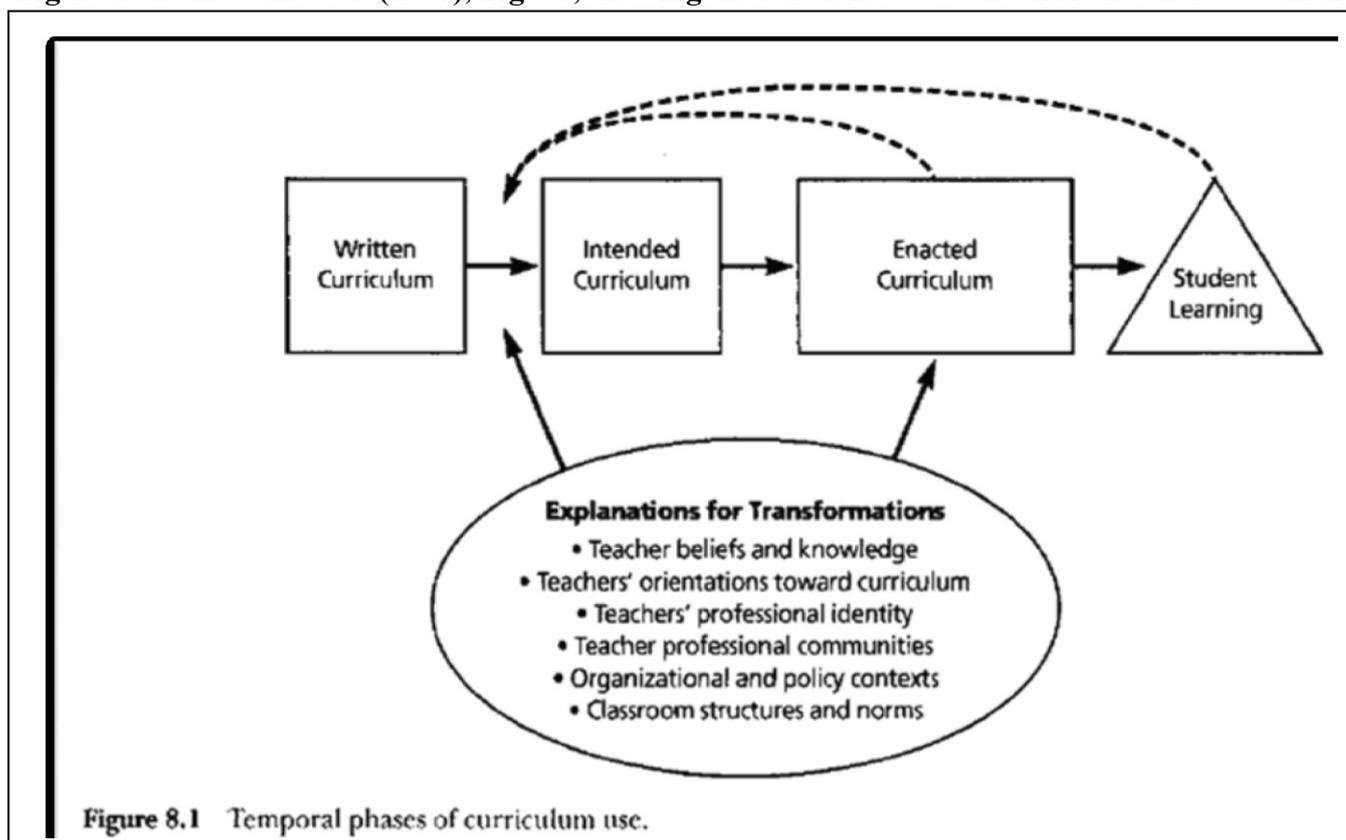
Inferences for a vision, however, must be made from fragments of empirical findings, and like perceptions must be constructed within a conceptual framework. As Lehrer (2011) claims of visual perception, the information flow from the brain to the eye is greater than that from the eye to the brain in one's construction of vision.

Take an illustrative example. Some children may come to believe that scientists are white males who wear white coats and work in labs; and some may believe mathematics is about doing hard calculations to little effect. What is learnt from the home and the media and what is learnt from schooling becomes enmeshed in these cultural models, (i.e. STEM is socio-culturally constituted in practice in a particular way in our society and in our times). Many learners do not choose STEM when they get the option, then, and of those who do, we know that girls tend to choose health sciences, and boys tend to choose technology and engineering. These effects correlate with surveyed attitudes of children from quite early on - well before STEM choices are made (e.g. see Figure 2). The facts show that some learners say they prefer to 'work with people' while others say they prefer to 'work with computers': and already at age 15 identities have formed alongside ideas about future careers. Such effects are present across almost all nations to a greater or lesser degree, but these effects can be changed over time, with sustained effort. For

example, respect for equity in Scandinavian cultures is thought to be reflected in the improving gender ratio of their STEM cohorts (Rasinen et al., 2009; Fredriksson et al., 2009; Van Langen, 2006; Hyde et al, 2008).

Figure 2: Showing gender differences in aspirations in STEM

Figure 14: The percentage of participating 15 year-old students expecting a career in a science, engineering or computing field by gender
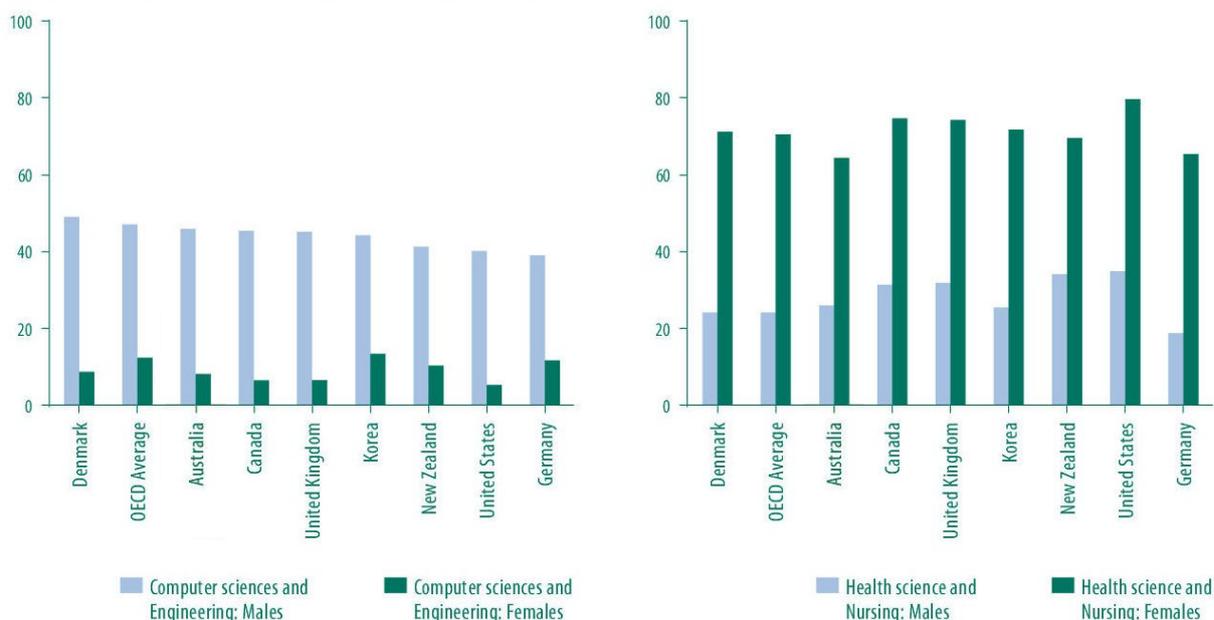
Source: OECD 2010e, Expectations for science-related careers by gender, *OECD Factbook 2010: Economic, Environmental and Social Statistics*, OECD Publishing, Paris.

An inference is then required suggesting a strategic vision for a STEM curriculum and assessment that widens access and deepens engagement of young people, one rooted in these findings: perhaps engineering and the physical sciences can be presented as subjects that *can* become 'good with people' as well as good for people (in the literature this is often described as becoming 'girl friendly'). They might usefully integrate with bioscience in medical, social or sustainability applications (Reiss, 2006); here chemistry claims to offer many possibilities for such connections and has a particular significance in medicine (Taber, 2010; Conboy and Fonseca, 2009). Similarly, it may be important for children to meet and hear about what mathematics graduates actually *do*, apart from teach (Williams, 2011); and perhaps the curriculum might usefully investigate inequities in the engineering industry rather than try to persuade women to enter it (Lyon et al., 2012; Sealey and Noyes, 2010).

In the above, we have extrapolated from some of the literature to address the issue of widening access to the curriculum, and this is our intended approach to surveying the literature. Given this, our direction of travel follows: in this part of our report we will outline relevant findings from the research literature, show how the literature (sometimes 'grey') makes sense of these, and draw out some seeds for the vision, with due caveats. In the second part we will be bolder with suggestions, but expand also on tensions, alternatives, and obstacles.

Methods

We addressed two overall questions: what is the state of play in the literature, and what innovative curriculum and assessment practice might contribute to a vision? For the first question, we adopted a

combination of survey and analysis of the research literature; for the second, we developed our thinking through particular case studies (further details on the method are given in appendix C).

A traditional literature search of recent English-language literature in refereed educational research journals, using key words and then selecting for relevance to the detailed research questions, produced a selection of some hundreds of papers: these were organised into themes, and in each of these we sought a recent meta-analysis or review that we could have confidence in. This survey then suggested some important gaps which we endeavoured to fill through wider searches. Additionally, we were concerned to focus on some specific questions about curriculum and assessment and integrated / interdisciplinary studies in different countries (see appendix A). Additionally the case studies invoked literature which the authors (sometimes with co-authors in conversation) deemed significant.

Research on effect of curriculum and curriculum reform
We draw on an excellent review of research on the effect of curriculum (Stein et al., 2007) which itself draws on over 150 citations. It identifies the methodological problems involved due to the way that the effect of written and intended curriculum on enacted curriculum is mediated by many complex factors including policy, teachers, management, and classroom norms. They specifically list: teacher beliefs and knowledge, teachers' orientation to the curriculum and their professional identity, their professional communities, organisational and policy contexts, and classroom structures and norms (p322). Note that teaching practice is central to all this. In educational research, the variables of greatest interest are often the most difficult to control.

Stein et al (ibid.) addresses what is argued to be the best curriculum effectiveness research to date, i.e. comparisons of the effect of 'reform' (i.e. US 'standards-based' reform) with traditional curricula on students experience and outcomes. On the whole, reform compares well with traditional curricula. But in practice this reform has very diverse meanings, even if it almost always involves some degree of 'problem solving and discussion' aimed at higher order learning challenges – and in middle schools it also can often involve a degree of integration. But they note here key limitations in this research:

(i)   there are very few genuine Randomised Control Trials involving students, and it seems never involving both teachers and students, as this proves impractical and ethically problematic;
(ii)   in about half the studies the measured effects of the differences were not statistically significant;
(iii)   evaluations whose measures centred on traditional / reform values favoured their respective curricula (i.e. the focus of the measurement used in evaluation is crucial);
(iv)   in most reports it is not possible to describe exactly what the enacted curricula were as well as measure their effects: usually there was an intention to enact the written curriculum as the teachers understood it, but there is evidence that many teachers are quite mistaken – or more accurately their views are quite at variance with those of experts- in their judgment of what the 'standards' are supposed to mean.

They conclude from this that (i) the curricula differ in significant ways, (ii) these differences *do* impact student learning; (iii) no curriculum is self-enacting (hence the 'mediating' variables); and (iv) the success of any reform curriculum is influenced by multiple factors, including the measures used to evaluate them. All this is supported by the literature evaluating the impact of "formative assessment", discussed in the next section.

Importantly one notes here several contextual factors: the 'standards' reform in the US was probably the largest and most comprehensively supported and evaluated systemic reform to date. No other similar research base is evident in the literature in relation to systemic change in STEM in other countries, where research and evaluation is even more contested (Penuel et al., 2009) . For instance, in the UK, national and international survey evidence of 14 years of National Numeracy (and other) Strategies agrees on a modest but significant rise in attainment at primary school level, and a fall in attitude, over the period of the intervention. However, there is as yet no clear answer as to whether it succeeded in really raising standards (e.g. Hodgen et al., 2010, 2013; Machin and McNally, 2009). There *is* a consensus that Primary children's standards in science did improve with the National Curriculum, but the evidence is weak.

Having said this, there is a body of work researching and evaluating smaller and shorter projects where the impact of focussed initiatives tend to achieve larger measurable effects (Ruiz-Primo et al., 2012; Marx et al., 2004 ; Furtak et al, 2012). There is a plethora of studies that evaluate relatively small scale innovations in the curriculum, but these all involve innovation in assessment as well, and we return to these later; but it is worth observing here that in every case reform in the curriculum designed to improve learning meets resistance from system inertia: in particular in (i) the assessment and examination ('what you test is what you get') (Suurtamm et al., 2008); (ii) teacher education and knowledge (Kidman, 2012); and (iii) institutional and system management practices (Marx et al., 2004). Thus, all these have to be considered in the reform if it is to begin to succeed at system level.

Additionally, where a system largely leaves assessment in the hands of well-prepared teachers, reform can happen relatively easily; on the other hand, where performance management of teachers is strong and dominated by a narrow set of traditional assessment data, the space for change or reform shrinks (Desimone et al., 2005; Labate, 2007). Even in such conditions, however, there may be room for development. In England for instance, one of the most closely externally tested populations in the world, the AS level "Extended project qualification" provides a (growing) number of students with an opportunity – however limited - to experience a supervised, but 'open' research project. (See case studies of innovations in assessment, B5a and B5b)

Assessment for learning - or not
In the area of assessment, one finds similarly an outstanding example of evaluation research of 'reform' in the case of Formative Assessment (FA): this research is rarely system-wide, but has been studied across many contexts in several systems internationally (Kim and Noh, 2010; Bull, 2010; Gioka, 2009; Ayala et al., 2008). Insofar as it complements the research on curriculum reform, it therefore adds to the above findings. We draw here on Wiliam's (2007) review (which also included 150+ citations).

Importantly for our review, Wiliam cites studies that show that the effectiveness of FA varies strikingly for different types of FA implementation: e.g. effect sizes from 0.16 for some 'weak' types, up to 0.5+ for 'strong formative assessment' types. (An effect size of 1 corresponds to a gain equal to a standard deviation of the population.) The latter involves teacher feedback to students identifying specific activities students should engage in to correct/advance their understandings, or more generally, "the creation of, and capitalization upon, moments of contingency in the regulation of learning processes" (Wiliam, personal website). This is a particular instance of the point made by Stein et al. about the significance of teachers' implementation of the reform in question (Gioka, 2009; Kingston and Nash, 2011).

Furthermore he points out that awarding grades/levels as part of the feedback, as in 'Assessing Pupil Progress' (APP) has a negative impact on the effectiveness of the feedback, and can effectively nullify any qualitative feedback designed to have a formative function (i.e. the learner sees the grade and ignores the comments!). Interestingly, the performance culture in England seems responsible for the evolution of formative assessment practices into the form of APP (Ardron and Monahan, 2010; Slade. 2009).

Further, evidence suggests that 'giving the answer/procedure or managerial information' when helping similarly nullifies the positive effects of offering 'elaborated explanations' in dialogues, so that teacher feedback can be less useful than peer feedback (Ding and Harskamp, 2011; Davis et al., 2007). Peer-tutoring has been found at least as effective (some effect sizes > 0.5) as regular teaching, presumably because of unhelpful teacher-feedback.

Other literature on assessment is consistent in pointing to its powerful effects on what is taught and learnt almost everywhere and throughout the school life-course, wherever it is prominent. In particular, high stakes external tests control teacher assessment (Hume and Coll, 2009; Li et al., 2006). In England this is a relatively prominent effect compared to Wales, and stronger than in the proposed Scottish system. In many other countries there is no high stakes external assessment until the end of lower secondary school or even the end of High School (in Finland teacher assessment is dominant throughout school, see more details in appendix A). We argue that relief from traditional high stakes assessment is a prerequisite for innovation generally and for the development of formative assessment in particular.

Innovations in assessment
There is a substantial literature involving attempts to innovate in assessment (Foegen et al., 2007; Kalyuga, 2006; Ashton et al., 2006; Fuchs et al., 2005; Kimbell et al., 2007; Kimbell, 2007): the key task is to find ways to assess a broader range of learning outcomes that can deliver a degree of politically acceptable reliability (see case study B2). The broad range of values might include the learner's creative work on open-ended innovative tasks, or their capacity to work together, or their attitudes and dispositions for further study (see case studies B5, and B6). The latter might be particularly important in view of the evidence that sometimes teaching to improve test scores can lead to declining attitudes (Pampaka et al., 2012a, 2012b). There is some positive work in this literature that suggests innovative assessment is possible, and can benefit from:

(i)   e-portfolios and comparative judgment techniques;
(ii)   group and peer assessments; and
(iii)   measuring a wider range of outcomes such as dispositions towards future study.

But much more research and development in practice is needed, particularly given the sensitivities to equity involved in the use of summative assessment to ration educational resources. Currently the resource going into R&D is small compared to the expenditure on assessment and testing, which is in turn a small fraction of the total cost of schooling.

Innovations within disciplinary curricula
Apart from those mentioned above, innovations in curricula within disciplines have been mainly concerned with two sorts of reform, often connected:

(i)   the introduction of 'real' (or 'hands-on' and 'practical') problem solving and modelling (Zhang and Campbell, 2012; Bintz et al., 2006; Marx et al., 2004); and

(ii)   the use of more challenging tasks requiring 'understanding' and application of concepts, or even critical thinking, rather than only rote learning of facts and procedures (Gijbels et al., 2005; Kanter, 2010; Pedersen et al., 2009; Fortus et al., 2004).

The term 'modelling' is particularly prominent worldwide in mathematics education, where it is conceived as an essential component of all problem solving reform especially with older learners, and those studying mathematics beyond compulsory school. However, it is also an important part of Primary school curricula, where research has shown how models such as the empty number line are important to children's building up mathematics in the Freudenthal tradition (see case study B4). In science, innovations often refer to 'hands on', but the process of 'modelling' is also used to describe the relation between theory and practical application (France et al., 2011).

There is also a body of work investigating mathematical practices in everyday life and workplaces which has implications for curricula (see case study B4): in everyday, non-academic settings mathematics is usually complex and messy, though not necessarily 'advanced'. Success relies essentially on one being disposed to see the world mathematically, analytically and rationally. When breaking for lunch at an upcoming seminar, observe how the coffee supply is spatially arranged to allow minimal access through one point at the end of a table, causing an unnecessarily long queue to form, rather than fractally arranged to allow multiple access points.

Then also science concepts model real world phenomena in ways that reveal insights and enable predictions to be made: its rationale can influence the way one sees the world. Understanding how science works in technology and society is the starting point of many curricular and assessment innovations. Local environmental investigations in the context of concerns about pollution, for example, provide a purposeful, collaborative context for development of laboratory, problem-solving and representation skills, particularly involving biology and chemistry, and for relating differently to the role of science in societal activity (Lee and Roth, 2001).

The literature points to gains from such reformed mathematics and science teaching, often in relatively small enrichment projects based on small numbers of schools: the most important gain seems to be in improving attitudes of learners and teachers, but there is also evidence in the literature of gains in higher order learning outcomes as one might expect when teaching is targeted at such outcomes (Manouchehri, 2007; French, 2004). The real challenge is to embed these enrichment projects in the everyday school curriculum and assessment system (See appendix B7.)

The case of 'Cognitive Acceleration' deserves mention as a strong study which has been widely followed up in England: the teaching of inquiry skills (including mathematical heuristics used in science) in supplementary science lessons led to significant gains some years later (Shayer, 1999; Adey et al., 1991). It is now widely thought that this was due to the experimental classrooms teaching children (and helping teachers) to critically inquire together; most significantly, the experimental gains spread across all GCSE subjects and were not restricted to science or STEM. Other studies in the cognitive acceleration frame have emulated this kind of experiment, including in the Primary school, and in mathematics curricula as well. These studies have also demonstrated positive gains but have rarely been conducted over such a long period (most research being completed within a project lifetime of a couple of years or so).

Research and innovation in integrated teaching and interdisciplinary STEM
If our reservations about evaluation/evidence of the effect of curriculum reform and formative assessment were hedged, "Evaluations of science learning that result from integrated programmes of work in schools

have produced notoriously ambivalent conclusions." (Venville et al., 2012, page 743). This is plausibly because 'integration' is even more difficult to pin down than formative assessment. STEM remains a misleading curriculum concept: it is not an integrated reality in high schools anywhere in the world that we know of, and STEM integration is not well understood by teachers (Brand and Triplett, 2012; Stinson et al.2009; Wallace et al., 2007). In many projects, the focus is on science and maths, leaving out engineering and technology (Hoachlander and Yanofski, 2011; Rickard and Wilson, 2006; Munier and Merle, 2009).

However, some meta-analyses show quite reasonable effect sizes even when the measure is traditional disciplinary knowledge. Thus, e.g., Venville et al. (2012) cites Hurley's (2001):

> "meta-analysis of quasi-experimental research that measured achievement in the science and/or mathematics disciplines. The results from 31 studies showed that, overall, student achievement effects for science were slightly larger than for mathematics (effect size d = 0.37 compared with d=0.27) ... Science achievement was greatest when mathematics was used in total integration with science or to enhance science. In contrast, both these forms had small effects for mathematics achievement. Student achievement effects were greatest for mathematics when it was taught in sequence with science, that is, when the subjects were planned together conceptually, but taught separately... (it also) noted anecdotal evidence that curriculum integration has a positive impact on attendance, student discipline, knowledge of academic resources, study habits, student enthusiasm and student engagement." (op cit p 743).

Ultimately however, one is left to wonder if it is the effect of innovation itself (at least in the short term) that is at work to produce such an impressive list of claimed benefits, raising doubts that systemic curriculum integration is the panacea for teaching the core disciplines (Applebee et al., 2007).

The sad case of the curriculum 'Yutori' in Japan, whereby 'room for thinking' was created and promoted through an expansion of Integrated Studies, is apt warning. Cave (2011) and others show how this remarkable system-wide innovation failed due to lack of proper investment in development, support, and infra-structure that might have facilitated genuine enactment of curriculum innovation. Essentially such reform requires that teachers get plenty of thinking time too.

On the other hand there is evidence from project evaluations that enrichment activities involving integration and interdisciplinary work can have positive effects on learners' interest in and attitudes to STEM (Riechert and Post, 2010). Tracing the history of integration of technology and science education, there are some conceptual similarities between 'enquiry' and 'design' processes (Lewis, 2006) and maybe also 'modelling' as well. Perhaps, then, the most significant goal of integration between science and technology or in STEM generally is for coherence of the curriculum for learners. Certainly, it is widely argued that meaningful integration of science and mathematics into technology must balance learning in those disciplines with clear intended learning outcomes from the integration or interdisciplinary work itself (Nikitina, 2006). This is evidenced in more vocational pathways (e.g. in Germany) (Marginson et al., 2013) but also in relation to the wider curriculum. E.g. some of these improved outcomes arise in helping learners to experience STEM as it is actually practised in work by 'real' engineers and scientists/mathematicians for instance (Kennedy and Schumacher, 2005; Meade et al., 2008). There are good reasons to integrate such workplace orientations into the STEM curriculum, so that students begin to appreciate the value a STEM career might have for them (Hodgen, 2013, cites persuasive literature evidencing this point).

However, experimentation has shown that STEM can be made meaningful to teachers and learners through change from the bottom up, with relatively little money for facilitation. Studies of collaborative teacher activity have found that effective pathways into STEM areas can be enhanced by a combination of: an integrated curriculum; changing teachers' practice; and more work-based learning (Brears et al., 2011; Schooler, 2004; Scott, 2012). Integration of STEM subjects in a school requires good partnerships between disciplinary staff. Integration must happen at the local level (Geraedts et al., 2006; Ritchie et al., 2012), and it requires innovation across STEM subjects to make the learner's experience coherent.

One focus for such integration has been the issue of sustainability (Clay et al., 2008; Villegas et al., 2010; Weyman, 2009). Sustainability education is a significant context also for Design and Technology activity (D&T); D&T focused in this way can provide a valuable space in which learners can develop both skills and knowledge, beyond simply practising or applying the core disciplines (Zoller, 2011). Evidence suggests that technology can be used to facilitate conceptual development in practice, with potential benefits for the learning of the disciplines: a mathematics infused technology curriculum has led to higher attainment in mathematics than when separated from a regular technology curriculum, for instance (Williams, 2010).

Integrated STEM can build from technology and engineering approaches that have contributed to educational engagement over the years. Engineering at school does not yet have a clear conceptual base, but can provide an authentic context for mathematics and science curricula. It is emerging as an alternative to technology education in some parts of the US. One engineering course moved High School students to engage in teamwork design processes over the course of a year for instance. Other cross-curricular themes (c.f. Primary and Middle school curricula) linking between separate subjects are possible (Bintz et al., 2006; Carrier et al., 2011; Romance and Vitale, 2012), and these might beeasier to implement than integration.

One particular project has been catching attention: a US originated project calls for STEM to integrate with Arts into a "STEAM" curriculum: and this has been picked up by the Korean government who have instituted a curriculum and teacher training programme on this basis. The rationale explicitly draws on the Steve Jobs formula of smart technology with 'cool design' (thus appealing to the agenda in Pacific rim countries for an innovation-based reform of education for a competitive knowledge economy). There is as yet not much research evidence to go on, but this formula might have great attractions for Arts-led work with learners who are not predisposed towards or interested in STEM, but who might thereby maintain contact with STEM in some form. Rather than seeing such students as having leaked from a STEM 'pipeline', perhaps we should see them as travelling by a different route, to a later rendezvous. As described in appendix B7, this approach could be developed initially as an enrichment that might subsequently become embedded in the mainstream curriculum, and there are such initial attempts described in the literature (Benedis-Grab, 2011; Costantino et al., 2010; Duncan et al., 2008; Stellflue et al., 2005; Suzuki, 2009). Although this is a long way from becoming systemic, a vision for STEM might include this approach if only as part of a science-for-all strategy that seeks to keep trying to reintroduce students to the attractions of STEM. We take up this thrust in our vision for school STEM below.

New technologies in STEM education
There has been a great deal of research and development (R&D) reporting the use of new technologies for enriching teaching and learning in science and mathematics, and STEM in general (Jones, et al., 2004;

Doherty et al., 2011; Buxton, 2007; Erhardt, 2007; Ruggirello et al., 2012; Carvalho, et al., 2008; Petre and Price, 2004;Schaefers, 2006; Danaia et al., 2012). Many innovations have made knowledge more accessible to learners, e.g. allowing them to inquire and explore concepts through powerful software, simulations and 'augmented reality' (Harlow and Nilsen, 2011; Chang et al., 2008; Ketelhut, and Nelson, 2010; Urban-Woldron, 2009). There is hardly an area of mathematics and science that cannot be so enriched, in principle.

To mention some really powerful examples being still researched on a case study basis: LOGO-type explorations and programming; geo-gebra software; modelling in chemistry and biology, and robots and solar-powered vehicles (Dori and Kaberman, 2012.; Wilder and Brinkerhoff, 2007;Klopfer et al., 2009.;Hohenwarter,et al., 2009). However, although the order of magnitude may be different, one notes that this has been the case for over 30 years: like gender differences in engineering it seems that under-usage of new technology in education just won't go away. Why? Research has repeatedly found that many of these powerful technologies and software packages (i) have too great an 'entry fee' in time required to use them effectively, either for teachers or learners, or both (i.e. to gain sufficient mastery to become effective); (ii) can be too costly for the schools to purchase and properly maintain; (iii) can be too diverse for effective instruction (e.g. the many different 'makes' of calculator); or (iv) do not become universal enough to embed them in assessment at institutional and system levels (Wachira and Keengwe, 2011).

However, it may be that the new generations of technologies will really become trivially available and accessible through mobile technology apps for all learners. Consider: Khan academy type tutorials and lectures for learning procedures and facts; applets/simulations for illustrating concepts in science; downloading data and other resources from the web for modelling; and accessing programming languages, drawing packages and 3D printing, via YouTube type demonstrations of everything instrumental the learner needs (Charles and Rice, 2012).

Even so, a narrow curriculum-and-assessment system imposed by test-and-inspection surveillance may successfully inoculate teaching from such developments, and might force them out of schooling into the homes of those privileged with the necessary resources, both financial and cultural. We argue that any strategic vision will need to find person-centred curriculum time and space where learners can access such resources and play with them.

Global/ large scale assessment and international comparisons
One way to look at global assessment data (e.g. PISA and TIMSS) is to think of comparator countries as those 'similar' to the UK but performing better: thus among the Anglophones England does about the same as Scotland and the US average, but Canada and the US-Massachusetts somewhat better. On the other hand among the industrialised EU economies: England does about the same as Germany but overall less well than Northern Ireland or Belgium. Differences between PISA and TIMSS tests are interesting as the former emphasise 'literacy' in ways that reflect curriculum policy conflicts. Thus, some schools here actually use PISA questions in class as means to broaden their learners' curriculum experiences.

But international comparisons in performance can be very difficult to interpret: there is a seductive tendency to cherry-pick those features of high performing systems that agree with author's well-formed prejudices (the down-side of what we have called making well-judged inferences from fragmentary findings). In addition, some features are shared by high and low performing systems (Hodgen et al.,

2013). The one clear explanatory variable of a country's success in league tables is a country's socio-economic performance (Meyer & Schiller, 2013).

Indeed, there are few educational system characteristics common to ALL high performing countries, and various authors have identified at least two very different characteristics of or 'routes to' their success: the 'social democratic' educational cultures emphasising equity (Canada, Finland) and the more 'traditional' educational cultures (Japan, Singapore, Korea) that sometimes appear as explanations of nations' performance at the top of league tables (Marginson et al., 2013). Of course it is more complicated than that: for instance, observers have remarked that the classroom practice in Finnish classrooms is highly 'traditional' (Andersen, 2010; Alexander, 2003), based on transmission of the text book whereas in some traditional cultures classrooms are relatively interactive – even dialogical - and intellectually challenging (Japan) (Roth and Garnier, 2007). Other characteristics, such as the near-universal education in mathematics up to age 18, are common to both high and low performance (e.g. Finland and Israel). Countries that encourage a broader range of subjects post-16 including mathematics or science do tend to get higher participation rates in STEM post-18 (Hodgen et al., 2013).

In trying to formulate characteristics of other cultures that might be relevant to change and improvement in the UK one looks to features that have been shown to 'transfer' in some form (e.g. 'lesson study' types of action research have transferred from Japan to the US and Canada (Lewis, Perry and Friedkin, 2011), and not just Pacific Rim countries) or that seem amenable because they might be culturally acceptable in the English context , given we are relatively – not to say deceptively - close in history, culture, and language (e.g. from Scotland or maybe Australia, New Zealand, and perhaps Canada and the US to England). Thus we can look to Scotland and New Zealand to see how we might increase participation in advanced mathematics and hence keep open access to STEM after 18: what is needed is an attractive and valued course for those that do not opt for the traditional A-level associated with STEM specialisation (akin to modelling with statistics in New Zealand). We note that the post-16 Use of Mathematics course has had some success here in making mathematics accessible (Pampaka et al., 2012) and that the *Mathematics for Education and Industry* project has begun work on another new modelling course for mathematics in England.

Many characteristics are common to most OECD countries, whether successful or not: compulsory mathematics and science to age 15 or 16, national examinations at end of compulsory school, etc. (OECD, 2009). Some apparent formal differences (e.g. age of beginning 'schooling') may be less real in terms of enacted classroom practices (e.g. teaching of science in kindergarten) than it appears (Ontario Ministry of Education, 2011). Similarly, the fact that the National Curriculum is no longer legally operative in much of England's schooling has not yet led to enacted curriculum differences, perhaps due to continuities in assessment and regulatory systems. Some characteristics are notably distinct in national systems but in themselves are hard to interpret: e.g. the Japanese written curriculum is specific and brief compared with the hundreds of pages of detail in California (California Department of Education, 2004), yet the curriculum in Japan is enacted in a relatively homogeneous way.

There are a few characteristics common to most of the successful countries, including: a professed national culture of respect for education and teachers' professionalism; a strong commitment to teachers' disciplinary knowledge; relatively well resourced education systems with proportional investment; and relatively highly qualified teachers (e.g. Korea, Japan, Finland) (Marginson et al., 2013; OECD, 2009).

Some contested issues remain: in some countries there is minimal examination and test-driven management, especially in the earlier phases of education (Canada, Finland, Japan and, increasingly,

Scotland) while some others are highly driven by test performance (Singapore, England) (IAEEA, 2011). Some of the policy perspectives have been well-aired in a recent Australian report that we summarise in appendix B1 (Marginson et al., 2013).

Discussion, conclusions, and recommendations

The most significant innovative developments in curriculum and assessment in STEM for the past 20 years have been to raise the expectation of students' opportunity to learn to include (i) problem solving and applications/modelling skills, investigation and creativity; and (ii) depth of understanding of concepts, and of intellectual challenge; (iii) 'talk', dialogue and argumentation; and (iv) good dispositions, attitudes, and self-knowledge, and so, 'learning to be a good learner' of STEM.

Many projects and much R&D has been expended round the world to show that these can be made to happen in practice, at least locally and periodically, and in managed tension with an emphasis on disciplinary knowledge (Marginson et al., 2013). On the other hand, innovations in this direction also tend to be degraded by time, systemic inertia, unhelpful changes in assessment, and occasionally other retrograde policy. The international evidence shows that some aspects of education systems can help in this regard, and probably no innovation in STEM will succeed while getting these elementary system factors wrong. Alas, these may be the stumbling block in England.

Cultures that respect education, learning and teaching (e.g. Confucian, but not only these) seem crucial – hence the adage that if you want the nation to take science, argumentation and evidence seriously, "get a blackboard in the House of Commons". High performing educational cultures also tend more to a culture of 'trust' in the teaching profession (Scandinavia; Japan) compared to the trend in Anglo-Saxon and many others towards inspection 'accountability', which can reduce risk-taking teaching behaviours.

A corollary is that many nations that prioritise education recruit the best graduates for teaching (Finland, Korea, to some extent Japan and others), and support their education and their education system in general financially (England is now approximately a median OECD country in %GDP funding for education as a whole and for universities for instance). We argue that the optimal governance of education would ensure that the profession is taken seriously, is expected to develop itself to a 10-20 year programme of goals, and prefers not to leave education to the political short-term (e.g. electoral timetables) or fluid market factors (e.g. text books and assessment left to private companies). See appendix B3.

Systems with a dominant emphasis on teacher assessment (such as Ontario and, increasingly, Scotland) create space for reform, such as a more open, problem-solving pedagogy and subject integration/inter-disciplinarity. The later that external, summative assessment can be delayed the more assessment might be formative, i.e. serve teaching and learning, and support intrinsic qualities and good educational dispositions. We argue that it is especially important to keep Primary or even middle school years free from teaching to the test.

The emerging world-wide consensus in educational reform does not universally extend to the public and policymakers; indeed in some countries there are moves towards a more discipline-based curriculum. But among educational specialists and researchers, a strong and developing contribution from STEM education is seen to include the following combination of requirements to act in concert:

- A balance in the curriculum and between lower level (facts and procedures) and higher level knowledge (conceptual, application and problem solving and metacognitive knowledge);

- The crucial importance of R&D support to provide curriculum materials and professional development tailored to programmes of reform, and the need for educational research, professional practitioners, and policy to work together in concert;
- Educational technology innovations that make science and mathematics topics and concepts more available/accessible;
- A balance between disciplinary curricula (science, mathematics, technology-computing) and problem-based, design, innovative or creative problem-centred (sometimes interdisciplinary) work;
- A balance in assessment consistent with this balance of curricula, and so also between summative paper-and-pencil facts-and-procedures testing/grading and project-based and alternative, practical and oral assessments.

(ii) A vision for the future STEM curriculum and assessment

How to start building a politically realisable 20-year vision for STEM that has roots in current R&D practice and in this evidence?

We argue the *key* aim should be to widen participation (potentially for all) and deepen engagement (in higher level learning) of students in STEM; attitudes and dispositions are crucial to this. But the curriculum must offer opportunities at all levels for students to engage or re-engage in meaningful science as a means to their educational development – it must never be 'too late now' for students to come back to science.

We note the evidence that genuine system change and improvement in curriculum and assessment often takes place on a time scale of decades, in the context of strategic leadership and R&D involving the whole profession in a relation of mutual trust. In the above review we have identified seeds that seem promising for development of the curriculum and assessment: e.g. in the influential US standards 'reform'; formative assessment based on dialogue and learners self-knowledge; in many and diverse 'modelling and problem solving' and 'science-in-society' projects and broader assessments; in technology- and engineering-led projects drawing in mathematics and science, with assessment fit for purpose, e.g. involving e-portfolio and comparative judgment assessments.

The main task here is to bring these together in a vision that seems realisable even though perhaps challenging.

Strategic leadership is vital. England is currently engaging in a very large, very loosely controlled curriculum experiment, with nearly half of English secondary schools freed from National Curriculum and many other formal constraints. There are no obvious plans for the reversal of this trend by other political parties, though one might expect it to be moderated by regulation and perhaps even local accountability. In principle this may provide opportunities for development and experiment that a new 'vision' might energise: but in such circumstances teachers and schools might become lost. In this political context, a respected STEM organisation or partnership of such organisations – with governmental support - might lead with a strategic vision for STEM, which in turn might be picked up by some groups of schools.

High quality research and development will be required. Educational research and teacher education arrangements are being radically shaken up with the intention that they become aligned with impact on

system policy and school practices. Thus educational research will feel pressure to align with R&D to survive. A vision that is both experimental and developmental might find resonance in educational research as well as in at least some school networks, as in Japanese lesson study (see appendix B3). A thousand flowers might bloom, and at least a few might be well researched. In the education business the spend on R&D remains the lowest of any knowledge industry, but this lack of confidence in educational research might be turned round if educational research firmly addresses problems of innovation and change. Educational researchers and institutions will rightly argue that an essential part of this vision must be rigorous research of real excellence if it is to identify and theorise practices that have potential for system change. Curriculum and Assessment, along with the education system and profession as a whole, needs to be subject to continuing research, development and improvement.

In line with the worldwide trends researched in part (i), we argue the following are essential to **improving the teaching and learning cultures**:

a.   The teaching profession needs nurturing and support; it needs to be trusted and loved;
b.   The bedrock of improvement is in the quality of teaching and teachers: any policy must recognise the importance of recruitment of potentially good teachers, and the lifelong professional development of teaching and teachers;
c.   Teachers and teaching develop best in schools and institutions (including professional networks and associations) that nurture a culture of continual improvement. This is best understood as a collective enterprise involving action research by teachers, shared collectively with professional researchers; and
d.   Integrated or interdisciplinary teaching can – along with disciplinary co-teaching/coaching – provide a rich resource for development of teaching and teachers, and a student- and person-centred curriculum that can engage all students.

A particular vision: integrating STEM into innovative student-centred projects
Our vision is that all students in 20 years time would leave school with a portfolio of 'projects' that shows who they are and what they can do as a result of a rich curriculum experience – based on a flexible, partly-integrated curriculum such as those in e.g. California, Ontario, Finland (see Appendix A). These students will enter schools in seven years from now, and there is a lot to be done to prepare the education system in such a short time.

The STEM curriculum requires both a "STEM for all" approach that keeps all students interested and involved, and further opportunities and encouragement for those keen to focus on STEM. The core disciplines in STEM will therefore need to be strengthened - adding opportunities to motivate STEM learning through its connections to enjoyment, aesthetics, societal needs (e.g. sustainability, ecology, social responsibility etc), and real problem solving that can engage all students' interests. This kind of work is being developed all round the world and has strong arguments as well as a base in R&D.

a.   Students might benefit from a core D&T or STE(A)M curriculum 'for all', but one in which aesthetic or creative projects are also included, while applying STEM;
b.   Students also need opportunities in the mainstream curriculum disciplines to work on quality projects that offer access to the higher order learning outcomes including conceptual understanding and problem solving/inquiry skills and dispositions ;
c.   Whatever the weaknesses of Extended Project Qualifications, these do show that an examination structure can allow institutions/schools to resource such open ended work. Any

vision for STEM needs to find 'elbow room' for problem-centred interdisciplinary/research project work which can be examined and accredited. See the Case study of Netherlands assessment where such work has rapidly expanded system-wide (appendix B6).

d. In addition, there is a need to improve the quality of texts and resources. A 'trip advisor' system based on teacher ratings might be introduced, with professional associations kite-marking or reporting their views on their rigour.

Inside but also outside the core discipline areas, we envisage interdisciplinary, integrated 'project work', facilitating outcomes of 'creativity' and 'problem-centred' learning thought to be required for an 'innovation' education and economy. This would include team and individual working, structured and unstructured tasks, formally and informally assessed by STEM criteria aiming always at a product such as an artefact, a solution to a real problem, a research project, a performance, or a report, etc. We are arguing here for an integrated studies space: the creation of 'space for thinking', but space that is richly supported and structured. These projects should be motivated by the students' interests (whether making a new kind of xmas card, or investigating the pollutants in the river next to their school, re-designing a park, or researching Escher's work) but enriched by STEM. Such work would need tutorial supervision, and make new demands on teachers to help students research and show how powerful STEM disciplinary knowledge can be. This creates an ambitious agenda for curriculum development and professional development: the Japanese yutori experience suggests that to go system wide on such a programme without adequate preparation would be dangerous. (Cave, 2011).

Crucially, assessment practice needs to be shaped to support these curriculum goals and developments. New, more valid assessment is needed that assesses higher level knowledge (understanding and application, problem solving and creativity) and of a broader range of learning outcomes (such as student attitude and dispositions). This requires research and development. Currently, most tests and summative assessment focus too much on low level facts and skills, which distort the enacted curriculum. Positive attitudes and dispositions and higher level cognitive objectives require that Assessment for Learning is protected from high stakes narrow-gauged testing, grading and accountability 'league tables' for as long as possible in early education. This requires development of teacher assessment that is validated by survey sampling and other rigorous moderation- see appendix B5a on innovative assessment. It also requires a broadening of the range of outcome measures to include attitudes and student destinations, see appendix B5b.

The Extended Project Qualification provides one current assessment space, but its criteria are entirely process focussed. We argue that project space needs to be posited in the context of a rich teaching-and-learning experience including eventually summative assessment against quality criteria that include STEM as well as process criteria. Similarly the e-scape project showed how portfolio assessment can provide reliable and valid assessments of D&T project work involving creativity, but these were not enriched or assessed by science and mathematics criteria. A STEM equivalent requires rich input and assessment from science and mathematics. Since D&T is or was one potential way forward in this respect, one could develop this 'project space' by integrating the core subjects into an enhanced, STEAM or STEM-enriched design-arts-technology course.

In post-16 education we argue that a norm could be established that most of the population study some STEM, and that a STEM core is included in all students' programmes. In adolescence interests increasingly include social relationships, and work. Course units related to these might draw in students who would otherwise leave school, and for those in education might enhance their experience, confidence, and future access to learning in the sciences (social or natural). The pre-university integrated

courses in Canada are interesting in this regard (see appendix A). One envisages, for example, a STEM-infused "architecture" course being developed (in collaboration with university architecture lecturers) leading to a design project- the outcome would not just be a grade and UCAS points, but an object and an e-portfolio that the student might use in university applications and interviews. We anticipate that trust in assessment in general will be a key concern for such innovations: we have pointed to some innovations such as that in e-portfolios and comparative judgment which have shown promise, and after all university architecture students' work is assessed in similar ways. There are ways of ensuring that teachers don't connive with students to deceive the summative assessment, but more R&D will be needed. Ultimately in such work there has to be a development of trust in teachers' professional judgments, and it is of course important to disentangle the assessment of teachers from the assessment of their learners.

References

Alexander, R. (2003). *The education of six year olds in England, Denmark and Finland: An international comparative study*. Ofsted.

Andersen, F.O. (2010). Danish and Finnish PISA Results in a Comparative, Qualitative Perspective: How Can the Stable and Distinct Differences between the Danish and Finnish PISA Results Be Explained? *Educational Assessment, Evaluation and Accountability*, 22(2), pp. 159-175.

Applebee, A.N., Adler, M. and Flihan, S. (2007). Interdisciplinary Curricula in Middle and High School Classrooms: Case Studies of Approaches to Curriculum and Instruction. *American Educational Research Journal,* 44(4), pp. 1002-1039.

Ardron, K. and Monahan, S. (2010). Assessing Pupils' Progress: Keeping It Real in the Primary Classroom. *Primary Science*, (115), pp. 9-13.

Ashton, H.S., Beevers, C.E., Korabinski, A.A. and Youngson, M.A. (2006). Incorporating Partial Credit in Computer-Aided Assessment of Mathematics in Secondary Education. *British Journal of Educational Technology*, 37(1), pp. 93-119.

Ayala, C.C., Shavelson, R.J., Ruiz-Primo, M., Brandon, P.R., Yin, Y., Furtak, E.M., Young, D.B. and Tomita, M.K. (2008). From Formal Embedded Assessments to Reflective Lessons: The Development of Formative Assessment Studies. *Applied Measurement in Education*, 21(4), pp. 315-334.

Adey, P., Shayer, M. and Yates, C. (1991). *Better learning. A report from the Cognitive Acceleration through Science Education (CASE) project*. London: Kings College - University of London.

Benedis-Grab, G. (2011). Making Art with Microscopes. *Science Scope*, 35(2), pp. 36-41.

Bintz, W.P., Moore, S.D., Hayhurst, E., Jones, R. and Tuttle, S. (2006). Integrating Literacy, Math, and Science to Make Learning Come Alive. *Middle School Journal*, 37(3), pp. 30-37.

Black, P., & Wiliam, D. (2005). Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal,* 16(2), pp.249-261.

Brand, B.R. and Triplett, C.F. (2012). Interdisciplinary Curriculum: An Abandoned Concept? *Teachers and Teaching: Theory and Practice*, 18(3), pp. 381-393.

Brears, L., Macintyre, B. and O'Sullivan, G. (2011). Preparing Teachers for the 21st Century Using PBL as an Integrating Strategy in Science and Technology Education. *Design and Technology Education*, 16(1), pp. 36-46.

Bull, A. (2010). Refocusing Assessment in Science. *Primary Science*, (115), pp. 22-24.

Buxton, R. (2007). The Laser Cutter: A Terrific Addition to Your Tech Program. *Tech Directions*, 66 (6), pp. 22-23.

California Department of Education. (2004). *Science framework for California public schools, K-12*. CDE.

Carrier, S., Wiebe, E.N., Gray, P. and Teachout, D. (2011). BioMusic in the Classroom: Interdisciplinary Elementary Science and Music Curriculum Development. *School Science and Mathematics*, 111(8), pp. 425-434.

Carvalho, C.C., dos Santos, M. J. and Marques, B. M. (2008). A Time-of-Flight Method to Measure the Speed of Sound Using a Stereo Sound Card. *Physics Teacher*, 46 (7), pp.428-431.

Cave, P. (2011). Explaining the Impact of Japan's Educational Reform: Or, Why Are Junior High Schools So Different from Elementary Schools? Social Science Japan Journal,14, 2, pp 145–163.

Chang, K., Chen, Y., Lin, H. and Sung, Y. (2008). Effects of Learning Support in Simulation-Based Physics Learning. *Computers & Education*, 51 (4), pp. 1486-1498.

Charles, K. and Rice, O. (2012). Emerging Issues: Open Educational Resources--How Science Teachers Can Use Open Educational Resources to Revitalize Lessons. *Science Educator*, 21 (2), pp. 55-56.

Clay, T.W., Fox, J.B., Grunbaum, D. and Jumars, P.A. (2008). How Plankton Swim: An Interdisciplinary Approach for Using Mathematics & Physics to Understand the Biology of the Natural World. *American Biology Teacher*, 70(6), pp. 363-370.

Conboy, J.E. and Fonseca, J.M., (2009). Student Generated Recommendations for Enhancing Success in Secondary Science and Mathematics. *EURASIA Journal of Mathematics, Science & Technology Education*, 5(1), pp. 3-14

Costantino, T., Kellam, N., Cramond, B. and Crowder, I. (2010). An Interdisciplinary Design Studio: How Can Art and Engineering Collaborate to Increase Students' Creativity? *Art Education*, 63(2), pp. 49-53.

Danaia, L., McKinnon, D., Parker, Q., Fitzgerald, M. and Stenning, P. (2012). Space to Grow: LCOGT.net and Improving Science Engagement in Schools. *Astronomy Education Review*, 11 (1), pp.

Davis, N.T., Kumtepe, E.G. and Aydeniz, M. (2007). Fostering Continuous Improvement and Learning through Peer Assessment: Part of an Integral Model of Assessment. *Educational Assessment*, 12(2), pp. 113-135.

Desimone, L.M., Smith, T., Baker, D. and Ueno, K. (2005). Assessing Barriers to the Reform of U.S. Mathematics Instruction from an International Perspective. *American Educational Research Journal*, 42(3), pp. 501-535.

Ding, N. and Harskamp, E.G. (2011). Collaboration and Peer Tutoring in Chemistry Laboratory Education. *International Journal of Science Education*, 33(6), pp. 839-863.

Doherty, M., Fish, V. L. and Needles, M. (2011). Revealing the Hidden Wave: Using the Very Small Radio Telescope to Teach High School Physics. *Physics Teacher*, 49 (9), pp. 546-549.

Dori, Y. J. and Kaberman, Z. (2012). Assessing High School Chemistry Students' Modeling Sub-Skills in a Computerized Molecular Modeling Learning Environment. *Instructional Science: An International Journal of the Learning Sciences*, 40 (1), pp. 69-91.

Duncan, K.A., Johnson, C., McEllhinny, K., NG, S., Cadwell, K.D., Zenner Petersen, G.M., Johnson, A., Horoszewski, D., Gentry, K., Lisensky, G. and Crone, W.C. (2010). Art as an Avenue to Science Literacy: Teaching Nanotechnology through Stained Glass. *Journal of Chemical Education*, 87(10), pp. 1031-1038.

Erhardt, W. (2007). Instrumental Analysis in the High School Classroom: UV-Vis Spectroscopy. *Journal of Chemical Education*, 84 (6), p. 1024.

Foegen, A., Jiban, C. and Deno, S. (2007). Progress Monitoring Measures in Mathematics: A Review of the Literature. *Journal of Special Education*, 41(2), pp. 121-139.

Fortus, D., Dershimer, C.R., Krajcik, J., Marx, R.W. and Mamlok-Naaman, R. (2004). Design-Based Science and Student Learning. *Journal of Research in Science Teaching*, 41(10), pp. 1081-1110.

France, B., Compton, V.J. and Gilbert, J.K. (2011). Understanding Modelling in Technology and Science: The Potential of Stories from the Field. *International Journal of Technology and Design Education*, 21(3), pp. 381-394.

Fredriksson, U., Holzer, T., McCluskey-Cavin, H. and Taube, K. (2009). Strengths and Weaknesses in the Swedish and Swiss Education Systems: A Comparative Analysis Based on PISA Data. *European Educational Research Journal*, 8(1), pp. 54-68.

French, L. (2004). Science as the Center of a Coherent, Integrated Early Childhood Curriculum. *Early Childhood Research Quarterly*, 19(1), pp. 138-149.

Fuchs, L.S., Fuchs, D. and Courey, S.J. (2005). Curriculum-Based Measurement of Mathematics Competence: From Computation to Concepts and Applications to Real-Life Problem Solving. *Assessment for Effective Intervention*, 30(2), pp. 33-46.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. *Review of Educational Research, 82*(3), 300-329.

Geraedts, C., Boersma, K.T. and Eijkelhof, H.M. (2006). Towards Coherent Science and Technology Education. *Journal of Curriculum Studies*, 38(3), pp. 307-325.

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of educational research*, 75(1), 27-61.

Gioka, O. (2009). Teacher or Examiner? The Tensions between Formative and Summative Assessment in the Case of Science Coursework. *Research in Science Education*, 39(4), pp. 411-428.

Harlow, D. and Nilsen, K. (2011). Virtual Inquiry Experiences. *Science and Children*, 49 (4), pp. 42-45.

Hoachlander, G. and Yanofsky, D. (2011). Making STEM Real. *Educational Leadership*, 68(6), pp. 60-65.

Hodgen, J.,Marks, R. And Pepper, D. (2013). Towards universal participation in post-16 mathematics: lessons from high performing countries. London: Nuffield Foundation.

Hodgen, J.,Pepper, D., Kuchemann, D., and Coe, R. (2010). Is the UK an outlier? An international comparison of upper secondary education. London: Nuffield Foundation.

Hohenwarter, J., Hohenwarter, M. and Lavicza, Z. (2009). Introducing Dynamic Mathematics Software to Secondary School Teachers: The Case of GeoGebra. *Journal of Computers in Mathematics and Science*

Teaching, 28 (2), pp. 135-146.

Hume, A. and Coll, R.K. (2009). Assessment of Learning, for Learning, and as Learning: New Zealand Case Studies. *Assessment in Education: Principles, Policy & Practice*, 16(3), pp. 269-290.

Hurley, M. M. (2001). Reviewing integrated science and mathematics: The search for evidence and definitions from new perspectives. *School Science and Mathematics*, 101, pp. 259–268.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321*(5888), 494-495.

International Association for the Evaluation of Educational Achievement. (2011). *TIMSS 2011 Encyclopaedia*, Vol. 1 & 2. IEA.

Jones, G. M., Andre, T., Kubasko, D., Bokinsky, A., Tretter, T., Negishi, A., Taylor, R. and Superfine, R. (2004). Remote Atomic Force Microscopy of Microscopic Organisms: Technological Innovations for Hands-On Science with Middle and High School Students. *Science Education,* 88 (1), pp. 55-71.

Kalyuga, S. (2006). Rapid Cognitive Assessment of Learners' Knowledge Structures. *Learning and Instruction*, 16(1), pp. 1-11.

Kanter, D.E. (2010). Doing the Project and Learning the Content: Designing Project-Based Science Curricula for Meaningful Understanding. *Science Education*, 94(3), pp. 525-551.

Kennedy, K. and Schumacher, P. (2005). A Collaborative Project to Increase the Participation of Women and Minorities in Higher Level Mathematics Courses. *Journal of Education for Business*, 80(4), pp. 189-193.

Ketelhut, D.J. and Nelson, B.C. (2010). Designing for Real-World Scientific Inquiry in Virtual Environments. *Educational Research*, 52 (2), pp. 151-167.

Kidman, G. (2012). Australia at the Crossroads: A Review of School Science Practical Work. *EURASIA Journal of Mathematics, Science & Technology Education*, 8(1), pp. 35-47.

Kim, M.K. and Noh, S. (2010). Alternative Mathematics Assessment: A Case Study of the Development of Descriptive Problems for Elementary School in Korea. *EURASIA Journal of Mathematics, Science & Technology Education*, 6(3), pp. 173-186.

Kimbell, R. (2007) E-assessment in project e-scape. *Design and Technology Education: an International Journal,12* (2)

Kimbell, R., Wheeler, T., Miller, A., & Pollitt, A. (2007). *e‐scape: e‐solutions for creative assessment in portfolio environments.* Technology Education Research Unit, Goldsmiths College, London.

Kingston, N. and Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), pp. 28-37.

Klopfer, E., Scheintaub, H., Huang, W. Wendel, D. and Roque, R. (2009). The Simulation Cycle: Combining Games, Simulations, Engineering and Science Using StarLogo TNG. *E-Learning*, 6 (1), pp. 71-96.

Labate, H. (2007). Science Education: A (Pending) Chapter in the Curriculum Transformation in Argentina. *Prospects: Quarterly Review of Comparative Education*, 37(4), pp. 469-488.

Lee, S., & Roth, W. M. (2001). How Ditch and Drain Become a Healthy Creek Re-Presentations, Translations and Agency during the Re/Design of a Watershed. *Social Studies of Science*, *31*(3), 315-356.

Lehrer, J. (2011). *Proust was a neuroscientist.* Edinburgh: Canongate.

Lewis, T. (2006). Design and inquiry: Bases for an accommodation between science and technology education in the curriculum?. *Journal of Research in Science Teaching*, *43*(3), 255-281.

Lewis, C.C., Perry, R.R. and Friedkin, S. (2011). Using Japanese Curriculum Materials to Support Lesson Study outside Japan: Toward Coherent Curriculum. *Educational Studies in Japan: International Yearbook*, (6), pp. 5-19.

Li, J., Klahr, D. and Siler, S. (2006). What Lies beneath the Science Achievement Gap: The Challenges of Aligning Science Instruction with Standards and Tests. *Science Educator*, 15(1), pp. 1-12.

Lyon, G.H., Jafri, J. and St. Louis, K. (2012). Beyond the Pipeline: STEM Pathways for Youth Development. *Afterschool Matters*, (16), pp. 48-57.

Machin, S., & McNally, S. (2009). *The three Rs: What scope is there for literacy and numeracy policies to raise pupil achievement?* , London, UK.

Manouchehri, A. (2007). Inquiry-Discourse Mathematics Instruction. *Mathematics Teacher*, 101(4), pp. 290-300.

Marginson, S, Tytler, R, Freeman, B and Roberts, K (2013). *STEM: Country comparisons*. Report for the Australian Council of Learned Academies, www.acola.org.au last accessed 27.06.2013.

Marx, R.W., Blumenfeld, P.C., Krajcik, J.S., Fishman, B., Soloway, E., Geier, R. and Tal, R.T. (2004). Inquiry-Based Science in the Middle Grades: Assessment of Learning in Urban Systemic Reform. *Journal of Research in Science Teaching*, 41(10), pp. 1063-1080.

Meade, S. D., Caron, D. W., Gray, A. and Weaver, K. (2008). The ITEA-NASA STS-118 Design Challenges. *Technology Teacher*, 67 (4), pp. 16-21.

Meyer, H.-D., & Schiller, K. (2013). "Gauging the role of non-educational effects in large scale assessments: socio-economics, culture and PISA outcomes. In H.-D. Meyer and A. Benavot (Eds) *PISA, power and policy*. Oxford: Symposium.

Millar, R. (2011). Reviewing the National Curriculum for Science: Opportunities and Challenges. *Curriculum Journal*, 22 (2), pp. 167-185.

Munier, V. and Merle, H. (2009). Interdisciplinary Mathematics-Physics Approaches to Teaching the Concept of Angle in Elementary School. *International Journal of Science Education*, 31(14), pp. 1857-1895.

Nikitina, S. (2006). Three strategies for interdisciplinary teaching: contextualizing, conceptualizing, and problem- centring. *Journal of Curriculum Studies*, 38(3), pp. 251-271

Ontario Ministry of Education (2011). *Ontario schools, K-12 policy and program requirements*. OME.

Organisation for Economic Co-operation and Development. (2009). *PISA 2009 results: What makes a school successful: How schooling is organised*. OECD

Pampaka, M., Williams, J., Hutcheson, G. D., Wake, G., Black, L., Davis, P., & Hernandez-Martinez, P. (2012a). The association between mathematics pedagogy and learners' dispositions for university study. *British Educational Research Journal, 38*(3), 473-496.

Pampaka, M., Williams, J. & Hutcheson, G. (2012b). Measuring students' transition into University and its association with learning outcomes. *British Educational Research Journal*, *38*(6), 1041-1071.

Pedersen, S., Arslanyilmaz, A. and Williams, D. (2009). Teachers' Assessment-Related Local Adaptations of a Problem-Based Learning Module. *Educational Technology Research and Development*, 57(2), pp. 229-249.

Penuel, W., Fishman, B.J., Gallagher, L.P., Korbak, C. and Lopez-Prado, B. (2009). Is Alignment Enough? Investigating the Effects of State Policies and Professional Development on Science Curriculum Implementation. *Science Education*, 93(4), pp. 656-677.

Petre, M. and Price, B. (2004). Using Robotics to Motivate 'Back Door' Learning. *Education and Information Technologies*, 9 (2), pp. 147-158.

Rasinen, A., Virtanen, S., Endepohls-Ulpe, M., Ikonen, P., Ebach, J. and Stahl-Von Zabern, J. (2009). Technology Education for Children in Primary Schools in Finland and Germany: Different School Systems, Similar Problems and How to Overcome Them. *International Journal of Technology and Design Education*, 19(4), pp. 367-379.

Reiss, M. (2006). Enthusing the next Generation: A Biology Curriculum for 3-11 Year Olds. *Primary Science Review*, (93), pp. 23-25.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211-246.

Rickard, L.N. and Wilson, C. (2006). "Lettuce" Learn Math: Teaching Mathematics with Seeds and Centimeters. *Mathematics Teaching in the Middle School*, 11(8), pp. 368-377.

Riechert, S.E. and Post, B.K. (2010) From Skeletons to Bridges & Other STEM Enrichment Exercises for High School Biology. *American Biology Teacher*, 72(1), pp. 20-22.

Ritchie, M.R., Keith, D., Jones, D., Ramnarine, S. and Spiers, A. (2012). The Donaldson Review and Scottish Baccalaureates--A Model for Implementation in Schools. *School Science Review*, 93(345), pp. 99-108.

Romance, N.R. and Vitale, M.R. (2012). Expanding the Role of K-5 Science Instruction in Educational Reform: Implications of an Interdisciplinary Model for Integrating Science and Reading. *School Science and Mathematics*, 112(8), pp. 506-515.

Roth, K. and Garnier, H. (2007). What science teaching looks like: An international perspective. *Educational Leadership*, 64(4), pp. 16-23.

Ruggirello, R.,M., Balcerzak, P., May, V.L. and Blankenship, R.E. (2012). Measurement of Solar Spectra Relating to Photosynthesis and Solar Cells: An Inquiry Lab for Secondary Science. *Biochemistry and Molecular Biology Education*, 40 (4), pp. 241-245.

Ruiz-Primo, M. A., Li, M., Wills, K., Giamellaro, M., Lan, M. C., Mason, H., & Sands, D. (2012). Developing and evaluating instructionally sensitive assessments in science. *Journal of Research in Science Teaching*, *49*(6), 691-712.

Schaefers, J. (2006). Infrared Thermometers. *Science Scope*, 29 (4), pp. 19-21.

Schooler, S.R. (2004). A "Chilling" Project Integrating Mathematics, Science, and Technology. *Mathematics Teaching in the Middle School*, 10 (3), pp. 116-121.

Scott, C. (2012). An Investigation of Science, Technology, Engineering and Mathematics (STEM) Focused High Schools in the U.S. *Journal of STEM Education: Innovations and Research*, 13(5), pp. 30-39.

Sealey, P. and Noyes, A. (2010). On the "Relevance" of the Mathematics Curriculum to Young People. *Curriculum Journal*, 21(3), pp. 239-253.

Shayer, M. (1999). *GCSE 1999: Added-value from schools adopting the CASE Intervention*. London: Centre for the Advancement of Thinking.

Slade, P. (2009). Assessing Pupils' Progress in Science. *Education in Science*, (231), pp. 10-11.

Stein, M.K., Remillard, J. and Smith, M.S. (2007) How Curriculum Influences Student Learning. In F.K. Lester Jr. (Ed.) *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 319-369). Greenwich, CT: Information Age Publishing.

Stellflue, P., Allen, M. and Gerber, T.D. (2005). Art & Science Grow Together. *Science and Children*, 43(1), pp. 33-35.

Stinson, K., Harkness, S.S., Meyer, H. and Stallworth, J. (2009). Mathematics and Science Integration: Models and Characterizations. *School Science and Mathematics*, 109(3), pp. 153-161.

Suurtamm, C., Lawson, A. and Koch, M. (2008). The Challenge of Maintaining the Integrity of Reform Mathematics in Large-Scale Assessment. *Studies in Educational Evaluation*, 34(1), pp. 31-43.

Suzuki, J. (2009). The Alliance: Mathematics and the Arts. *Mathematics Teacher*, 103 (5), pp. 318-319.

Taber, K.S. (2010). Challenging Gifted Learners: General Principles for Science Educators; and Exemplification in the Context of Teaching Chemistry. *Science Education International*, 21(1), pp. 5-30.

Urban-Woldron, H. (2009). Interactive Simulations for the Effective Learning of Physics. *Journal of Computers in Mathematics and Science Teaching*, 28 (2), pp.163-176.

Van Langen, A., Bosker, R. and Dekkers, H.(2006). Exploring Cross-National Differences in Gender Gaps in Education. *Educational Research and Evaluation*, 12 (2), pp. 155-177.

Venville, G., Rennie, L.J. and Wallace, J. (2012) Curriculum Integration: Challenging the Assumption of School Science as Powerful Knowledge. in B.J. Fraser et al. (eds.), Second International Handbook of Science Education, Springer International Handbooks of Education 24.

Villegas, J.C., Morrison, C.T., Gerst, K.L., Beal, C.R., Espeleta, J.E. and Adamson, M. (2010). Impact of an Ecohydrology Classroom Activity on Middle School Students' Understanding of Evapotranspiration. *Journal of Natural Resources and Life Sciences Education*, 39, pp. 150-156.

Wachira, P. and Keengwe, J. (2011). Technology Integration Barriers: Urban School Mathematics Teachers Perspectives. *Journal of Science Education and Technology*, 20 (1),pp. 17-25.

Wallace, J., Sheffield, R., Rennie, L. and Venville, G. (2007). Looking Back, Looking Forward: Re-Searching the Conditions for Curriculum Integration in the Middle Years of Schooling. *Australian Educational Researcher*, 34(2), pp. 29-49.

Weyman, P.D. (2009). The Interdisciplinary Study of Biofuels. *Science Teacher*, 76 (2), pp. 29-34.

Wilder, A. and Brinkerhoff, J. (2007). Supporting Representational Competence in High School Biology with Computer-Based Biomolecular Visualizations. *Journal of Computers in Mathematics and Science Teaching*, 26 (1), pp. 5-26.

Wiliam, D. (2007). Keeping learning on track: Formative assessment and the regulation of learning. In F. K. Lester, Jr. (Ed.) *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Greenwich, CT: Information Age Publishing.

Williams, J.P. (2011). STEM Education: Proceed with Caution. *Design and Technology Education*, 16 (1), pp. 26-35.

Williams, J.P. (2010) Technology Education to Engineering: A Good Move?. *Journal of Technology Studies*, 36 (2), pp.10-19

Zhang, D. and Campbell, T. (2012). An Exploration of the Potential Impact of the Integrated Experiential Learning Curriculum in Beijing, China. *International Journal of Science Education*, 34 (7), pp. 1093-1123.

Zoller, U. (2011). Science and Technology Education in the STES Context in Primary Schools: What Should It Take? *Journal of Science Education and Technology*, 20 (5), pp. 444-453.

Appendices

Appendix A: International comparisons of curriculum, assessment, interdisciplinarity and integration

Appendix B: Case studies

B1   Policy literature from 20+ countries: a summary of the ACOLA report
B2   Innovating in mathematics curriculum and assessment and the wider political context: a case study of two contrasting states in Australia
B3   Primary Lessons from Japan: towards a high quality curriculum (intended and enacted).
B4   Modelling and problem solving in STEM (with Mathematics)
B5   Innovations in assessment:
      B5a   E-portfolios and comparative judgment techniques
      B5b   Measurement of dispositions/affect and learners' trajectories
B6   Case study of NLT integrated science and maths subject, Netherlands
B7   Enrichment strategies in science and mathematics

Appendix C   Methodology: literature search process

Appendix A: International comparisons of Curriculum, assessment, interdisciplinarity and integration

David Swanson

What is the state of play (in curriculum, assessment, interdisciplinarity and integration in STEM age 5 to 18) and the evidence base for effective system practices across the UK and internationally?

We answer these questions through analysis of the educational systems of the 4 UK nations and 6 other states/jurisdictions which are relatively high performing in terms of international comparisons (i.e. TIMMS and PISA). Reference to such measures should not however be taken as acceptance of the performativity agenda associated with them, nor indeed, that these tests show what they purport to, i.e. a valid indication of the relative attainment of the populations they refer to. The non-UK systems analysed include Japan, Singapore, the Netherlands, Finland, Ontario (Canada) and California (U. S.).

Curriculum

What do compulsory science and mathematics education pathways to age 18 look like in other countries?

See the attached table for an overview

Are there discernible patterns?

All systems in our sample have some form of common national curriculum for the early stages of education however these vary in their level of prescribed detail. For example, the Dutch have eleven 'Core Objectives' for mathematics in primary schools, which can be listed on one page, whereas in California there are over 100 pages of 'Mathematics Content Standards' and 'Grade-Level Considerations' for the equivalent period. Some nations also prescribe both the instructional methods and the curriculum materials used (Japan, Singapore). This period of common curriculum, also varies in extent. At one end we have Singapore where distinct pathways appear at age 12, while at the other, we have Finland (age 16) but also Japan (age 15).

In the various systems, at upper secondary level, we essentially find three strands: vocational, higher vocational/general academic and higher academic. The systems vary between those which institutionalise these strands very clearly (e.g. the Netherlands) and those which keep it all under one roof (e.g. Californian High Schools).These latter perhaps have greater room for flexibility and allow more differentiation in pathways. The systems also have minor variations in their distribution of compulsory and optional elements within institutions. All have some compulsory mathematics and science courses in this phase, although sometimes with different levels available.

Science options, at this level, are usually physics, chemistry and biology, or a 'general science' (particularly in the early stages of upper secondary) which is essentially subdivided similarly. General institutions are more likely to offer a wider variety than this (such as those in Ontario which include options in Earth and Space Science, Environmental Science and others, through Interdisciplinary Studies). Vocational and upper vocational strands may also include wider disciplines, although often as applied science.

Almost all students have compulsory maths and science, in some form, to at least 15. Some vocational pathways and institutions (for example, in Singapore, around 15% of students) do not include compulsory science in latter stages of upper secondary.

What determines these patterns?

If we put aside its role in babysitting children, the main purpose of mass schooling is to prepare students for the workplace and, to some extent, wider society. One central aspect of this is the allocation of positions within a stratified labour force. This is the main determinant of the general stratification we see within education systems. 40 years of neo-liberalism have led to both an increase in pressure for

competitive individualism, yet at the same time a massive increase in bureaucracy to push through this performativity agenda. Nation states tend to see their global economic competitiveness as dependant on developing markets linked to technologies. Overall this leads to an emphasis on advanced STEM training for a minority and the measurement of success based on exam results (whether at individual, school, national or international level).

Some states have, for various historical, political and cultural reasons greater emphasis on societal features, such as equality, which schooling is also encouraged to develop. The interplay between such factors and economic demands shape, in a complex way, various issues such as school and teacher autonomy or at what age stratified education is introduced.

What are the post-compulsory pathways that allow young people to study STEM subjects to age 18?

See the attached table

Non-compulsory pathways follow the same pattern as the compulsory pathways described. However, separate pathways are more likely to take place in separate institutions.

At what age does compulsory science and mathematics begin in other countries?

See attached table. Maths and science generally begin in the first year of schooling, typically year 6 in our sample. Many nations have large-scale participation in kindergarten prior to this age, some with nationally defined curricula including maths and science (e.g. Ontario from age 4). Some nations, such as Singapore and Japan, only begin science in grade 3. However, the differences in activities undertaken, say at age 6, may be minimal in practice.

To what extent does practical science form a part of the formal science curriculum in the UK and other countries?

The table below shows the approximate percentage of students whose teachers use science equipment and materials as a basis for instruction (or, in brackets, as a supplement to instruction), as reported by teachers in the TIMMS 2011 Science survey for grades 4 and 8. The equivalent figures for textbooks, workbooks/worksheets and computer software are also given for comparison. Alongside the inherent problems in self-reporting, it should also borne in mind that these percentages do not necessarily imply any student engagement with the materials or equipment.

Country/State	Science equipment and materials		Textbooks		Workbooks/ Worksheets		Computer software	
	G4	G8	G4	G8	G4	G8	G4	G8
California	-	36 (64)	-	53 (46)	-	22 (73)	-	18 (65)
England	62 (38)	62 (37)	4 (45)	8 (78)	4 (82)	21 (76)	15 (74)	29 (67)
Finland	7 (90)	38 (62)	94 (6)	78 (22)	40 (54)	26 (67)	1 (61)	3 (64)
Japan	62 (38)	65 (35)	82 (17)	71 (29)	17 (76)	30 (69)	2 (52)	3 (49)
Netherlands	4 (78)	-	74 (13)	-	72 (26)	-	3 (31)	-
Ontario	36 (61)	34 (66)	33 (54)	54 (44)	28 (69)	15 (78)	7 (51)	5 (69)
Singapore	60 (40)	32 (67)	68 (27)	52 (39)	69 (31)	66 (34)	19 (78)	23 (69)

The next table, taken from the same source, shows the approximate percentage of students where science investigation is emphasised in their classroom **about half of the class time or more**. This includes

students watching, designing or conducting experiments, observing natural phenomena or relating science to their daily lives. Again these figures are self-reported by teachers in the survey.

% of students in	Japan	Singapore	U.S.	England	Ontario	Finland	Netherlands
Grade 4	51	50	41	41	32	13	5
Grade 8	32	29	47	37	22	32	-

Although both sets of figures give some indication of the extent of practical science, the lack of differentiation between teacher activity, teacher-led student activity, and student- led practical activity undermines the possible conclusions which may be drawn from this sort of data. Similarly, the widely seen emphasis within formal curricula on investigation and experimentation suffers from similar evidential defects. Further discussion on this topic can be seen in the main report.

Assessment

How do the UK nations and other countries assess STEM learning at a national level?

See the attached table for an overview of non-UK systems.

There are clear differences between those systems which rely on high-stakes examinations for assessment of individual students and those which emphasise teacher judgement (in relation to criteria). For example, Finland has no high stakes testing until the matriculation exam at the end of upper secondary, and, even at that stage, Ontario assessment only relies 30% on (teacher marked) test scores. Such an approach seems to better allow assessment in relation to problem-solving and project based learning. Another noticeable difference is Japan's reliance on entrance (as opposed to exit) exams for progression. Combined with their multiplicity this appears to temper pressures to 'teach to the test'.

Apart from the Netherlands, all of the states in our sample engage in regional and/or national standardised testing in order to test school and system-wide capabilities, although at varying frequencies (see attached table).These are not always distinct from examinations with implications for individual progress (for example, the Primary Leaving Exam in Singapore). They also participate in international sampled standardised testing (PISA and TIMMS).

A further reflection of the degree of trust in schools and teachers is seen in the lack of inspection visits to assess curriculum implementation in Finland, Ontario and Singapore. Finally, another tool for assessing STEM education, an official research program, is only absent in the UK and Singapore.

Integration of STEM subjects

To what extent is teaching and learning about STEM subjects integrated within the teaching of the arts and humanities (e.g. history, English, modern foreign languages, creative and performing arts) across the UK and more widely internationally?

There is little evidence of such integration, either within formal curricula or in practice, outside of isolated, small projects. Primary curricula sometimes suggest ways that science can be integrated within wider activities (for example, in Ontario's primary curriculum). Some attention to numeracy or mathematics as cross-curricular does appear in most written curricula, or in guidance. Aspects of or applications of mathematics and science do occasionally appear (briefly) within arts and humanities courses (for example dance science, acoustics for music students, or as aspects of cultural and philosophical history). This issue is discussed in more detail in the main report.

Subjects loosely categorised as Design and Technology are more likely to include other STEM elements to the extent that they are necessary, but science and maths are not explicitly taught through this type of course, except perhaps in Finland and the Netherlands where integration is more likely to occur.

Is there evidence that such integrated approaches benefit pupils, and how reliable is this evidence?

See the main study for a discussion of this question.

Interdisciplinarity

To what extent are STEM subjects taught in an interdisciplinary manner across the UK and in other countries?

Thematic education is more common at Primary level which allows scope for interdisciplinary STEM in most systems. Systems with extended general education, such as Finland and now Scotland, can continue such practices much later into school life, because the curriculum and assessment regime impose less of a disciplinary focus.

There is little evidence of interdisciplinary STEM in the curricula of nations which rely on individual subject examinations (such as England and Singapore), even where Science is taught as one subject this tends to involve physics, chemistry and biology being taught serially.

Interdisciplinarity is most advanced at secondary level in matriculation/leaving certificate type systems which allow room for over-arching connecting modules such as in Ontario's interdisciplinary studies, or California's Career Technical Education. In Ontario, this extends to other forms of Integrated Curricular Programs, which thematise groups of varied modules, sometimes to the stage of integration.

A final aspect of interdisciplinarity is in the role of national science centres, such as in Singapore or the National STEM Centre here in England. These often work with schools on developing interdisciplinary projects or provide out-of-school activities which do like-wise.

What evidence is there that interdisciplinary teaching and learning of STEM benefited pupils?

See the main study for a discussion of this question.

Bibliography:

Australian Council of Learned Academies. (2011). Consultancy Report. Securing Australia's future STEM. Country Comparisons: Japan. ACOLA.

California Department of Education. (2004). Science framework for California public schools, K-12. CDE.

California Department of Education. (2006). Mathematics framework for California public schools, K-12. CDE.

Eurydice. (2006). Science Teaching in Schools in Europe. Brussels: Eurydice.

Eurydice. (2011a). Mathematics Education in Europe: Common Challenges and National Policies. Brussels:Education, Audiovisual and Culture Executive Agency, European Commission.

Eurydice. (2011b). Science Education in Europe: National Policies, Practice and Research. Brussels:Education, Audiovisual and Culture Executive Agency, European Commission.

Grice, J. (2011). Curriculum Integration in Ontario High Schools. *Pathways: The Ontario Journal of Outdoor Education*, *24*(1), 4-8.

Hodgen, J., Pepper, D., Sturman, L., Ruddock, G.(Nuffield Foundation) (2010). An international comparison of upper secondary mathematics education.

International Association for the Evaluation of Educational Achievement. (2011a). TIMSS 2011 encyclopedia, Vol. 1. IEA.

International Association for the Evaluation of Educational Achievement. (2011b). TIMSS 2011 encyclopedia, Vol. 2. IEA.

Nezvalova, D. (2010). Science Education in Europe. European dimension in science education, 11.

Ofqual. (2012). International comparisons in senior secondary assessment. Full report. Ofqual

Ofsted (2010). Finnish Pupils' Success in Mathematics. Factors that Contribute to Finnish Pupils' Success in Mathematics. Manchester: The Office for Standards in Education, Children's Services and Skills.

Ontario Ministry of Education. (2002). The Ontario Curriculum Grades 11 and 12, Interdisciplinary studies. OME.

Ontario Ministry of Education. (2007). The Ontario Curriculum Grades 11 and 12, Mathematics. OME.

Ontario Ministry of Education. (2008). The Ontario Curriculum Grades 11 and 12, Science. OME.

Ontario Ministry of Education. (2011). Ontario schools, K-12 policy and program requirements. OME.

Organisation for Economic Co-operation and Development. (2007). PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis. OECD Publishing.

Organisation for Economic Co-operation and Development. (2009). PISA 2009 results: What makes a school successful: How schooling is organised. OECD

Osborne, J., & Dillon, J. (2008). Science education in Europe: Critical reflections: a report to the Nuffield Foundation. Nuffield Foundation.

Queensland Department of Education, Traning and Employment. (2013). P-12 Curriculum, assessment and reporting framework.

Science Community Representing Education. (2008). Practical work in science: A report and proposal for a strategic framework. London: SCORE.

ScienceMath group (2009). The ScienceMath project: Results and material. (Available online at http://www.sciencemath.ph-gmuend.de/Download/Sciencemathvolume.pdf)

Websites:
Eurydice: http://eacea.ec.europa.eu/education/eurydice/
OECD: http://www.oecd.org/education/
TIMMS: http://nces.ed.gov/Timss/
California: http://www.cde.ca.gov/index.asp
Finland: http://www.minedu.fi/OPM/Koulutus/?lang=en
Japan: http://www.mext.go.jp/english/
Netherlands: http://www.government.nl/issues/education
Northern Ireland: http://www.nicurriculum.org.uk
Ontario: http://www.edu.gov.on.ca/eng/
Scotland: http://www.educationscotland.gov.uk and http://www.scotland.gov.uk/Topics/Education
Singapore: http://www.moe.gov.sg/
UK: http://www.education.gov.uk
Wales: http://learning.wales.gov.uk/?skip=1&lang=en and http://wales.gov.uk/topics/educationandskills/?lang=en

Country/ Region	Curriculum								Assessment
	Age compulsory begins				Pathways to 18				
	Maths		Science			Grades	Pathways	Science/Maths	
	Grade	Age	Grade	Age	Compulsory?				
Ontario, Canada	1[1]	6	1	6	C	1-8	Common national curriculum		Assessment at all levels primarily relies on teacher judgement. Secondary modules = 70% criteria based continuous assessment/observation. 30% final test. General standardised testing: Pan-Canadian, rolling program focused on one of maths/science/reading every 3 years. Ontario standardised testing of maths in grades 3, 6 and 9.
					C	9-10	Applied, Academic, Open (mixed)	Secondary school diploma: Credit based- Minimum 30. 18 compulsory including 3 Maths (at least 1 in grade 11/12) and 2 Science. Secondary school certificate: Minimum 14 credits (7 compulsory including 1 Maths, 1 Science) Different pathways have different courses.	
					C	11-12	Uni., College, Uni/College, Workplace, Open		
Singapore	1	6	3	8	C	1-6	Common national curriculum		Assessment is primarily exam based. Primary school leaving exam in 4 subjects including science and maths. Secondary: O-level or N-level Pre-University: A-level or IB
						7-10/11[2]	Express, Normal-academic, Normal-vocational	Science and Maths compulsory to grades 9-10 for majority. Science compulsory to grade 8 in normal vocational.	
						11-	Pre-university, vocational, polytechnic		
Netherlands	1	6	1	6	C	1-8	National curriculum of core objectives		Primary- most schools use national multiple choice tests combined with teacher recommendation for pathway. At end of secondary (grade 10/11 or12) national exams dependant on pathway
					C[3]	9-	Pre-vocational (2yrs) → Voc. Sec. or Snr Gen. Senior General (3yrs) → Pre-Uni or Higher Voc. Pre-University (4yrs) → Uni.	In latter stages of Senior General and Pre-University pathways there are 4 possible strands with fixed study programs. 2 are science related and 3 contain mathematics.	
Japan	1	6	3	8	C	1-9	Compulsory curriculum		Primary and lower secondary use teacher criteria based assessment but entrance exams for upper secondary. Entrance exams for university.
						10-12[2]	General secondary (academic) or Specialised secondary (vocational/career). General secondary usually has two tracks with emphasis on science or humanities.	Maths compulsory to grade 10 Varied maths and science options in grades 11, 12	
Finland	1	7	1	7	C	1-9	Basic education - National curriculum		Basic education teacher assessed mix of tests and continuous assessment/ observation. National descriptions of
						10-12	Upper Secondary General or Vocational	Within upper secondary some	

[1] 90% of children attend kindergarten from age 4 with a state guided 'Core day program' which includes play-based learning in maths and science.

[2] Secondary school is non-compulsory but near universal

[3] Partial-compulsory at age 16/17 (minimum two days per week)

									science courses are compulsory. Matriculation exam: Of four tests one mathematics (at two possible levels) and one which includes optional science questions.	good performance at grades 2, 5 and 9 but no grade descriptors. Individual courses at upper secondary level are similarly teacher assessed. General upper secondary matriculation exam. University entrance exams though. General standardised assessment: Sample based national assessments in science and maths at grade 9 and maths at grade 6 every few years. (Science much less frequently).
California, U.S.	1[4]	6	1	6	C	1-8	Common core standards		Standardised tests including maths and science every year, grades 2-11. Teacher assessment for individual courses primarily test based. High school exit exam including maths required for high school diploma. Commercial standardised tests inc. 'SATS' play a role in university entry. General standardised assessment: See grade 2-11 tests above. Plus national periodical sampled testing in maths and science.
					C	9-12	High School	Minimum requirements of high school diploma include successful completion of two year-long courses in maths and two in science in grades 9-12	

Assessment in the UK:

In England, Wales and N. Ireland student assessment is exam-based (GCSE/ A-levels at age 15/17) and currently moving away from any elements of coursework/teacher assessment. In Scotland, the equivalents (Standard/Higher/Advanced Higher at age 15/16/17) are being phased out under the new Curriculum for Excellence. National 4 and 5, replacing Standard, are either fully or partly class work based, and the upcoming new Higher and Standard Higher are expected to move in the same direction.

England: SATS in Maths and Science at age 7, 11 and 14. These involve teacher assessments at age 7 and 14 but national tests at age 11.
Northern Ireland: Teacher assessed 'cross-curricular' Use of Mathematics at age 8, 11 and 14. Wales: Teacher assessments for Maths and Science at age 7, 11 and 14.
Scotland: Provides a National Assessment Resource. School's self- report (moderated) progress.

---

[4] Science and Maths content standards exist for kindergarten from age 5, which most attend.

Country/Region	Curriculum: Extent of practical Science	STEM integration D&T	Interdisciplinary STEM
Ontario, Canada	Extensive within curriculum	Wide variety of technology courses available at secondary level.	In grades 11/12 options include Interdisciplinary Studies e.g. Biotechnology, Learning and Mathematics. These include mix of other modules plus a unifying project-type module.
Singapore	Schools well-equipped with labs	Post-secondary vocational pathways may begin pre-18 and include engineering at Institute of Technical Education or diploma at polytechnic. Design and Technology an option at secondary.	Singapore Science Centre and A*STAR
Netherlands	Most secondary schools, but not primaries, have science labs	'Technology' in lower secondary. Integrated with science in upper secondary.	Advanced science, mathematics and technology an optional interdisciplinary subject in Senior General or Pre-University pathways
Japan	Curriculum emphasis on experiment and observation	Various specialised D&T, Engineering and other science courses are possible at upper secondary level at applied or theoretical levels.	At primary and lower and upper secondary, all schools have space for school-determined 'Integrated studies'. These are, effectively, (class) projects which should interdisciplinary work across all subjects including STEM
Finland	One appropriately equipped science lab per 27 hrs of science lessons is prescribed	'Technology Education' problem-centred design tasks during Basic education.	LUMA project- National Science Education Centre
California, U.S.	Investigation and experimentation included in standards at all grade levels	See Career Technical Education under interdisciplinarity.	Career Technical Education in high schools. For example students can do a Biotechnology research and development course with science and maths modules and CTE courses which integrate these.
England Wales N. Ireland	80%of teachers indicate that they spent more than 40% of lesson time at KS3 doing practical work, though only 56% and 45% report that they spent more than 40% of time at KS4 and KS5 respectively (Scope, 2008, p. 2). 'Well-documented evidence of the shortcomings of equipment funding and replacement of laboratories' (p. 16).	Design and Technology - part of the national curriculum for KS1-3 Optional subject at GCSE and A-level	National STEM Centre
Scotland		Technologies – part of the new Curriculum for excellence at general education stage. National 4 to Advanced Higher qualifications available in Design & Technology, Design & Manufacturing, Engineering Science from grade 10	Interdisciplinarity encouraged at all levels, and general education extended to age 14. Planned redesign of qualifications includes greater coursework and scope for projects on subjects of interest.

Appendix B

B1   Policy literature from 20+ countries: a summary of the ACOLA report

B2   Innovating in mathematics curriculum and assessment and the wider political context: a case study of two contrasting states in Australia

B3   Primary Lessons from Japan: towards a high quality curriculum (intended and enacted).

B4   Modelling and problem solving in STEM (with Mathematics)

B5   Innovations in assessment:

      B5a   E-portfolios and comparative judgment techniques

      B5b   Measurement of dispositions/affect and learners' trajectories

B6   Case study of NLT integrated science and maths subject, Netherlands

B7   Enrichment strategies in science and mathematics

Appendix B1 Policy literature from 20+ countries: a summary of the ACOLA report

This case study is still under construction. It began with a conversation with the second author of the report (Tytler) and a summary is still under consideration.

We drew on the ACOLA "STEM: country comparisons" and their reports from 28 international consultants' published in Marginson, S, Tytler, R, Freeman, B and Roberts, K (2013). *STEM: Country comparisons*. Report for the Australian Council of Learned Academies, www.acola.org.au last accessed 27.06.2013. This report was produced as a synthesis from the reports of consultants and not from the literature per se, but is of course informed by a great deal of expertise.

The principal issues for English policy in curriculum and assessment in STEM appeared in this report to be:

(i)   Cultural-political factors: 'trust in professionalism' versus 'accountability and league tables' (the consultants report from UK and others);

(ii)   The apparent conflict between 'elite' science/STEM and 'science/STEM for all' is not seen as a real difficulty, in practice these go together. (But maybe there are some issues with shortages of teacher supply that could bring conflict – just as schools prioritise certain years (e.g. in England years 6, 9 and 11);

(iii)   The apparent tension between 'discipline-content focused' and the integrated/interdisciplinary innovation/D&T and problem-focussed problem solving;

(iv)   The apparent tension between vocational (tech/eng) and academic STEM (science/D&T): some countries invest significantly in a strong vocational sector in schools and this seems to help them get STEM 'coverage' – but in Germany for example, mathematics is compulsory in these vocational courses. This is probably important for England too.

(v)   The importance of attitudes and dispositions to STEM, supported by innovative curricula etc, but also the dangers of 'yutori'.

(vi)   The section on gender is powerful in making the point that countries that have improved (claim they) have made great efforts to do so;

(vii)   Technological innovation in learning and assessment is a significant factor;

(viii) A STEM community and professional support is necessary for STEM improvement in schools: we know that the cultural capital of a school makes the difference.

Some key passages/ quotes from the report follow:

"Countries strong in STEM

While the countries strong in STEM are diverse in their economies, political and social cultures and their educational traditions, certain features recur in common.

 First, school teachers enjoy high esteem, are better paid and work within more meritocratic career structures than found elsewhere. An outstanding example is Finland, where all teachers have a Masters degree, teaching is harder to enter than most other professions, and the strongest teachers are paid to work in school districts serving poor families and students with the most learning difficulties. In China, STEM teachers receive salary increases not on the basis of seniority but via continuing professional development programs, specific to the discipline. To be promoted China's teachers must demonstrate an improving standard of work.

Second, these countries have an unbreakable commitment to disciplinary contents. They do not equate teaching with class management and credentialing alone. They focus on knowledge. STEM teachers are expected to be fully qualified in their discipline and to teach in that field and not others. This contrasts sharply with Australia. Professional development is primarily focused on the discipline rather than generic programs, which again contrasts with Australia.

Third, the most successful countries have instituted active programs of reform in curriculum and pedagogy that are focused on making science and mathematics more engaging and practical, through problem-based and inquiry-based learning, and emphases on creativity and critical thinking. These themes also run through the best Australian classrooms in STEM. The main South Korean program for building participation and achievement in STEM has incorporated the arts, to strengthen the focuses on creativity and design. The program is titled STEAM. These more student-centred approaches are being employed without diluting content. In Japan, where mandatory hours and standards in STEM were successively lowered for two decades and PISA performance declined, since 2008 there has been a return to stronger content requirements and less open choice.

Fourth, many of these countries have developed innovative policies to lift STEM participation among formerly excluded groups. Finland's focus on low achieving students has been mentioned. The consultants' report on the strategies used among first nation students in the Province of Saskatchewan suggests that the Canadian experience has lessons for indigenous STEM education in Australia. Finally, STEM-strong countries have developed strategic national STEM policy frameworks which provide favourable conditions for a range of activities: centrally driven and funded programs, including curriculum reform and new teaching standards; world class university programs, the recruitment of foreign science talent and new doctoral cohorts; decentralised program initiatives and partnerships and engagement that link STEM activities in schools, vocational and higher education with industry, business and the professions. Often STEM programs are led or facilitated or informed by institutes, centres or other agencies that have been specifically created to progress and resource the shared national STEM agenda."

And from the key findings for curriculum and assessment, this:

"Key Finding 5.1: Broadening STEM engagement and achievement

In all strong STEM comparator countries, broadening STEM engagement and achievement entails improving participation in the STEM disciplines through 'T' policies (i.e. learning in both breadth and depth) and covering the full spectrum of prior student achievement levels. In particular:

- Provision of at least some discipline based STEM learning for all school students, up to and including students in senior secondary education.
- Improving the engagement and performance of students from groups currently underrepresented in STEM, that on average perform relatively poorly in mathematics and science.
- Lifting the size and average achievements of the group of students engaged in intensive STEM learning in depth, in both schooling and higher education.

Key finding 5.2: STEM-specific tracking in secondary education

Many of Australia's comparator countries achieve strong participation in STEM through bifurcation at secondary school level between STEM and non-STEM tracks, and vocational tracks leading to significant STEM training. There may be benefits in significant discussion in Australia concerning the potential for, and the pros and cons of:

- Firm bifurcation between a comprehensive STEM track, and a non-STEM track, in the final two years of secondary education.
- Development of STEM-heavy technical and vocational schools and tertiary institutes, alongside academic secondary schools and universities (the latter also including some STEM). ...

Key Finding 5.4: STEM-specific prerequisites for higher education

In a number of high performing countries STEM subjects at upper secondary school level are strongly linked to university entrance. One way of lifting the level of study of STEM in both senior secondary and higher education would be the reintroduction of more comprehensive prerequisite requirements for university programs requiring advanced STEM knowledge, optimising preparation in the disciplines. ...

Attitudes to STEM

Key finding 6.1: Building awareness of STEM disciplines and STEM-related occupations among young people

For most countries, initiatives targeted at student attitudes and identity were a significant part of the strategic mix. This included initiatives to increase awareness of the nature of STEM professions. Based on the consultants' reports, strategies and programs could be further developed and extended so as to encourage in students positive attitudes to study of mathematics and science, and to STEM-related work and careers. Such strategies would need to take into account the diversity of students' contexts, including their gender, ethnicity/ cultural background, SES status and indigeneity.

Such strategies could include:

- Awareness campaigns to enrich public understanding of career options in STEM and the nature of STEM work, and to alert young people to the range of possible future STEM lives and identities. • Strategies at school level designed to involve families in mathematics and science learning and in building positive attitudes to STEM related careers.
- Role models, in the form of student interaction with practicing STEM professionals, or web-based presentations of narratives of STEM professionals (such as those on the Academy of Technological Sciences and Engineering [ATSE] Science and Technology Education Leveraging Relevance [STELR] website).
- Career advice that includes images of people working in STEM-related careers, delivered through information workshops for careers teachers, and mathematics and science teachers.
- The inclusion, in curriculum resources, of images of people working in STEM-related careers.
- The inclusion, in curriculum resources, of materials that speak to the identity needs of the diverse range of students. This includes girls (e.g. science material related to health, or the environment.), indigenous students (e.g. materials that embody respect for indigenous knowledge), and contextual science that relates to youth interests.
- The expansion of opportunities for families and the general public to engage positively with science and mathematics through events, exhibitions and other approaches.
- Enrichment programs whereby students are engaged in science or mathematics projects that entail linking to members of local communities. ...

Framing national STEM policy and strategy

Key finding 7.1: National STEM policy

A number of countries articulate through national policy a government commitment to STEM or a broader science and technology agenda. In these cases national policy establishes a framework for STEM-specific objectives and facilitates the implementation of coherent STEM specific strategies and programs. National STEM policy tends to span more than one government ministry, and in many instances is supported by structures coordinating STEM or science and technology activity across jurisdictions and agencies. National STEM or science and technology policy is generally conceived in human capital terms. A national STEM policy could provide a coherent framework for identifying and articulating STEM specific strategies and programs spanning the school, VET, higher education and research and development sectors, and also relevant programs in relation to innovation, employment and industry development.

School curriculum and pedagogy

Key finding 8.1: Inquiry, reasoning, and creativity and design in STEM curricula

Many comparator countries with strong STEM agendas and results have a well-developed curriculum focus on innovation, creativity and reasoning, accompanied by a strong commitment to disciplinary knowledge. In relation to school curricula, teaching, learning and educational policy and organisation could usefully address elements such as:

- Strong disciplinary frameworks, noting that disciplinary thinking and disciplinary literacies are central to creative problem solving in STEM-related learning and work.
- At the core of learning, methods of problem solving, inquiry, critical thinking and creativity, all of which can enhance both students' attitudes to, and practical competencies, in STEM fields.
- Design tasks into school science and mathematics curricula, in order to support the development in students of problem solving skills, flexibility in thinking, and awareness of engineering design activities.
- Consideration of the inclusion of the visual and performing arts alongside strategies and programs designed to enhance the orthodox STEM-related disciplines, as in the successful STEAM policy in Korea.
- Development of assessment regimes that support the commitment to problem solving, inquiry-based approaches, critical thinking and creativity. "

A number of other passages form the report were also found useful: e.g. the passages on girls and gender that also informed appendix B8.

B2   Innovating in mathematics curriculum and assessment and the wider political context: a case study of two contrasting states in Australia

Merrilyn Goos (University of Queensland) and Julian Williams

Introduction

In England the recent political agenda has been dominated by standards of a particularly specific and narrow kind, based on test results and league tables of schools and teachers performance: and these have been imposed by a performative management system within schools that have left little room for challenge. The political intention (under first the Thatcher government and later Blair and Brown governments) was to 'drive up standards' and challenge low expectations of learners; the benchmark was test outcomes that left little room for negotiability. In many countries of the world now this is becoming a recognisable pattern, and TIMMS and PISA are becoming used as technologies for this purpose.

One of the side effects of this culture was the widespread abandonment of authentic assessment of problem solving projects in Science and Mathematics as a serious possibility within these high stakes testing cultures, both in the UK and in some other parts of the world. Initially some niches were left in the margins, with a few examinations incorporating modelling or project work in coursework assessments, these became increasingly formulaic and lacking in room for individual creativity. At the same time, increasingly teachers themselves have started to form anti - coursework sentiments: assessing coursework in the UK became seen as another burden.

Thus, even though some recent research has shown the potential value of modelling coursework, the prognosis for problem solving and modelling may be poor (see Hernandez-Martinez et al., 2011). On the other hand developments in Australia were both initially stronger and have been in some places – but not all - sustained.

The question posed, then, is how different jurisdictions, even in the same country, can afford different political contexts in which innovations might flourish or decay?

The comparison between Queensland and Victoria

Australia provides an interesting case study of attempts in different states to embed mathematical modelling into the secondary school curriculum. Until recently, each state had its own education system and curriculum and assessment policies, and this led to significantly different experiences in embedding a modelling approach into mathematics curricula, particularly in the upper secondary school where subject choices and high stakes assessment determine which students gain entry to university.[5] The comparison between the states of Victoria and Queensland is instructive in identifying influences on sustained implementation of a modelling approach. This analysis draws on work undertaken by Stillman (2004, 2007).

In the 1980s a range of innovative projects and curriculum developments around Australia were emphasising an applications and modelling approach. These included the *Mathematics Curriculum and Teaching Program* (Lovitt, Clarke, & Stephens, 1986), designed for teachers in lower secondary and primary classrooms, and the Australian Academy of Science's series on *Mathematics at Work: Modelling your World* (Lowe, 1988, 1991), which provided modelling scenarios at the upper secondary level. There were also modelling courses designed and taught by enthusiasts and implemented within a standard, non-modelling, curriculum framework (e.g., Galbraith and Clatworthy, 1990). However, none of these lighthouse initiatives resulted in system-wide curriculum change in the high stakes subjects at senior secondary level that determine entry to university.

---

[5] A national curriculum with common agreed content has now been developed for the primary and secondary years of schooling, but implementation in the senior secondary years is not expected until 2016. Because states will be allowed to add content and de-emphasise some of the common content it is possible that there will still be a degree of diversity in senior secondary mathematics curricula in the years to come.

By the early 1990s, every state in Australia was experiencing impetus for curriculum reform. This was due to several factors, such as increasing retention rates leading to a more diverse student cohort in the upper secondary years, and a desire to make the subject matter and learning approaches more relevant, challenging and meaningful for all students, not just those aspiring to university entry. **In the state of Victoria**, a major review of post-compulsory schooling led to changes in both curricula and assessment for all subjects. Until this time, assessment was conducted solely via external examinations. It was now proposed that there would be a mixture of school-based, moderated, and external assessments in all subjects. At the same time, an innovative curriculum was introduced that incorporated mathematical modelling activities, all of which had to be formally assessed. According to Stillman's (2007) analysis of this implementation effort, *both the extent and pace of* change threatened the ability of Victoria's education system to sustain curriculum innovation.

The *extent* of change **in Victoria** was significant, with teachers describing it variously as "a revolution", "a radical change" and "a huge shock" as they struggled to reconceptualise what mathematics is about (Stillman, 2007). Even more challenging was the new expectation that their judgments were integral to high stakes assessment in the upper secondary years. Assessment underwent several changes throughout the implementation period. At first, there was a mix of Common Assessment Tasks designed by the state's curriculum authority, an Investigative Project on a centrally determined theme, and a Challenging Problem that schools chose from a centrally provided set of tasks. The latter tasks were school assessed and undertaken by students over an extended period of time. They were supplemented by end of year examinations. Within a couple of years, however, the school assessed tasks were suspended or downgraded to a series of smaller tasks, with a corresponding decline in the contribution of the school assessed component of the course to students' final result (from 50% to around 33%). These changes were ostensibly a response to concerns about student and teacher workload and the difficulty in authenticating students' work. A decade later, a further curriculum review replaced extended Common Assessment Tasks with school-assessed coursework, which Stacey (2001) claims signalled the end of "a bold experiment of assessment driven change and a real focus of problem solving and modelling" (p. 48).

The *pace* of change in Victoria also hindered sustained implementation. The new mathematics courses went from a pilot in a limited number of schools in 1989 to full implementation the following year. Insufficient time was allowed for teachers to come to grips with new expectations about changed practice, and there was low tolerance by the education system for incremental, evolutionary change over time.

On the other hand, **the State of Queensland** introduced an applications and modelling focus in its senior secondary mathematics courses at about the same time as Victoria, following a review of the previous syllabus begun in 1987. The new syllabuses, published in 1992, organised course general objectives into three categories, one of which was Mathematical Applications. Students were required to apply mathematics in situations ranging from purely mathematical to life-related. Interestingly, many teachers who embraced the modelling approaches promoted by these initial syllabuses were also technology enthusiasts, and in-service workshops on the use of graphics calculators played a role in disseminating modelling materials (Stillman & Galbraith, 2009). The syllabuses were revised in 2000 and again in 2009, each time increasing the explicit focus on mathematical modelling. For example, one of the course general objectives is now labelled Modelling and Problem Solving, and another Communication and Justification.

Clearly, modelling has become embedded into senior secondary mathematics curricula in Queensland throughout a series of syllabus revisions. Why is this so, when concurrent implementation attempts failed in Victoria? Stillman (2007) suggests the following possible factors influencing the sustainability of curriculum change:

(a) how teachers were positioned as professionals by the implementing authority and implementation model used (e.g., externally induced change with participant involvement);

(b) pressures from conservative political groups (e.g., general public, parents, teachers, academics);

(c)   considerations of discipline oriented versus whole curriculum approach to curriculum change (e.g., system-wide assessment requirements); and

(d)   changes in society (e.g., desire for equity with respect to socio-economic status). (p. 500)

In **Queensland, external examinations at the end of secondary school had been abolished in 1972,** and replaced by a system of externally moderated school-based assessment. This was part of a move towards designing a curriculum that was more relevant to an increasingly diverse student cohort and assessment that could capture a broader range of learning outcomes than those that could be demonstrated in timed written tests. In each subject, schools design work programs consistent with broad syllabus requirements. They design assessment tasks and award final grades to students for each subject in which they are enrolled, but task design and assessment judgments are moderated at district and state level by panels of experienced teachers to ensure comparability of outcomes. These panels are convened by, and accountable to, the state's curriculum authority. School-based assessment was initially norm-referenced, but the requirement to adhere to pre-determined grade distributions within each school district as well for the state as a whole soon proved to be unworkable. Teachers also struggled with the unaccustomed responsibility of making high stakes assessment judgments, and disliked having to reconcile their professional role of supporting learning with the pressures of a "grade rationing" regime that seemed to disadvantage students. A subsequent review of the state's school-based assessment system recommended a change to criteria and standards based assessment, and this approach has been in place ever since. Under this system, student performance on assessment tasks is compared with explicitly stated standards rather than with the performance of other students. Although the impetus for reform in Queensland came from public dissatisfaction with a narrowly focused curriculum and increasing failure rates on external examinations, the effects of reform undoubtedly led to increased teacher professionalism with regard to curriculum and assessment design and arguably increased public trust in teachers' ability to make sound assessment judgments.

When the first modelling syllabuses were introduced in Queensland in 1992, mathematics teachers were already accustomed to making assessment judgments based on a range of evidence, including results of supervised examinations conducted in school time and students' responses to extended, unsupervised tasks undertaken in class or at home. Although modelling was not well understood by most teachers at first, with many simply grafting an applications approach onto existing teaching methods and tasks (Stillman & Galbraith, 2009), the state's curriculum authority tolerated a slow evolutionary process through successive syllabus revisions. Moderation panels played a part in this process, gradually exerting more pressure to ensure that all modelling-related assessment criteria were being addressed in task design. However, there is still far from uniform implementation across the state, with uptake being described as on a continuum from "minimalist approaches to very rich" (Stillman & Galbraith, 2009).

In Queensland, then, unlike Victoria**, teachers were already positioned as professionals** capable of making high stakes assessment judgments and involved as participants in syllabus change and implementation. With respect to mathematical modelling, the extent of change in Queensland was much less than in Victoria, where a wholesale curriculum review affected every school subject at the same time as a modelling emphasis was introduced in the mathematics curriculum. Also, in Queensland there was greater tolerance on the part of the central curriculum authority for slow change and diverse implementation in schools. While this did not lead to a comprehensive, uniform, radical change in teaching practice, it did seem to reinforce at least the possibility for change that is sustainable.

Nevertheless, teacher autonomy and professionalism in Queensland, and Australia as a whole, is currently under threat due to an increasing policy emphasis on testing at all levels of schooling. This is seen, for example, in the recent introduction of the National Assessment Program – Literacy and Numeracy (NAPLAN) testing regime implemented annually for all students in Years 3, 5, 7, and 9. Rather than being seen as providing a moment-in-time snapshot of student capabilities, the test itself has become the object of teaching, measuring teacher and school performance as well as that of students. At the senior secondary level, the introduction of a national curriculum to be implemented from 2016 has raised the spectre of uniform national assessment at the end of secondary schooling. Although it is unlikely that the

states will be asked to adopt a common assessment process for some years yet, Queensland stands to lose its unique approach to school-based assessment if uniformity dictates acceptance of external examinations to the extent they are used in other states.

Conclusion and discussion

The key difference between the two states then seems to be at the cultural-political level regarding the public trust in which the teaching profession is held (or not): how Queensland came to differ from Victoria in this respect is somewhat subject to speculation, but the effect of a teaching profession that is not well-respected is inimical to progression and development, as the case study shows.

Additionally, however, the case suggests the importance of building partnerships and consensus across the relevant communities for reform policies; development takes time.

This is entirely consistent with the other key difference noted, i.e. that some nations hold teaching in high esteem, getting into the profession is sometimes very difficult, and the quality of qualifications in the profession may be more akin to medics or lawyers. And the degree of autonomy granted the profession in such nations may be a key factor. Thus, for instance, the time table for change in the UK, and many countries where teaching is a political football, tends to be less than 5 years, or an electoral cycle; whereas it is well understood in countries such as Japan that development takes place effectively on at least a ten (more like twenty) year timetable.

References

Lovitt, C., Clarke, D., & Stephens, M. (1986). Mathematics curriculum and teaching program - an invitation for you to be involved. In K. V. Swinson (Ed.), *Proceedings of the 11th Biennial Conference of the Australian Association of Mathematics Teachers* (pp. 4-5). Brisbane: AAMT.

Lowe, I. (1988). *Mathematics at work: Modelling your world* (Vol. 1). Canberra: Australian Academy of Science.

Lowe, I. (1991). *Mathematics at work: Modelling your world* (Vol. 2). Canberra: Australian Academy of Science.

Stillman, G. (2004). Sustained curriculum change: The example of implementation of applications and modelling curricula in two Australian states. In H.-W. Henn & W. Blum (Eds.), *ICMI Study 14: Applications and Modelling in Mathematics Education* (pp. 261-266). Dortmund: University of Dortmund.

Stillman, G. (2007). Implementation case study: Sustaining curriculum change. In W. Blum, P. Galbraith, H.-W. Henn, & M. Noss (Eds.) *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 497-502). New York: Springer.

Stillman, G., & Galbraith, P. (2009). Softly, softly: Curriculum change in applications and modelling in the senior secondary curriculum in Queensland. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Crossing divides* (Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia, pp. 515-522). Palmerston North, NZ: MERGA.

Williams, J.S. (2010). *Audit and evaluation of pedagogy: towards a cultural-historical perspective*. In (eds.) T. Rowlands and K. Ruthven, Mathematical Knowledge in Teaching, (pp 161-178). Dordrecht, Netherlands: Springer, Mathematics Education Library 50 .

B3   Primary Lessons from Japan: towards a high quality curriculum (intended and enacted). Julian Williams and Darinka Radovic

The Japanese education system is justly admired for several qualities and became well known world wide through the TIMSS video project reported by Stigler et al. in 'The Teaching Gap'. Its use of lesson study is widely understood to have had a major influence on the development of quality teaching and learning, especially – but not only - in elementary mathematics (Fernandez & Yoshida, 2004; Lewis & Tsuchida, 1998; Lewis, 2000). Their focus in mathematics has largely been on deepening the engagement of learners with intellectual challenge through problem solving, while at the same time learning new mathematics through an inquiry-oriented learning process focused on one main problem during the lesson (Stigler et al., 1999; Takahashi, 2006). Consequently, the Japanese lesson study approach has recently been spreading round the world as a possible way of dealing with Professional development (see the APEC website, especially East Asia and Anglophone countries are active).

What is less well known however is the way the Japanese curriculum and pedagogic system has worked to benefit from these lesson study 'research lessons', and how this has influenced their texts and professional development. The Japanese curriculum has typically evolved on a ten year timetable: great care is taken in the process to benefit from lesson research, and the curriculum provides a framework for the text books that are produced under government license. There are about 7 Primary mathematics text book series, and along with each text there are – most important – teacher manuals. Many in fact draw on the same research and analysis which has established a consensus about some 'good practices'.

The way in which Japanese curriculum and pedagogic system has worked to benefit from these lesson study 'research lessons', is rooted in their cultural view of curriculum change and professional development (e.g. Lewis, 2000; Lewis & Tsuchida, 1997). The Japanese curriculum has typically evolved on a ten-year timetable. Great care is taken in the process to benefit from lesson research: these spread from teacher to teacher and outside commentators take ideas to national curriculum committees. This curriculum provides a framework for the textbooks that are produced under government license. There are about 7 Primary mathematics textbook series, and along with each text there are – most important – teacher manuals. Many in fact draw on the same research lessons and analysis, which has established a consensus about some 'good practice' (this is why it is not surprising that most of these books follow even the same sequences of questions, see Doig et al. 2011).

For instance lesson study research into the key conceptual lesson on teaching subtraction-through-bridging, or bridging-through-ten has revealed the many and varied responses of children when presented with tasks like $13 - 9$, $12 - 3$, and $14 - 8$. A well-known lesson study shows that the task 13-9 leads to the strategy of subtraction-addition, so 13 is $10 + 3$, and 10 sub 9 leaves 1, and 3 add 1= 4. The manual prepares the teacher for the many other responses that children will make, but this is the dominant one that the teacher may want to discuss with the class. However $12 - 3$ generates a different dominant strategy, subtraction-subtraction: (12-2) -1… while 14- 8 typically splits the class with other strategies (Doig et al., 2011) The manual thus prepares the teacher for the likely results of the problem solving work of the children, who is then prepared to have the various children's mathematical strategies revealed on the blackboard before the class: a key part of the lesson (the major part) occurs when the teacher enables the children to discuss the various methods and help the children to evaluate them. NB the creative problem solving and critical evaluation processes of the children are engaged while building up the basic mathematics: this is no postponement of creative problem solving until the children have 'mastered the basics'.

There has been much debate in Japan about how varied a collection of responses to a question/task is desirable: one manual concluded that a new teacher might better to limit the variety of children's solutions, as this might be easier for them to manage in the 'neriage'. The neriage is the plenary part of the lesson where the teacher tries to 'knit together' the problem solving work of the learners – ALL the learners - into a public revelation of the more advanced conceptualisations involved.

Interestingly, some research in the US (Lewis et al) involves American teachers studying English translations of such Japanese teacher manuals. Also, it is significant that this work goes beyond mathematics and includes science.

The point? One might conclude that Japanese curriculum, as enacted in the classroom, is tied together with the intended (and innovative, problem solving) curriculum through slow, careful work on practice, and on texts and manuals connected to professional R&D work in classrooms. One might also point here to some defects in many other parts of the world, including recent history in the UK. Curriculum change is often highly political, rushed, and often conducted without any deep engagement of teachers or other professionals (if not against their advice). Text books are left to the market, of poor quality and with no basis in research or practice in real lessons, and so with ineffective (if any manuals) Rather than incrementally improving materials, some teachers even look back to materials 20+ years old for quality curriculum material.

What vision – to go back 20 years? But perhaps the current abandonment of a National Curriculum for many schools here in England offers a new opportunity as well as the obvious risks. Perhaps what might be more forward looking is a new policy on Royal Society (or some other bodies) "kite marking": ensuring quality in school curricula, texts and manuals, that could bring together academics and researchers to award some official approval for curricula, texts and manuals coming on the market that meet some quality criteria… as regards evidence-base, professional evaluations, etc.

This is where the analysis of Lesson Study initiatives in different context may provide useful information for upcoming vision. Most of the recent literature on Lesson Studies implementation in different contexts has shown how it can provide a useful model for teachers' professional development, actually improving practices (e.g. Cavey & Berenson, 2005; Inoue, 2011; Fernandez, 2005; etc.). Advances have been made in the translation of the model to different education systems, with increasing information on difficulties encountered during implementation in schools (e.g. teachers lack a research lenses – see Fernandez, 2002; Fernandez et al., 2003).

But what is lacking is how this local research-professional development practices can build into changing policy, instead of 'from reform to the classroom', rather 'from classroom research to reform', in these different contexts (for initial discussions of this see Williams et al., in press)..

This is clearly a ten or twenty year project at least, and requires the collective commitment of thoughtful teachers, researchers, and policy-makers in joint activity focussed on the interests and real experiences of the learners in classrooms. There are lessons here for curriculum development, policy, professional development, and for educational research (Williams et al, 2013).

References
Cavey, L. O., & Berenson, S. B. (2005). Learning to teach high school mathematics: Patterns of growth in understanding right triangle trigonometry during lesson plan study. *The Journal of Mathematical Behavior*, *24*(2), 171–190.
Doig, B., Groves, S. & Fujii, T. (2011). The Critical Role of Task Development in Lesson Study. In L. C. Hart Alston, A. & Murata, A. (eds.), *Lesson Study Research and Practice in Mathematics Education*, (pp. 181-199), Springer Science and Business Media.
Fernandez, C. (2002). Learning from Japanese Approaches to Professional Development: The Case of Lesson Study. *Journal of Teacher Education*, *53*(5), 393–405.
Fernandez, C., Cannon, J., & Chokshi, S. (2003). A US–Japan lesson study collaboration reveals critical lenses for examining practice. *Teaching and Teacher Education*, *19*(2), 171–185.
Fernandez, C., & Yoshida, M. (2004). *Lesson Study: A Japanese Approach to Improving Mathematics Teaching and Learning*. New Jersey: Lawrence Erlbaum Associates.
Fernandez, C. (2005). Lesson Study : A Means for Elementary Teachers to Develop the Knowledge of Mathematics Needed for Reform-Minded Teaching? *Mathematical thinking and learning*, *7*(4), 265–289.
Inoue, N. (2010). Zen and the art of neriage: Facilitating consensus building in mathematics inquiry lessons through lesson study. *Journal of Mathematics Teacher Education*, *14*(1), 5–23.

Lewis, C., & Tsuchida, I. (1998). A lesson is like a swiftly flowing river. *American Educator*, 12–52.

Lewis, C., & Tsuchida, I. (1997). Planned educational change in Japan: the case of elementary science instruction. *Journal of education policy*, *12*(5), 313–331.

Lewis, C. (2000). *Lesson Study: The Core of Japanese Professional Development*. New Orleans: American Educational Research Association Meetings. Retrieved from http://www.csudh.edu/math/syoshinobu/107web/aera2000.pdf

Takahashi, A. (2009). *Characteristics of Japanese mathematics lessons*. Colección Digital Eudoxus, (18). Retrieved from: http://cimm.ucr.ac.cr/ojs/index.php/eudoxus/article/view/379/380.pdf

Williams, J., and Ryan, J. (2013). Research, policy, and practice, In V. Farnsworth & Y. Solomon (eds.) Reframing educational research. London: Routledge.

Williams, J., Ryan, J., and Morgan, S. (in press). Lesson study in a performative culture. In MacNamara et al.

B4   Modelling and problem solving in STEM (with Mathematics)
**   Geoff Wake and Julian Williams**

There has been a continuing worldwide effort since at least the 1970/80s to develop mathematical modelling and problem solving (MMPS) courses with the emphasis on problem solving skills, and there is now a huge body of international research focussed on this, with many examples of curriculum developments both within STEM and across the curriculum. The literature originated in two sources. In Primary schools and upwards, 'modelling' was understood as a part of the process of 'mathematising' realistic situations that 'beg to be organised' mathematically and allow 'guided re-invention' (after Freudenthal, Treffers, Streefland etc) and this approach effectively protected the Netherlands from the 'new math' in the 60s and 70s. On the other hand, in universities to begin with, 'modelling' was seen as the applied version of Polya's problem solving heuristics, and became popular in the Open University and elsewhere in Britain in polytechnics., The "International Conference on Mathematical Applications" (ICTMA) became the international focus of this movement and spread these ideas and practices in Europe and then internationally (see references to ICTMA papers).

In the early days the ICTMA conferences addressed the philosophy of modelling and practical needs of teachers of modelling at all levels for examples of models and contexts; later the research has shifted more to system questions like teacher education, curriculum specification, and crucially its assessment. Of course, authentic assessment of MMPS has proved problematic in some systems: the modelling innovations in the UK of the 1990s almost died when coursework assessment at A-level was restricted. A similar issue arose in Victoria, Australia (but see case study of a different trajectory in Queensland, appendix B2). Interestingly, a resurgence of MMPS in Britain is underway in some universities and in the new curriculum developments at AS level (those considered 'lower stakes', i.e. not intended for entry to university STEM subject, e.g. such as "Uses of mathematics" and the new MEI initiative for a post-compulsory AS Mathematics focussed more on social science and every day contexts).

The two understandings of modelling are now conceived to be parts of the same mathematical process, i.e. problem solving with, and within mathematics. Thus pure mathematics uses 'modelling' too in the sense of formulating new representations and applying heuristics such as constructing diagrams: one learns new mathematics by applying mathematics and developing models to new situations – albeit within mathematics. (Treffers and others refer to horizontal and vertical mathematisation: but they engage the same mathematising processes.)

The problems addressed by mathematical modellers can be rich and varied, and touch on the whole of STEM, and real problems such as sustainability, engineering design, medicine etc. In this trend there has been a body of research into the mathematics and modelling required in working practices – this has led to new perspectives on mathematics and the curriculum (e.g. Noss et al., Williams & Wake, etc). But also this approach engages the academic with the vocational, at least in principle, and the 'employability' agenda.

Real modellers work can look quite traditional and academic compared to some of the modern trends in mathematics problem solving outside mathematics class in school - they deal with very open ended and complex problems that are under specified and uncertain, and much closer to the work in D&T (such as E-scapes work).

A somewhat parallel development in science subjects has been underway under the banner of 'process' science and 'science in society'. (See Fraser et al. for various review chapters.) The 'process' of science similarly raises awareness of the methods of science, whereas the SiS situates science in wider interests, in its practice and effects.

Indeed, sociologists/philosophers of science also now argue that perhaps the traditional science 'disciplines' of Physics/chemistry-biology make little sense. Real STEM now looks more like GRAIN – genomics, robotics, artificial intelligence and nanotechnology' (Ravetz).

The extensive research in this area suggests that curricula can be made more coherent and relevant to learners through modelling and problem solving, but that the obstacles in terms of assessment (and other conservative system effects) tend to resist such changes. This reinforces the need for R&D of assessment and examination processes (see also appendix B6 for an example).

In conclusion, considerable work has been done world-wide to develop mathematical modelling and real problem solving in mathematics education throughout 5-18: any vision for STEM should integrate this. It includes the notion of problem solving and the use of models and emergent modelling in the earlier years, and heuristics for problem solving later on. It links with integration across STEM, and with vocational links and applications as well as 'everyday' and citizenship tasks.

Bibliography
Blum, W., Galbraith, P.L., Henn, H.-W., & Niss, M. (2007). (Eds.) Modelling and Applications in Mathematics Education. The ICMI 14th Study. Springer
Burkhardt, H. (1981) *The real world and mathematics.* Glasgow: Blackie.
Engeström, Y. (1991). Non scolae sed vitae discimus: Toward overcoming the encapsulation of school learning. *Learning and Instruction, 1*(3) 243-259.
Fraser, Tobin, McRobbie (Eds.) *Handbook on science Education.* Springer.
Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures.* Mathematics Education Library . Dordrecht, Nl: Kluwer.
Hoyles, C., Noss, R., & Pozzi, S. (2001). Proportional Reasoning in Nursing Practice. *Journal for Research in Mathematics Education, 32*, 4-27
Kaiser, G., Blum, W., Ferri, R., & Stillman, (Eds.) (2011). *Trends in Teaching and Learning of Mathematical Modelling.* ICTMA-14. Springer.
Lesh, R. & Zawojewski, J. (2007). Problem solving and modelling. In F. Lester (ed.) *Second Handbook of Research on Mathematics Teaching and Learning.* (pp763-804) Reston, Va., NCTM.
Noss, R. , Bakker, A., Hoyles, C., & Kent, P., (2007) Situating graphs as workplace knowledge. *Educational Studies in Mathematics, 65*, 3, 367-384.
Polya, G. (1957) *How to solve it.* Princeton, NJ: Princeton Univ Press.
Roth, W.-M., & McGinn, M. K. (1997). Graphing: Cognitive Ability or Practice? Science Education, 81 (1), 91-106.
Schoenfeld, A. (1992). Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (ed) Handbook of Research on Mathematics Teaching and Learning. (pp 334-370.) Reston, Va., NCTM.
Streefland, L. (1991). Fractions in Realistic Mathematics Education. Mathematics Education Library 8. Dordrecht, Nl: Kluwer.
Treffers, A. (1987). *Three Dimensions: A Model of Goal and Theory Description in Mathematics Instruction – The Wiskobas Project.* Dordrecht: Reidel.
Wake, G.D., & Williams, J.S. (2003). Using Workplace Practice to Inform Curriculum Design, In S.J. Lamon, W.A. Parker & S.K.Houstone (Eds.), *Mathematical modelling a way of life,* Chichester, U.K.: Horwood Publishing.
Wake, G.D. (2005). Understanding workplace mathematics from a curriculum development perspective, paper presented at International conference, In Annual Proceedings of Adults Learning Mathematics Group. Gothenburg, Sweden.
Wake, G.D. (2005). Functional mathematics: More than "back to basics". Nuffield Review of 14–19 Education and Training, pp. 1–11.
Wake, G.D. (2007). Considering workplace activity from a mathematical modelling perspective. In Blum W., Galbraith P.L., Henn, H.-W., & Niss M. (Eds.), Modelling and Applications in Mathematics Education. The 14th ICMI Study. (pp. 395 – 402). New York: Springer.
Wake, G. D. (2011) Modelling in an integrated mathematics and science curriculum: Bridging the divide, Paper presented at the Proceedings of the Seventh Conference for European Research in Mathematics Education

(CERME 7). Rzeszow, Poland, Retrieved 10th August, 2011from
http://www.cerme7.univ.rzeszow.pl/WG/6/CERME7-Wake.pdf.

Wake, G. D. (2013). Drawing on understanding of workplace practice to inform design of general mathematics curricula. In Damlamian, A., Rodrigues J.-F., Sträßer R. (Eds.), Educational Interfaces between Mathematics and Industry (EIMI). An ICMI-ICIAM Study 20. Berlin – Heidelberg: Springer.

Usiskin, Z. (2007). The Arithmetic operations as mathematical models. In W. Blum, P.L. Galbraith, H.-W., Henn, and M. Niss. (Eds.) Modelling and Applications in Mathematics Education. The ICMI 14th Study. (pp 257-264). Springer.

B5 Innovations in assessment

B5a E-portfolios and comparative judgment techniques

The E-scape projects (2003-2010) have shown how creativity, teamwork and innovation can be assessed in semi-structured contexts in science and technology (and beyond) in UK schools. The two key innovations are (i) the online capture of students work (photos, oral reporting, etc.) on semi-structured design type tasks, which both structures their work and captures evidence of the learners' work-in-progress in real time as they attempt to solve tasks; and (ii) adaptive comparative judgement (CJ) of the resulting portfolio of information.

There has been recent growth of interest in CJ wherein evaluators/assessors make evaluations of texts by simply comparing 'pairs' of texts and judging which is better. By appropriately focussing repeated judgements on those 'difficult' judgments (i.e. where two 'close' texts A and B are judged A>B by some and B>A by other judges) the multiple judgements converge on an ordering (and a scale) that is as reliable as an analytic marking or criterion referenced assessments that have become the norm. For instance in one study an average of seven paired judgements were necessary per script, but the numbers concentrated on some judgments were as high as 20 (E-scape report.)

The validity of these scores is, of course, an emergent consensus from (and so a product of) the collection of judges selected to make the judgments: but the process does allow analysts to identify outlier, or misfit judges, and hence train these to agree with the community consensus (or better, evaluate the reasons for the differences in judgments of 'quality').

These methods are ideal for assessing projects and assignments where open-endedness allowing for creativity makes traditional criterion-referenced analytic rubrics impossible to construct without constraining the very creativity the tasks are designed to encourage. But the method can be (and has been) tested and calibrated against traditional methods. For instance Jones (ref) showed that a group of mathematics students could evaluate their own work and arrive at judgments that were consistent with that of experts and traditional rubric-based assessment.
The literature supporting the efficacy of this approach now extends to school students D&T, science, geography, etc projects, and to students and student teachers involved in self-assessment.

Bibliography
Bramley, T. (2007) Paired Comparison Method. In Newton, P., Baird, J.A., Goldstein, H., Patrick, H., Tymms, P. (Eds.) *Techniques for monitoring the comparability of examination standards.* London: Qualifications and Curriculum Authority.

Heldsinger, S., Humphry, S. (2010) Using the Method of Pairwise Comparison to Obtain Reliable Teacher Assessments. *The Australian Educational Researcher, 37* (2), pp.1-19.

Jones, I, Inglis, M., Gilmore, C., Hodgen, J. (in press) *Measuring Conceptual Understanding: The Case of Fractions.* In Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education, Vol. 1, Kiel, Germany: PME.

Jones, I, Swan, M., Pollitt, A. (Forthcoming) Assessing mathematical problem solving using Comparative Judgement. *International Journal of Science and Mathematics Education.*

Jones, I. (Forthcoming) Peer assessment without assessment criteria. *Studies in Higher Education.*

Kimbell, R. (2007) E-assessment in project e-scape. *Design and Technology Education: an International Journal, 12* (2)

Kimbell, R., Wheeler, T., Miller, A., & Pollitt, A. (2007). *E-scape: e-solutions for creative assessment in portfolio environments. Technology* Education Research Unit, Goldsmiths College, London.

Pollit, A. (2012) The Method of Adaptive Comparative Judgement. *Assessment in Education: Principles, Policy & Practice,* 1–20

B5b Measurement of dispositions/affect and learners' trajectories
Maria Pampaka and Julian Williams

It is widely claimed in the literature on attitudes and dispositions that the current curriculum and assessment structures in schools over-emphasise cognitive learning outcomes at the expense of the affective and metacognitive; much research and development has been done in this area, and some robust measures of these alternative learning outcomes are available and in use in research and in surveys such as PISA (in which self-efficacy has been used, inter alia). Recent work in England also developed robust instruments for the measurement of new STEM attitudinal measures, such as mathematics self-efficacy, confidence, dispositions to study more STEM, dispositions towards modelling applicable from secondary to HE level (Pampaka et al, 2011, 2012, 2013; Pampaka & Williams, 2010)

Clearly some dispositions are likely to be important in indicating whether a student will choose to engage deeply with a subject, or choose to study it in future when the option arises (Hutcheson, Pampaka & Williams, 2011; Pampaka, Williams & Hutcheson, 2012). All this is clearly salient to encouraging students to study STEM subjects and enter STEM degrees etc.

Additionally there is evidence that some teaching has negative effects on relevant dispositional outcomes (Pampaka, Williams et al, 2012, Pampaka, Wo et al 2012; Wolf, 2000). It is even argued that those countries that seek to 'drive up standards' can sometimes pay a price in terms of learners attitudes to science (e.g. Ruthven et al., 2010; Pampaka, Pepin & Sikko, 2012). One might label this effect 'drive up standards and drive them out' of science/STEM.

This may be because the focus on 'driving up standards' at the political level leads to an off-putting over-emphasis on test performativity in the school and classroom levels that students dislike. But one antidote to this might be to make measures of student disposition and choices they make into accountability measures that might counterbalance grade-measures.

Against this, one would not want to see teachers training their students to fill in survey measures 'in the desired manner' as sometimes is said (anecdotally) to occur around university NSS surveys.

Bibliography

Fraser, B. J. (1998). Classroom Environment Instruments: Development, Validity and Applications. *Learning Environments Research, 1,* 7-33.

Fraser, B. J., McRobbie, C. J., & Giddings, G. J. (1993). Development and Cross-National Validation of a Laboratory Classroom Envronment Instrument for Senior High School Science. *Science Education, 77*(1), 1-24.

Gwilliam, L. R., & Betz, N. E. (2001). Validity of Measures of Math- and Science-Related Self-Efficacy for African Americans and European Americans. *Journal of Career Assessment, 9*(3), 261-281.

Hutcheson, G. D., Pampaka, M., & Williams, J. (2011). Enrolment, achievement and retention on 'traditional' and 'Use of mathematics' pre-university courses. *Research in Mathematics Education, 13* (2), 147-168.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (Third ed., pp. 13-103). USA: American Council of Education and the Oryx Press.

Pampaka, M., & Williams, J. (2010). Measuring Mathematics Self Efficacy of students at the beginning of their Higher Education Studies. In M. Joubert & P. Andrews (Eds), Proceedings of the British Congress for Mathematics Education (BCME), Manchester, April 2010, pp. 159-166

Pampaka, M., Kleanthous, I., Hutcheson, G. & Wake, G. (2011). Measuring mathematics self efficacy as a learning outcome. *Research in Mathematics Education, 13*(2), 169-190.

Pampaka, M., Pepin, B., & Sikko, S.A. (2012). Supporting students' transition to HE or alienating them: Pedagogies and Transitional Experiences in two educational contexts. *Paper presented under the Keynote Symposium "Lost in transition: student alienation in the transition to University", at the Annual Conference of the British Educational Research Association (BERA 2012).*

Pampaka, M., Williams, J., Hutcheson, G., Black, L., Davis, P., Hernandez-Martinez, P., & Wake, G. (2013). Measuring Alternative Learning Outcomes: Dispositions to study in Higher Education. *Journal of Applied Measurement, 14*(2), 197-218.

Pampaka, M., Williams, J. & Hutcheson, G. (2012). Measuring students' transition into University and its association with learning outcomes. *British Educational Research Journal*, *38*(6), 1041-1071.

Pampaka, M., Williams, J., Hutcheson, G. D., Wake, G., Black, L., Davis, P., & Hernandez-Martinez, P. (2012). The association between mathematics pedagogy and learners' dispositions for university study. *British Educational Research Journal, 38*(3), 473-496.

Pampaka, M., Wo, L., Kalambouka, A., Qasim, S., & Swanson D. (2012). Teaching and Learning Practices in Secondary Mathematics: measuring teaching from teachers' and students' perspectives. *Paper presented at the Annual Conference of the British Educational Research Association (BERA 2012)*.

Ruthven, K. (2010). Using longitudinal, cross-system and between-subject analysis of the TIMSS study series to calibrate the performance of lower-secondary mathematics education in England. *Proceedings of the 7th. British Congress of Mathematics Education* [BCME 7], published as *Proceedings of the British Society for Research into Learning Mathematics* 30(1), 183-190.

Williams, T., & Williams, K. (2010). Self-Efficacy and Performance in Mathematics: Reciprocal Determinism in 33 Nations. *Journal of Educational Psychology, 102*(2), 453-466.

B6   Case study of NLT integrated science and maths subject, Netherlands

Nature, Life and Science (NLT) is an integrated science and maths subject, which was introduced in upper secondary education in the Netherlands in 2007. In a surprisingly short time, the course has expanded to become a well-supported and recognised course and qualification. Significantly, the assessment is school-based, the choice of modules for study is broad, and the development of curriculum materials is ongoing.

In the Netherlands, the separate curricula of science and mathematics in upper secondary education are fairly traditional, and there is a lack of interdisciplinarity. Mathematics in particular is an isolated subject. In general, there are few contacts with colleges or university or with professionals outside education. To address this situation, the NLT project was instigated in 2006. The general aims of NLT are the following: to make the natural sciences and technology more attractive to a broader range of students, and to create coherence in the different subjects of the sciences and mathematics. It is a supplement to the study of physics, chemistry, biology and mathematics, and can be regarded as preparation for higher education in the field of science and technology.

NLT is an optional subject within the science stream of **upper secondary education**, to be completed with a **school examination**. The mandatory team-teaching approach, and the lack of a fixed and established curriculum were fundamental but also potentially problematic features which risked a low take-up of the new course. Teaching teams were required to include at least three teachers (from biology, chemistry, earth science, mathematics, physics) with time for preparation and networking. There were associated demands on the timetable and on support by a technician. Surprisingly, 34% of schools offered NLT in its first year, and high numbers of students opted for it. These numbers have continued to rise (2011: 233 schools - 45%)

NLT schools are encouraged to participate in networks with a university, a college and a number of schools. The coordinators of the network, from the universities and colleges, support teachers with the content of modules, facilitate meetings of teachers, provide maintenance of modules, and help with contacts in the region for out-of-school activities. Some of the key characteristics of NLT are

- the content is mainly interdisciplinary;
- it is taught by teams of teachers with a degree in biology, chemistry, earth sciences, mathematics and/or physics;
- students experience the usefulness/necessity of mathematics in sciences;
- most subjects start from a context: concepts follow from the context;
- the subject depends on structural contacts between secondary and higher education;
- content is developed in networks of teachers from secondary education and experts from colleges and universities, sometimes involvement of other institutes or business enterprises
- content is flexible and up-to-date;
- school based exam only.

The curriculum development and implementation of this programme are worth further consideration, and have been studied by researchers to better understand the factors which enhanced the advanced and interdisciplinary character of NLT (Eijkelhof, and Krüger, 2009). They found that the key elements of the process were interaction with higher education; stimulating teacher cooperation through a combination of support and requirement; material development & implementation

The Steering Committee certifies new modules following an extensive trialling and feedback process. The Steering Committee also takes care of the *implementation* of the subject including

assessment, evaluation and adaptation of the draft curriculum during the first few years after the introduction.

Assessment

The subject NLT is only examined in a school examination, and therefore there are no binding regulations in the field of assessment or an end-of course test, though it is subject to a Programme for Assessment and Completion with key recommendations. Written tests are not sufficient: the students will carry out many activities, the learning outcomes of which cannot easily be recognised by means of a written test. But omission of written tests would not do justice to students' learning. Testing includes: digital diagnostic tests; a description of eg. reports which the students have to submit while working on the module, and modes of assessment of these, including peer-review; a final written test with model answers, including short, closed/multiple choice questions in order to test the knowledge-related objectives; short, simple calculations and questions relating to application; complex and more open assignments, where the acquired knowledge and skills are placed in a new context. They can also be assignments requiring the analysis of a set of data, the design of research or the realisation of part of the design cycle; one or more (choice) essay-questions, in order to test whether the candidate is able to use the learned concepts to reflect on a new situation.

NLT is examined in at the school level. The lack of a final central exam requires extra guarantees for the quality of the subject. During the development phase (2007-2010), schools were clustered into regional NLT-networks with at least one institution in higher education. The teams of teachers in the participating schools and the institution in higher education organised NLT- network meetings with the following functions:

- to inspire and to stimulate each other by discussing practical situations in their own school and to exchange educational resources and student materials
- to support each other by discussing organisational and content-based problems, which arise during the implementation of the subject
- to exchange programmes for assessment and completion, to compare and discuss those also with respect to their compliance with the examination programme
- to evaluate the tests including assessment and standards.

Quality assurance through the NLT–network involves schools in evaluating and assessing each other's school examinations. Schools can use this network evaluation to give account of the quality of their school exam to the Education Inspectorate. This process is supplemented by external visits and supervision from the education inspectorate.

To date, more than 66 modules have been certified, including modules by other institutes, certified according to the same criteria. This has involved teachers from 152 schools (98 development / testing; 54 testing); experts from 47 institutes (10 universities, 12 colleges, 25 other institutes, including research institutes and companies). The character of the modules is suggested in the following list: *Forensic Science; Driving and drinking; Heart disease; Life in the International Space Station; Glue and attach; Medical imaging; Brain and behaviour; Measuring galaxies; Bio-informatics; Pharmacology; Molecules of life; 2nd generation biofuel; Molecular gastronomy*

References:

Eijkelhof, H. and Krüger, J. (2009) Improving the quality of innovative science teaching materials, Paper presented at the ESERA 2009 Conference, Istanbul   www.betavak-nlt.nl/english

Modules
[Molecular Gastronomy - Science in the Kitchen - manual](#) (Pdf)
[Molecular Gastronomy - Practical 1](#) (Pdf)
[Molecular Gastronomy - Practical 2](#) (Pdf)

[Food or Fuel? - manual](#) (Pdf)

[The Molecules of Life](#) (Pdf)
[The Dynamic Earth](#) (Pdf)

B7   Enrichment strategies in science and mathematics

Rachel Sharkey and Julian Williams

Enrichment has been defined in different ways by different people often focusing on one particular aspect of teaching and learning or curriculum. It is argued that a conceptual framework for enrichment should not include just one aspect of teaching and learning or the curriculum but should incorporate the following four elements: (i) curriculum (intended, implemented and achieved (Mullis, Martin et al. 2009)), (ii) teaching and learning (presentation by the teacher and experience of the students), (iii) environment and organisation (setting, location and organisation of the lesson as well as organisation and management of the school), and (iv) assessment (identifying what has been learnt and inferring the students' subsequent needs).

Enrichment can be seen as a stimulating experience where the teaching approach and environment are as important as the learning process. For instance, "Enrichment acts to counter concerns about mathematics teaching and learning and to compensate for the perceived shortfall in the curriculum" ((Feng 2009)) providing for a wider participation, which now includes "almost every .. group of students" (p26).

Within the context of gifted and talented, enrichment has been used as an intervention strategy in order to avoid talented students repeating unchallenging work (Renzulli and Reis 1994). Renzulli's Schoolwide Enrichment Model is underpinned by three types of enrichment (Renzulli and Reis 1994; Renzulli and Renzulli 2010): Type 1 is receptive and research based and includes guest speakers, field trips etc.; Type 2 encourages the development of skills; Type 3 is about being an independent and creative thinker and investigator. In contrast to the idea of pull-out lessons in the Schoolwide Enrichment Model (Moon, Feldhusen et al. 1994; Renzulli and Reis 1994), the eNRICH project intervention was through out-of-school workshops and occasional visits and lectures (Smith 2006). Enrichment can be provided for all students through enrichment clusters, the only criteria being an interest in the topic. The aim of enrichment clusters is to provide engaging and enjoyable activities as part of 'authentic learning' which involves thinking effectively about real-life problems and how to solve these (Renzulli 1997). A good problem should be posed which engages the learner in an environment that supports the learner and teaches the learner not just to solve the problems posed but also the process of how to approach such problems (Piggott 2004a; 2004b;).

"Enrichment is any type of activity or learning which is outside the core of learning which most children undertake…and the introduction of new methods for approaching work" (p18) (Eyre and Marjoram, 1990). The common themes in the literature about enrichment are that problem solving and metacognition are key parts of many enrichment activities, as enrichment activities are devised to fill a perceived gap or deficiency in the curriculum, and supplement what is being taught.

Enrichment activity in STEM often engages students in out-of-school activity or working with visiting ambassadors or neighbourhood engineers; thus enrichment benefits from the community playing its part in education, in a way that is difficult to embed.

However, experimental work across the curriculum, or with new technologies can find a role n enrichment activity too, and sometimes this can and does become embedded in school practices; the literature often mentions the important role here for broadening teacher's perspectives on education (by working outside their comfort zone) as well as learners..

The implications for R&D are clear: a relatively small investment in enrichment can provide a focus for identifying and developing areas of the curriculum that deserve attention. **Seen as an R&D opportunity, enrichment can provide a part of a strategy for continual development and improvement. As such research should attend more to the process of embedding lessons learnt from enrichment into the curriculum and everyday practices of teaching, learning, and assessment.**

Bibliography:

Eyre, D. and T. Marjoram (1990). *Enriching and Extending the National Curriculum*. London, Kogan Page Ltd.

Feng, W. Y. (2009). *Developing a Framework for Understanding Mathematics Enrichment in the UK*. Doctor of Philosophy in Education Thesis, University of Cambridge.

Moon, S. M., J. F. Feldhusen, et al. (1994). Long-term effects on an Enrichment Program Based on the Purdue Three-Stage Model. *Gifted Child Quarterly 38*(1): 38-48.

Mullis, I. V. S., M. O. Martin, et al. (2009). *TIMSS 2011 Assessment Frameworks*. Retrieved 5 May 2012, from http://timssandpirls.bc.edu/timss2011/downloads/TIMSS2011_Frameworks.pdf.

Piggott, J. (2004a). *Developing a Framework for mathematical enrichment*. Retrieved 9 April, 2012, from http://www.icme-organisers.dk/tsg18/S63J.Piggott.pdf.

Piggott, J. (2004b). "*Mathematics enrichment: What is it and who is it for*?" Retrieved 9 April, 2012, from http://www.leeds.ac.uk/educol/documents/00003649.htm.

Piggott, J. (2007). The nature of mathematical enrichment: a case study of implementation. *Educate 7*(2): 16.

Renzulli, J. S. (1997). *How To Develop an Authentic Enrichment Cluster*. Retrieved 9 April, 2012, from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED420954.

Renzulli, J. S. and S. M. Reis (1994). Research Related to the Schoolwide Enrichment Triad Model. Gifted Child Quarterly 38(1): 7-20.

Renzulli, J. S. and S. R. Renzulli (2010). The Schoolwide Enrichment Model: A Focus on Student Strengths and Interests. *Gifted Education International 26*(2-3): 140-157.

Smith, C. (2006, October 2006). "*eNRICH Mathematics Project Evaluation Interim Report.*" Retrieved 9 May, 2012, from http://nrich.maths.org/2719.

Appendix C

Methodology: the literature search process

Dimitrina Kaneva

Relevant literatures were collected at several stages. The preliminary stage of gathering evidence involved collating lists of peer-reviewed academic articles by keywords searching in electronic databases (ProQuest – BEI, AEI, ERIC), websites and other sources. The key terms used included: STEM education, science, mathematics, technology, engineering, curriculum, assessment, integrated studies, cross-curricular learning, STEM vision, innovation, future, international comparisons, interdisciplinarity and synonyms. Ten searches were performed to obtain initial ideas about the state of play in the researched field in terms of curriculum, assessment, integration and interdisciplinarity. There was a specific focus on existing literature reviews and international comparisons of curriculum and assessment (2 searches). Initial searches were restricted to 2004-13, on the assumption that most Handbook and other literature reviews (e.g. taken from the "Review of Educational Research") would provide a summation of the earlier work.

13 additional searches were performed in the area of curriculum innovation and specific case studies looking at robotics, nanotechnology, sophisticated science apparatus, micro science. After each database search the obtained raw results (titles and abstracts) were scanned for articles which either related directly to the study (to the literature review questions or state of the art case studies), or indirectly but could be potentially useful in framing innovative practice.

A total of number of 8003 articles were returned from all the searches. After exclusion of those that were not relevant to the report, the research team selected 516 articles to go to the next stage of review. The details of these potentially relevant papers were exported from the search databases for further use by the research team. Details of each paper (reference, web-link, key words and abstract) were stored in a shared online document.

The second stage included a more detailed read of the titles and abstracts and organisation of the papers under thematic headings relevant to the research questions and the sections of the report. A further 40 papers were excluded at this stage.

In addition, several other methods were used to obtain evidence:

- Extensive search for evidence of international educational systems on the internet – Eurydice, TIMMS Vol. 1 and 2 Encyclopaedia, European Commission (see Appendix A)
- Manual search of key journals: Journal of Curriculum Studies –all volumes since 2004; Review of Educational Research – volumes since 2004; Design and Technology Education: An International Journal – since 2004; International Journal of Science Education - all volumes from 2013 and 2012, Handbook of Science Education – 2011 edition. This search returned a total of 55 relevant papers. The articles returned from the manual search were cross-checked with the database and the majority were already added. Details about the new relevant articles were uploaded. Consequently, the team identified why these articles were not found with the initial searches. The most obvious one was the key words of the articles – they were different from our initial search terms. Alternatively, as with the Handbook of Science Education, the papers were not available through online databases.
- Snowballing and following up on the references of relevant papers, especially these based on literature reviews and meta-analyses returned a further 21 results.
- Experts in the various respective fields were asked to recommend papers they were aware of. As a result we were given access not only to existing published materials available in the databases but also papers which are due to be published in the very near future. 10 such papers were obtained.

Details of the articles obtained with the systematic search and used in the main report are available as a reference list. The case studies of developments in appendix B have their own reference lists consisting of useful literature.

Table: Inclusion and exclusion criteria

	Inclusion	Exclusion
Review period	2004-2013	Anything before 2004
National/international scope	Anglophone countries Specific focus on: UK USA/Canada/Australia Finland, Singapore Japan	No set exclusion criteria. Articles from a number of additional countries were obtained to assess the evidence but these were not satisfactory in terms of quality or scope (e.g. the Turkish textbook analyses)
Language	English	Non-English
Type of data in initial survey	Peer reviewed journals, Lit reviews,	Commentaries, Speeches
Age range	Compulsory school and post-16 5-18 years	Higher / Adult education Teacher education University/ Undergraduate students
School type	Mainstream Primary/secondary/Middle/ High	Special schools Other non-traditional educational establishments, Summer schools
Case studies	Interdisciplinarity/Integration Problem solving Credibility	Science in Society/SATIS (due to lack of time)
STEM Provision	Integrated studies Cross curricular approaches Curriculum and Assessment International comparisons Science and Mathematics, STEAM	University STEM provision Digital technology teacher supply