Cambridge Assessment

A Cambridge Approach to *improving education*

Using international insights to manage complexity

December 2017

Contents

Acknowledgements

Thanks to Abbi Barnett, Dan Bray and Karen Kester, from Cambridge Assessment International Education, for reviewing and commenting on the drafts of this document; their support has been invaluable. Thanks go to Bill Schmidt, whose original work on TIMSS data provided a vital catalyst for wider examination of 'curriculum coherence'. Thanks also to Gabriel Heller Sahlgren at the Centre for the Study of Market Reform of Education – we were right to push against the 'simple' story of Finnish success. Thanks to Sirkka Ahonen, University of Helsinki, for her tireless responses to endless questions. Thanks also to Lucy Crehan for many great conversations about the issues in this text. And finally, thanks to all the teachers and officials I have interviewed; those from Singapore and Finland in particular.

Tim Oates CBE, Group Director
Assessment Research and Development

Reading this document

Educational improvement is a key objective of policy makers and educationalists around the world. There is increasing transnational dialogue about how to identify the most promising improvement strategy, and how to manage effective implementation. This document is founded on systematic analysis of transnational comparative research, and focuses particularly on the importance of coherent, well-managed change. It presents a framework of 'control factors' and 'explanatory factors', derived from systematic analysis of international comparisons.

Our approach draws on meticulous and wide-ranging study of the history of improvement across a range of jurisdictions, alongside insights from specific examples of effective system-level change, as well as instances where policy aims have not been fully realised.

It does not give 'ready-made' solutions. It rejects naïve 'cherry picking' and 'policy borrowing' from one system to another. Rather, it provides a powerful framework for understanding the specific operation of different systems at specific times, and for policy formation. It underpins our commitment to support policy makers and educationalists in capacity building, and to enhance transnational exchange regarding improvement and innovation.

In this document we do not examine the detail of curriculum specifications, assessments, learning resources and so on. Each of these of course requires careful design, management and evaluation. The details matter, but our research suggests that a coherent overall 'take' on system performance is essential. It is vital 'framing' for more specific actions, without which detailed policies may have reduced effectiveness. The document helps with developing overall strategy for improvement – a process which historically has proved to be frustrating and demanding.

While our approach emphasises that effective improvement policy demands an understanding of complex relations and interactions within each national setting, it does not mean that rapid, modest action is not possible. Far from it, it suggests that full understanding of system relations and context can guide highly targeted and specific action, and maximise the impact of effort and expenditure.

The analysis provides a basis for:

- formulating policy options
- assessing what interventions can and should be made, and likelihood of success
- anticipating dependencies, interactions, and impact of externalities
- monitoring and evaluating impact and formulating options for 'fine tuning' policy actions
- determining actions on communication, intelligence gathering and consultation.

Using this document

This document does not give precise steps to formulating policy or managing implementation. We think that to do so would be quite wrong. Different nations, at different times, face different challenges, have different resources available and are presented with contrasting opportunities to effect change. Sometimes urgent action is required, sometimes the long view needs to be taken. In recognition of this, we do not here recommend a fixed approach to using the insights and approaches outlined in this document. Instead, the text asserts some strong principles and models, underpinned by research, to support effective policy formation and implementation strategy. This is intended to guide thinking on policy formation, making sure that policy formation takes a more comprehensive view of the forces and factors at work in education systems.

The models outlined here have led to governments adopting new approaches to policy formation and management; for example leading one administration to set up a formal committee to review and better align different aspects of government policy on inspection, accountability, curriculum and assessment – something which had been neglected in the past, and had led to inefficiencies and contradictions.

This document offers 'high level organising principles' – they are no less useful for being high level. They have extremely practical applications.

1 A Cambridge Approach

For over 160 years, Cambridge Assessment has worked with nations around the world to improve education. Interest in international comparisons has blossomed in recent years, and this is therefore an area that has continued to be an important part of our effort as we review the field and make use of the latest developments in comparative methods. We always have worked in close collaboration with national governments and schools, and in doing so always seek to ensure that the advice and services which we provide are grounded in the local context and reflect the specific improvement objectives they are designed to support.

The policy support which we offer in this document has been developed, over the past decade, through constant review of transnational research and practical development work with a range of nations. We strongly believe approaches must take into account the specific pressures and possibilities which make up the context in each national setting. In response to this, the frameworks in this document explicitly are designed to support sound analysis of context and circumstance, and to enable evidence-based policy formation for educational improvement.

With the rapid growth of interest in international comparisons we thought it important to provide research not just for illumination and reflection on the ways things are, and why, but also with the practical objective of supporting on-the-ground action to improve educational attainment, equity and engagement in learning. We very much hope that A Cambridge Approach to Improving Education will assist you in that endeavour.

Simon Lebus
Group Chief Executive

'designed to support sound analysis of context and circumstance, and to enable evidence-based policy formation for educational improvement'.

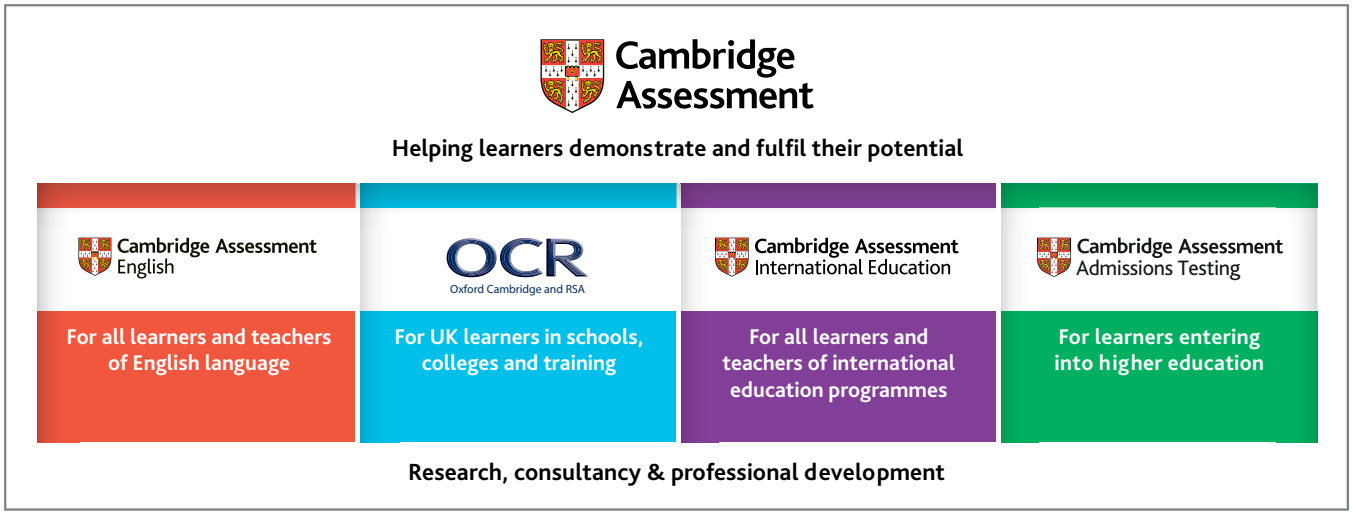
2 Cambridge Assessment, our principles and values

At Cambridge Assessment, our purpose is to help learners demonstrate and fulfil their potential. We care about making a difference for every learner.

As a department of the University of Cambridge, we provide education programmes and exams in over 170 countries offering global recognition. We unlock the power of education for millions of learners of all ages and abilities.

We have unrivalled depth of experience in national education systems, international education and English language learning. We are an international not-for-profit organisation with unique strengths and 160 years of expertise. Our qualifications are backed by the largest research capability of its kind.

We support and learn from teachers, schools and governments. Together, we are shaping education and creating a confident future for learners and a real and lasting impact on the world.



Organisational structure of the Cambridge Assessment Group

Our research underpins all our qualifications and education programmes. Across Cambridge Assessment we have a team of more than 100 researchers, which makes our research capability the largest of its kind. It is this research strength that enables us to help teachers, learners and governments stay at the forefront of education and unlock its power.

But our research is not just about ensuring our qualifications and services are the very best for learners. It's also designed to add to knowledge and understanding about assessment in education, both nationally and internationally. We also carry out research for governments and agencies to inform their education reform programmes. It's all with one goal in mind – helping learners.

We strive to open doors for learners, to unlock the power of education and give them the confidence to thrive. We work with many national educational organisations and ministries through our international organisations, Cambridge English Language Assessment and Cambridge International Examinations. We work to improve standards of education, creating opportunity for learners around the world.

Cambridge Assessment has a high number of experts, with proven experience in curriculum and assessment design, and as a part of the University of Cambridge has access to world-leading resources, skills and research. We work in collaboration with institutions such as University of Cambridge Faculty of Education, Cambridge University Press and Fluentify to offer a comprehensive service to our partners.

In the UK we have been working with industry leaders to develop real industry projects for our Cambridge Technical and Cambridge National qualifications to give learners a head start in their chosen career.

3 Ethical considerations – an important starting point

There are important ethical considerations when undertaking analysis of the performance of education arrangements. These ethical issues have, in turn, important practical consequences.

Ethical considerations apply to the actions and recommendations of those undertaking analysis as well as those with executive responsibility. Deciding to analyse curriculum content rather than teaching quality, assessment rather than teacher workload; these kinds of decisions carry important responsibilities. There are an increasing number of international organisations offering analysis services, coinciding with a drive by governments to enhance the performance of their education arrangements. What to focus on, what to examine, should be considered extremely carefully. Undertaking curriculum review, or review of other key aspects of education arrangements, can be extremely disruptive and costly – we see it as our role to support policy makers and educationalists in improvement which they have elected to undertake; we do not see it as our role to stimulate review on an unsolicited basis.

Our 'control factors' approach recommends that 'analysis should precede action'. In this document we highlight the fact that the scope of this analysis is itself important, and offer advice which can increase its effectiveness.

There now is substantial international discussion of educational performance and the means of securing improvement. 'Single factor' discussions arise all too frequently:

'school autonomy', '21st century skills' or 'assessment reform'[1]. We argue that not only do the assumptions built into these discussions need to be scrutinised and the background evidence examined, but even if a 'single factor' approach holds well-evidenced promise (such as a specific approach to early reading, or a revised approach to science practical assessment), effective implementation will only be likely to occur if the wider context is considered. This document offers a systematic approach to the understanding, analysis and management of this wider context.

4 Improvement: the importance of context

A focus on determining the impact of specific actions and instruments predominates in education. The following enquiries are typical: What is the impact of this particular early reading intervention? What is the effect of having a specialist maths teacher in primary education? Did this change in a qualification benefit boys more than girls? Such studies predominate, and many of them produce invaluable evidence to support system improvement and enhancement of education. But they also have another effect: they focus policy makers and educationalists on very specific aspects of education. Effort to improve education can become narrowly 'initiative based'. In education, it certainly is not wrong to examine practice, structures and instruments in meticulous detail. It is important to understand how specific things work, and the theory and assumptions which lie behind them. But too frequently, specific initiatives on reading, management reorganisation, assessment, and curriculum reform fail to achieve expected levels of improvement. When rolled out to whole systems, approaches which bore promise on the basis of small-scale research and pilot programmes fail to realise the promised gains. There can be a number of reasons for this: poor implementation, failure of professionals to understand the background rationale of change, and so on. But in addition, frequently there is failure to understand complexity and context. Reform effort focused on international comparisons throws light on this. Even implemented with great commitment, efforts to use something which worked well in one country frequently can result in disappointment when used in another. A principal reason for this is the challenge of interactions and relations. Whilst research can cause us to focus on the form of a specific aspect of the totality of education arrangements in a jurisdiction, the overall performance of those arrangements is determined not only by the specific form of each element of those arrangements (of assessment, of pedagogy, of inspection, and so on) but by the relations between them.

Reform policy, and on-going policy directed at maintaining quality, needs to incorporate a recognition of these relationships and their complexity. There is compelling evidence for the importance of this.

1 Benton, T. (2014) *A re-evaluation of the link between autonomy, accountability and achievement in PISA 2009*. Cambridge Assessment; Allen, R. (2010) *Does school autonomy improve educational outcomes? Judging the performance of foundation secondary schools in England.* DoQSS working paper no 10–12, Institute of Education, London; '21st century skills': Suto, I. (2013) *21st Century Skills: ancient, ubiquitous, enigmatic?* Cambridge Assessment; 'assessment reform': Oates, T. (2016) Assessment: the need to 'do nothing'. In Pring, R. and Roberts, M. (Eds) *A generation of radical educational change*. Routledge.

Firstly, evidence from the aforementioned history of the failure of specific initiatives; where high-quality research strongly suggests that a high effect size will be yielded by a specific approach to learning or assessment, but this effect is not realised in practice. This occurred in England with Assessment for Learning. An initial wide-ranging and thorough international research review was undertaken, and this made clear the high potential of a set of practices focusing on formative assessment. In a trial in two education areas, results of the intervention fell dramatically short of the anticipated outcomes. The evaluation of the intervention attributed this shortfall in great part to the very strong influence of specific external accountability measures, which continued to dominate teachers' practices. Some commentators would interpret this as evidence against accountability per se[2]. However, more sophisticated examination of the performance of systems suggests the accountability is an important feature of developed education arrangements[3] and can assume different forms[4]. This suggests that the cause of underperformance in this instance was a lack of alignment between accountability and formative assessment practice. This highlights the issue of complexity of relations between key elements of arrangements and the need for policy to consider and manage such relations.

Secondly, extraordinarily powerful research on the performance of education systems was completed by Bill Schmidt and William Prawat[5]. This used data from TIMSS studies (Trends in International Mathematics and Science Study), a worldwide study which began in 1995 and by 2017 had gone through six cycles. Schmidt and Prawat looked at TIMSS data to explore common features of high-performing jurisdictions. Their work yielded two vital insights. That alignment – between pedagogy, assessment, textbooks, and so on – was essential. They termed this 'curriculum coherence'. So too with appropriate age-related sequencing in subject discipline content, arranged into coherent progressions. This was an additional dimension of their concept of 'curriculum coherence'. The second insight related to 'curriculum control'. Namely, that to obtain 'curriculum coherence', systems needed active policy enactment and constant monitoring: 'curriculum control'. Their use of the term 'control' has caused immediate misunderstanding for some readers. It was assumed – wrongly – that Schmidt was suggesting that 'coherence' can only be obtained through 'top down' control arrangements. However, they make clear in their seminal paper 'Curriculum Coherence and national control of education: – issue or non-issue'[6] that different systems exercise curriculum control through very different patterns of political organisation and public administration.

What Schmidt and Prawat's analysis emphasises is the importance of not only managing the form of specific elements of education arrangements, but also managing the relations between these elements. These relations need to be a deliberate object of policy. Schmidt and Prawat's work on coherence initially focused on the relationship between curriculum aims and content, teaching and learning materials, and teacher practice. We have used their analysis as the basis for a wider consideration of the factors in operation in education systems.

The third and final body of evidence which emphasises the importance of the form of relations in educational arrangements comes from comparison of the highest-performing jurisdictions. In the last three decades, a number of countries have

2 Ravitz, D. (2010) *The death and life of the great American School System*. Basic Books.
3 OECD (2013) *What makes schools successful? Resources, policies and practice*. OECD.
4 CERP (2013) *Examining school accountability*. CERP.
5 Schmidt, W. and Prawat, R. (2006) Curriculum coherence and national control of education: issue or non-issue? *Journal of Curriculum Studies* vol 38 no 6, pp641–658.
6 op. cit.

emerged as outstanding in respect of standards which they achieve (attainment) and the distribution of attainment (equity). What is extraordinary about these different jurisdictions is the extreme differences in the form of arrangements (selective/non-selective, for example), underpinning learning models, and forms of administration – contrast Hong Kong, Shanghai, Finland, Alberta and Massachusetts. These disparities in form yet commonalities in improvement and performance suggest that careful management of the relations between elements of each system are as important as the specific form of assessment, learning, etc in each setting.

5 Complexity and resilience in education arrangements

Transnational and historical comparisons of the performance of national educational arrangements highlight the importance of two key features: complexity and resilience.

Complexity

It is essential to differentiate a *complex system* from a *complicated system*[7]. Complicated systems have many parts and many interactions, but give predictable outcomes. A chronograph is complicated, but gives a highly regulated and consistent output: a measurement of time. By contrast, complex systems possess a large number of interacting components, with outcomes which are not a simple function of the interaction of the parts: 'a complex system is any system featuring a large number of interacting components (agents, processes, etc) whose aggregate activity is nonlinear (not derivable from the summations of the activity of individual components) and typically exhibits hierarchical self-organisation under selective pressures' (Complex Systems Modeling: www.informatics.indiana.edu). Social systems such as education and finance differ from natural systems in a fundamental respect – the operation of a social system is determined in part by the ideas which are held by people within those systems[8] – the behaviour of financial systems is affected by ideas of confidence and risk, the behaviour of education systems is affected by ideas of the value of education, ideas about ability, and so on. Education policy is made *complicated* by the extent to which other aspects of social policy, social development etc impinge on education (welfare policy, health policy, economic policy) but it is made *complex* by the nature of the interactions in and around the system, including the role of aligned and conflicting ideas about education.

This has two extremely important implications for policy makers and those managing educational improvement:

1 Educational improvement cannot be directed towards a static ideal state, but requires constant monitoring, fine-tuning and 'shepherding' in order to secure outcomes such as high equity and high attainment. The Singapore case study

7 Mitleton-Kelly, E. (2003) Ten principles of complexity and enabling infrastructures. In Mitleton-Kelly, E. (Ed) *Complex systems and evolutionary perspectives on organisations: the application of complexity theory to organisations*. Elsevier.

8 Bhaskhar, R. (1998) *The possibility of naturalism*. Third edition. Routledge.

contained in this document highlights how this constant attention to fine-tuning requires clarity of aim and purpose, a focus on evidence of effectiveness, and a careful balancing of the assets of existing and new aspects of arrangements. It is not constant arbitrary 'tinkering'.

2 While attention to the detail of each element of an education system is important, the 'coherence' research suggests that the interaction and alignment of a system should be a deliberate and constant focus of monitoring activity and policy attention – the complex and constant interaction of factors in the system determines the outcomes which it provides.

That we are stating the existence of this complexity could be seen to be a trivial point, except that the complexity so frequently is ignored in efforts to improve the quality of education. Hopes are often pinned on a single initiative, or 'cherry picking' from other systems. For example, out-of-hours study in schools has been shown to enhance attainment, but depends on appropriate premises, availability of supervising or supporting staff, good behaviour management in semi-structured learning settings, suitable tasks to be completed, safe transport home, and so on. A policy initiative such as increasing out-of-hours study in schools needs to be designed in the light of a number of interacting factors – policy formation needs to confront this complexity.

A lack of attention to this complexity stems from the way in which innovation is produced and the practical limits on the possible evaluation of that innovation. Research indeed needs to 'drill down' into specifics, to examine how something causes improvement, not just that it tends to be associated with it. This requires focus – with researchers understandably concentrating on specifics which hold the greatest potential. Funding and practical limits on research and innovation exacerbate this tendency. It is not wrong, since we need forensic analysis of how and why things work. But it is a tendency and, without due care, can result in undue dependence on one-dimensional initiatives.

Avoiding this 'one-dimensional' approach to system improvement is important; one policy approach which seeks to avoid this has emerged in Germany, in response to the OECD's PISA survey. In 2000, its results in PISA came as a 'shock' to Germany[9]. Despite the 'PISA shock' arising from the survey, the German Government treated PISA as a source of data, not as a source of definitive policy solutions. It commissioned key research groups, and in particular groups at the German Institute for International Educational Research (DIPF, founded in 1951), to scrutinise PISA data and OECD's analysis, and link this to wider domestic and international research, only then beginning to strive to understand the causes of poor performance in the system and the potential policy responses to it. The German policy community thus has seen the importance of highly sensitive policy formation which takes into account the detail and complexity of its system, using both international and domestic analyses. The 2000 PISA results caused extensive media scrutiny and widespread social discussion. But the policy response was carefully considered, and evidence based; despite the controversy and concern, policy makers did not rush into premature action. Measures taken – such as increasing access to early years child care, extending the duration of the primary school day (a very significant policy development), national standards-based assessment (every five years in primary schools and six in secondary schools) – all have carefully been monitored, and have had positive impact – attainment has improved in subsequent PISA administrations, as has equity.

9 Waldow, F. (2009) What PISA did and did not do: Germany after the 'PISA-shock'. *European EducationalResearch Journal* vol 8 no 3, pp476–483.

The German experience highlights the subtlety of sophisticated policy formation – putting effort into understanding the interactions in the system, and only then introducing specific and focused improvement measures. Although it is vital to consider educational attainment as the outcomes of a complex system, once context has been considered, highly focused intervention IS possible: for example, the introduction of early reading schemes; diagnostic assessment on transfer from primary into secondary schooling; the introduction of carefully varied practice activities in mathematics.

Complexity should be reflected in our understanding of key terms such as 'curriculum'. Frequently reduced to the idea of 'content', 'curriculum' should be viewed in a far more rich way. Michael Eraut[10] defines 'curriculum' as consisting of:

- aims
- content
- methods (pedagogy and didactics)
- assessment (formative and summative)
- evaluation.

This usefully enables us to capture subtle but powerful relationships, such as that between assessment and curriculum, where assessment usefully can operationally define the depth of treatment of specific topics for pupils of a particular age or stage, and the extent to which assessment can drive curriculum priorities. Hattie's work makes clear the importance of what actually happens in the classroom, and Dan Willingham and ED Hirsch point us to the importance of what an individual child derives from a specific learning experience. This enables us to see the importance of the distinctions which can arise between:

- the intended curriculum (the formal statement of curriculum, whether national or at school level)
- the taught curriculum (what a teacher delivers)
- the learned curriculum (what a pupil derives from the learning experience – learning outcomes)
- the informal curriculum (untaught experiences such as societies, sports teams, etc)
- the unstated curriculum (the ethos, or culture of a school)[11]

and we can extend this to the difference between the National Curriculum (general standards) and the school curriculum (the way in which a school decides to deliver the requirements of the National Curriculum, the way they manage time and priorities, the contexts which they use to explain ideas and so on).

Why is this complexity needed? Because policy can assume that things happen (the intended curriculum) but realities in schools can play out differently – the realistic recognition of the distinction between policy intentions and emerging realities.

The German experience also highlights the importance of analysis preceding action, even in a context where poor performance has created a call for rapid response. Premature

10 Eraut, M. (1997) Curriculum frameworks and assumptions in 14–19 education. *Research in Compulsory Education* vol 2 no 3, pp281–298.

11 Schmidt, W., McKnight, C., Houang, R., Wang, H. C., Wiley, D., Cogan, L., et al (2001) *Why schools matter: A cross-national comparison of curriculum and learning*. Jossey-Bass.

action runs very great risk, since it not only can be an inadequate response to the real causes of poor performance – but also, by being enacted it can affect the system, creating new problems rather than remedying existing ones. Frank Achtenhagen[12] has characterised this as perpetuating a 'cycle of planned failure'.

His model emphasises the following cycle:

Cycle of planned failure

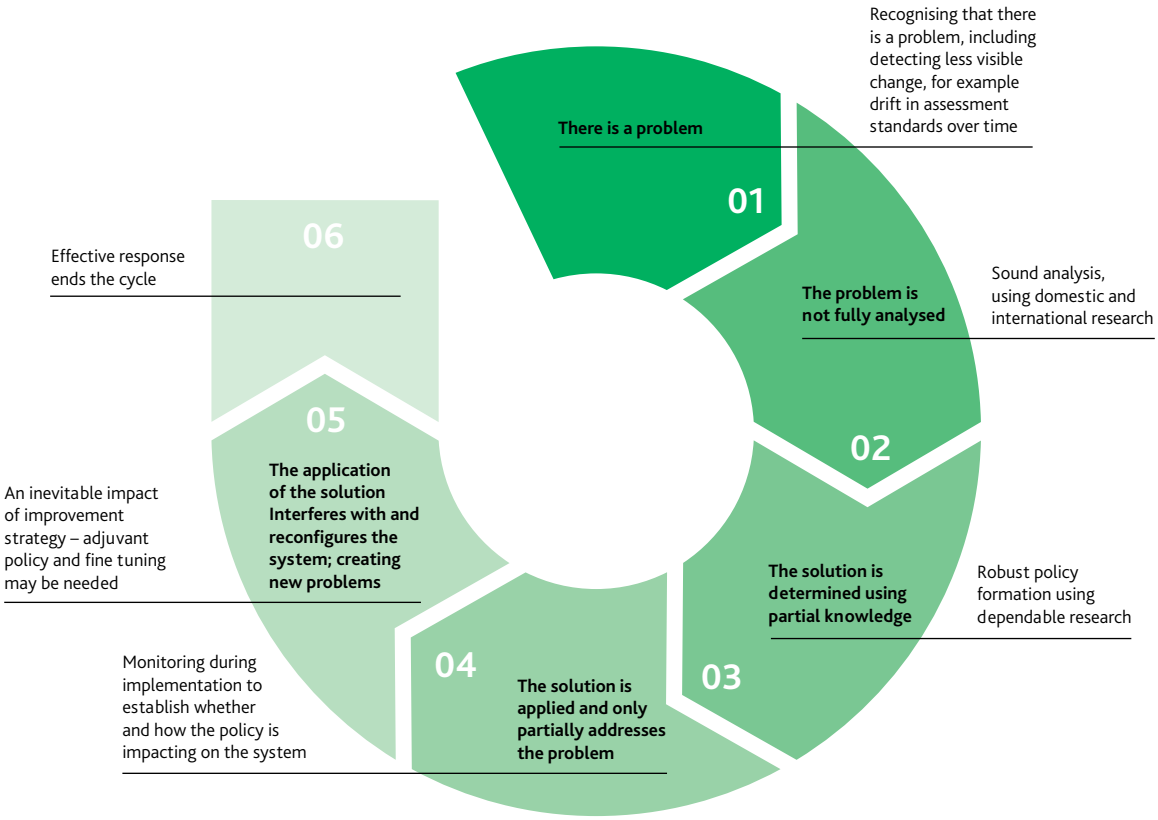
The 'cycle of planned failure' in practice:
An example of this cycle occurred in national qualifications in England. Problems of lack of dependability in examination components relying on school-based assessment were leading to 'grade inflation', decreasing confidence in the examinations, and professional dissatisfaction with the assessments. A government agency was charged with finding a solution to the problem of retaining this kind of assessment (eg practical work in science; writing tasks in English; fieldwork in geography) yet improving its dependability. The model proposed and adopted was 'controlled assessment', which focused on tightening all the conditions of the assessment. However, the analysis failed to take into due account the professional contradictions operating on teachers – who on the one hand were supposed to improve their schools' examination results in conditions of high pressure from accountability measures, and on the other hand operate as wholly objective assessors acting on behalf of an external exam agency. Failing to take this professional contradiction into account, the elaboration of control in the new 'controlled assessments' did not improve dependability. In addition, it created new problems: pupils perceived the new tasks as boring and over-constrained, whilst teachers experienced an increase in workload and a sense of being 'policed', and assessment was viewed as less relevant to active learning. Not only was the original problem not resolved, new

12 Achtenhagen, F. (1994) Presentation to Third International Conference of Learning at Work. Milan, June 1994.

problems were introduced. This cycle was broken only with a far more thorough analysis of the nature of the problem – taking into account professional roles, external pressures etc. This analysis underpinned the introduction of a new, more effective model for examinations at 16 and 18. This was piloted in science in 2015–17[13], with extremely positive outcomes for both assessment and learning.

A key part of breaking the 'cycle of planned failure' is adequate analysis of the nature and cause of poor performance, combined with well-managed implementation, clear communication (which, history tells us, can in certain circumstances include wide-ranging debate and discussion) and effective monitoring:

Cycle of planned failure – breaking the cycle



This highlights the extent to which well-grounded policy formation is not the first nor only step in effective improvement strategy. In addition, highly practical action needs to be taken in respect of designing and managing practical steps to realise the aims of improvement policy – this represents legitimate 'managerialist' focus on *how to get things done*. Whilst Finland's period of improvement certainly involved wide social discussion and careful policy formation, it also included a wide raft of highly practical and well-managed implementation measures: a highly active inspection service to examine how each school was implementing the new pedagogy; national tests to monitor impact; a five-year collaborative process developing a new national curriculum; an intensive staff development programme for all staff in all schools; approved learning materials; and a carefully designed 'roll-out' from north to south, over a five-year period 1972–77[14].

13 Ofqual (2015) *Assessment of practical work in new science GCSEs – summary*. Ofqual.

14 OECD (2010) Finland: slow and steady reform for consistently high results. In *Strong performers and successful reformers in education: lessons from PISA for the United States*. OECD.

The 'control factors' which we present in this document capture the 'areas for action' in respect of improvement strategy – institutional development, professional development, curriculum reform etc. The historical study of educational innovation points to a different balance of emphasis across the factors at any one time in any one jurisdiction. But the analysis also shows an interesting issue: if one specific factor is not deemed as being 'available for use' in innovation, then other factors will need to carry the policy load regarding revisions to arrangements. For example, in using structural reform of schools as an improvement strategy in England, the 2010 Coalition Government removed the requirement for certain classes of State-funded schools to follow the subject-content requirements of the National Curriculum. The factors used to drive schools to teach a 'broad and balanced' curriculum shifted to accountability measures – required combinations of qualifications at 16, and proportions of pupils gaining higher grades in those qualifications. In the system in England, in the absence of 'steering mechanisms' such as approved learning materials, national programmes of staff development associated with reform, and so on, assessment, accountability (particularly school attainment data and national school inspection) carry a very high 'policy load' regarding the aims and objectives of educational improvement[15].

Resilience

The complexity of education systems also gives rise to a further important feature: resilience.

There are two meanings of resilience – 'quickly recovering from difficulty'; and 'the capability of a strained body to recover its shape and size'. It is the second of these to which we refer here. We considered using the term 'resistance to change', but this implies active resistance to improvement policy. Rather, education systems are complex and interconnected, and it is this which tends to make them systemically resistant to change. Many education systems have seen 'initiative' based innovation which addresses a few factors only, and sometimes address only limited aspects of one factor at work in the education system. As stated previously, 'Assessment for Learning' promised much, on the basis of comprehensive research synthesis. But applied in practice in two large educational authorities in England, the anticipated gains failed to materialise. It was an expensive, large-scale intervention, but the researchers concluded that the potential gains were swamped by teachers' concerns regarding assessment and exam grades. The teachers were asked to change their practices: they themselves thought that they had changed their practices, but practices simply moved back into their previous 'shape and size' – systemic resilience.

In the 1980s and 1990s, there emerged in educational research a body of work on 'self-improving schools', emphasising the importance of institutional-level improvement by the concerted actions of teachers in reviewing their own practice, observing one another's practices, and establishing clear shared aims and objectives for improvement. This model of improvement conceptualises the school as the 'unit of improvement'. In one important sense this is consistent with work such as John Hattie's international research synthesis which emphasises the importance to educational quality of what happens in the classroom on a day-to-day, minute-by-minute basis. But schools are

15 Simons, J., Oates, T. and McCulloch, G. (2017) *Nice aims, shame the law's a mess*. Policy Exchange.

not disconnected institutions and do not exist in isolation from the schools which they supply in the next phase of education, or from which they receive pupils: they are connected remotely or directly to higher education, to the economy, and so on. While the idea of the 'self-improving school' usefully has been supportive of institutional and professional development, it ultimately fails to explain the performance of national education systems since, while each school is itself a complex system, the scale of the school does not reproduce all of the interactions and processes which require management or response at the scale of a national education system. 'Self-improving schools' have not added up to significant improvements in national arrangements. The idea does not account for the pressures and drivers which act on schools – the set of important relations in which schools sit. Resilience can originate in the established beliefs and practices of teachers, pupils and parents, but comes also from the network of pressures and drivers in the system.

Policy makers frequently construct carefully considered national priorities for education which reflect social consensus about the desirable aims of education: higher rates of participation; higher attainment in national examinations; higher equity in outcomes; and so on. National aspects of systems, such as school inspection criteria, reflect these aims. National instruments, such as a National Curriculum, seek to provide support and guidance to individual schools – indeed, to individual teachers. It is important to look at the time taken to establish innovation and change in arrangements. Finland provides a model of good practice. Finland began discussions of comprehensive schooling in the late 1940s; with a number of 'false starts' the innovation began at scale in the late 1960s, taking over 15 years to be fully established, and that after radical attention to almost all aspects of the education system and massive implementation effort. The aims and objectives of the reforms were clear and well understood across society. Coherent and effective change was secured only after a substantial period of large-scale, concerted effort. A particularly important part of the Finnish implementation strategy was the focus on professionals' ideas about ability and attainment.

By contrast, many education systems have seen 'initiative' based innovation which addresses a few factors only, and sometimes addresses only limited aspects of one factor at work in the education system. For example, recent Federal attempts at improvement in the USA have focused on standards-based strategy (core standards, tied to accountability measures), with Federal ability to affect more factors (such as the subject knowledge of teachers; pedagogic models; etc) strictly limited by restrictions around the extent to which Federal requirements can be placed on State governance – a perennial issue in the US context. The huge Federal investment in educational reform has not borne immediate fruit; there has been no 'step change' in educational attainment. In the Charter Schools initiative – effort which focuses on improvement through structural change, by the introduction of systematic competition between schools – there are emerging both high-performing, improving Charter schools, and Charter schools whose outcomes have deteriorated since founding. The innovations, though hugely costly, have not yielded universal benefits[16]. Standards were not by themselves enough. Nor was enhanced funding. Nor, in isolation, was increased school competition.

16 Petersen, P. (2016) Post-regulatory school reform. *Harvard Magazine*. Sept–Oct 2016.

These perspectives from Finland and the USA give insights into the scale, scope and duration of improvement policy aimed at substantial system change. Policy formation is complex, yet policy alone is not enough, no matter how well formulated and evidence based it is. Implementation strategy requires the same level of careful thought and grounding in evidence as policy formation, if policy aims are to be realised. The history of educational innovation tells us that educational systems are highly resilient, and thus difficult to change. For example, a specific policy aim or innovation may be highly compelling, well grounded in research, and enjoy wide social and professional consent and support, but drivers and incentives deriving from funding patterns, accountability measures and so on may provide a contrary set of pressures on professionals, diluting the policy aims. Alongside this, professional practice in the classroom, and in school management, relies on heavily internalised sets of practices – these are necessarily automatic and ingrained, so that they are efficient and effective. At one level, professionals may support a new set of practices, but may at the same time operationally lapse to existing practices. Although complex, the existing patterns of operation of an education system tend to be ingrained, mutually reinforcing, and highly persistent. In other words, systems tend to be highly resilient. The development of implementation strategy, commissioning of monitoring processes to detect the impact of innovation, and the management of 'fine tuning' and responsive, adaptive implementation strategy are not to be underestimated.

6 The frequency of curriculum and assessment reform

Reform and transformation of education reduces capacity in the system during the time of change. As teachers and managers work to understand and adopt new working processes, this uses time and resource re-directed from existing practices into new processes. New processes may have distinct advantages and assets, and may address known, persistent problems of existing arrangements, but it is vital not to underestimate the impact of transformation. For example, modelling of possible transitional challenges was not done in a variety of initiatives – such as the implementation of 'levels' in English assessment practice, and the implementation of the reform of A levels in 2000 – and considerable problems arose as a consequence. Cambridge Assessment also has argued that it is vital to ensure that national curricula only change when there is a fundamental shift in foundational knowledge in key disciplines. Research by INCA at NFER suggests that nations change their national curricula, on average, every 10 years[17]. It is important not to reify this figure – it simply is an average of existing 'habits' in curriculum renewal; it does not fit, for example, the frequency of change in fundamental paradigms in key disciplines; which typically have occurred much less often than this. Cambridge has highlighted the error of confusing

17 O'Donnell, S. et al (2010) *Thematic probe: curriculum review in the INCA countries*. May 2010. NFER.

'concepts' and 'contexts' and has argued that a national curriculum should focus on a parsimonious listing of key concepts, principles, fundamental operations and core knowledge – not on the contemporary contexts and settings which teachers might use best in teaching and learning around these concepts – this is the now widely recognised 'national curriculum'–'school curriculum' distinction. We would argue that national curricula have been changed more frequently than necessary; change can be entirely necessary when there are shifts in the content of foundational knowledge in key disciplines – new fundamental discoveries in physics, biology, geography, etc. Change can be necessary when there is a need to address curriculum overload, or correct accumulating problems. Impetus to change may come from research – for example on reading, on sequencing in maths, etc – but this is likely to be discipline specific, and not warrant wholesale curriculum revision across all subjects. Change should not be undertaken without due cause – there can be powerful negative impact of unwarranted change in content in national curricula and national assessments: teachers' carefully accumulated practices and materials can be pushed aside; time and resource has to be allocated to managing change; school leaders' energy can be directed away from necessary management activities, and so on.

Our framework of 'control factors' does not suggest that effective improvement only occurs when policy aims to shift all key aspects of arrangements simultaneously – funding, inspection, etc. We are not asserting that. Rather, we suggest that change and refinements in single aspects of arrangements – in curriculum, in assessment, etc – can be entirely appropriate, but need to be undertaken with awareness of how key elements of arrangements line up and interact.

We argue that change in a national curriculum or national assessment (including examinations) should be relatively infrequent[18], always research based, and its implementation carefully monitored. Cambridge has examined the processes of change over time in various national curricula and frameworks of national standards, and laid down the research-based principles for the revision of the National Curriculum in England. The work indicated that change in discipline content tends to occur within individual subjects, and this further introduces a rationale for always considering whether incremental change is necessary, instead of wholesale change in national curriculum frameworks. The principles also highlight the fact that change in sciences and maths can occasionally affect one another, where one subject demands a foundation of concepts or operations from another. Such change is less disruptive than wholesale, regular change across the whole of the national framework. If the national system is exhibiting wholesale weakness, due to poor design, or accumulated problems deriving from pressures outside the framework, then there may be a case for total framework review. But historically, such total review has been conducted more frequently than genuinely is necessary, with negative consequences for capacity and resource.

Much has been written about 'washback effects' from change in national standards and national qualifications. Standards and assessments are relatively easy to change compared with many other 'control factors', but without careful consideration of other factors, these washback effects can be unpredictable in precise impact, and the transition costs exceptionally high. This endorses our emphasis on effective implementation strategy as well as formulation of well-grounded policy.

18 Oates, T. (2010) *Could do better: using international comparisons to refine the National Curriculum in England.* Cambridge Assessment.

7 Curriculum coherence

Modern analysis of the performance of education systems suggests that 'curriculum coherence' is vital, and is associated with high-performing systems. This is not just a trivial, common-language use of the term 'coherence'. A system is regarded as 'coherent' when the national curriculum content, textbooks, teaching content, pedagogy, assessment and drivers and incentives all are aligned and reinforce one another. '...Curricular materials in high-performing nations focus on fewer topics, but also communicate the expectation that those topics will be taught in a deeper, more profound way...'[19]

We have extended this analysis beyond the alignment of curriculum standards, curriculum materials and teaching, to look at alignment across a large set of dimensions of education arrangements – assessment, funding, professional development, and so on. Deriving from study of the factors which emerge across transnational surveys and research, the framework has proved a powerful mechanism for looking at beneficial coherence in education systems and dysfunctional lack of alignment. Schmidt's work suggests that a level of control must be exercised in a system in order to promote a necessary level of curriculum coherence. Once again, it is vital to recognise that a national curriculum cannot, by itself, guarantee curriculum coherence in the system. As stated above: a system is regarded as 'coherent' when 'factors' are aligned: national curriculum content, textbooks, teaching content, pedagogy, assessment and drivers and incentives all are aligned and reinforce one another. For this to be the case, a certain level of control is necessary. Crucially, Schmidt and Prawat's comparative work suggests that this level of control need not necessarily derive from top-down measures. It is more that the system must exercise control, not that individual agencies always should take control:

> **'...our purpose in introducing alternative ways to govern curriculum...is not to advocate one approach or another. As analysis by Cochran-Smith and Fries (2001) indicates, disagreements about teaching and, by implication, curriculum, often divides along ideological lines, an outcome that occurs no matter how pragmatic the veneer. A functional approach, by specifying in advance the criteria that an effective curriculum-governance system must meet, lessens the tendency to judge these systems in terms of the political values they represent (eg regulation vs deregulation, public interest vs private interest...'.**[20]

19 Schmidt, W. and Prawat, R. (2006) Curriculum control and national control of education: issue or non-issue? *Journal of Curriculum Studies* vol 38 no 6, pp641–658.

20 Ibid, p656.

Their analysis suggests that while the existence of curriculum coherence through curriculum control is essential, the precise institutional and system form to achieve this can vary from one jurisdiction to another. The evidence for this is the extent to which different jurisdictions have improved their systems using different relationships between central government and schools; have consulted with and involved education interests in very different ways; and have used contrasting mixes of different forms of 'restriction' – for example, Finland placing great emphasis on teacher training in establishing adherence to curriculum goals (an emphasis on front-loaded control); other nations using on-going accountability measures and assessment (an emphasis on 'end-point' control).

While the 'control factors' analysis is informed and underpinned by Schmidt and Prawat's work on 'curriculum coherence' and 'curriculum control', the identification of specific factors was undertaken through analysis of literature on transnational comparison of the performance of education systems. Key analyses were identified using criteria related to citation and prestige, and explanatory and causal power. The number of discrete and interlinked factors were then identified in texts such as Green's 'Education and State Formation'[21], Alexander's 'Culture and Pedagogy'[22] and Raffe's 'Policy learning from "home international" comparisons'[23]. When collated and organised, these factors were allocated to one of two categories: 'control factors' (those most amenable to policy action) and 'explanatory factors' (those which condition the context of that policy but which are distinctly resistant to direct action in educational policy). The listings were tested and refined through discussions with researchers, opinion-formers in education, policy makers and teachers. As with the background theory on 'curriculum control', the analysis of 'control factors' does not imply a specific form of political organisation, a specific form of construction of policy, nor specific models of enactment, implementation or evaluation. As with 'curriculum control' the word 'control' in 'control factors' does not imply or presuppose 'top down' models of policy formation and management.

8 Control factors

Please note that the analysis asserts strongly the interaction of these factors in educational arrangements. We cannot capture all the interactions in a given system in a simple diagram. So here we present the control factors in the form of a table. There naturally is overlap between categories – for example, formative assessment in the form of rich questions asked in the classroom can be an intrinsic aspect of deliberate pedagogy. We therefore have not pursued 'perfect separation' of the content. Nonetheless we have developed the listing to be helpful – a valuable heuristic for understanding, for policy formation and for system management.

21 Green, A. (2013) *Education and state formation*. Second edition. Palgrave.
22 Alexander, R. (2001) *Culture and pedagogy*. Blackwell.
23 Raffe, D. and Byrne, D. (2005) *Policy learning from 'home international' comparisons*. CES Briefing.

Control factors

1 curriculum content	national standards; curriculum 'frameworks'; aims statements; subject specifications; textbooks; schemes of work; support materials; subject discipline models/domain specifications; subject sequencing; subject combination criteria/baccalaureate models
2 pedagogy	teaching and learning approaches; implicit and explicit theory driving teaching and learning; didactics; models of ability; models of progression; setting and streaming; classroom culture; homework and practice models
3 assessment and qualifications	summative assessment; formative assessment; diagnostic assessment; assessed elements of the curriculum versus non-assessed; assessment and measurement models including principles on the use of data and information from assessment; teacher assessment and external assessment; sample-based measurement and related approaches to measuring national standards
4 institutional development	leadership; management models; institutional policy formation; teacher allocation; programme and institutional evaluation; lesson observation and methods for identification and dissemination of good practice
5 institutional forms and structures	size of schools; school type; phase/age range; class size; building forms; facilities; institutional specialisms; services available (social, health; etc); inter-institutional collaboration/competition
6 governance	national control arrangements; inter-departmental collaboration at government level; inter-agency collaboration; student/pupil allocation arrangements; governance tiers and powers; governance composition and membership; institutional status
7 professional development	professional roles and responsibilities; teacher selection, training and preparation; continuing professional development and support; unionisation and association; professional progression; remuneration; performance measurement
8 accountability	accountability model (teacher level; school level; national level); targets, goals and criteria; data collection processes; publication of data; advice and guidance on interpretation of data; analysis, interpretation, consequences, sanctions and action
9 inspection	framework for inspection; publication and reporting arrangements; composition and competence of inspection profession; governance; frequency of inspections
10 funding	levels and patterns; funding sources; allocation; financial control and financial indicators; links to accountability; funding of development projects and innovations; evaluation and review
11 national framework	legal attendance requirements; routes in education and training arrangements; route allocation points; route flows; route transfer arrangements/flexibility; relative status of different routes: • routes in arrangements – academic, vocational; allocation to different school types; etc • route allocation points – pre-school to primary; primary to secondary; etc • route flows – the number and type of pupils on each route
12 selection and gatekeeping	methods of allocation to routes; entry requirements; governance of entry requirements; qualifications equivalence specifications and rules
13 information and guidance about routes and choices	focus, level and detail of information and guidance; entitlement to guidance services; guidance professionals' responsibilities and their means of updating on labour market etc; links between schools and destinations in education, training and labour market
14 allied social measures	development of incentives through fiscal policy; family support; regional development support; educationally related service provision (health etc); incentives and washback effects from labour market policy

9 Explanatory factors

If 'control factors' describe things with which education policy directly can engage, 'explanatory factors' describe things which tend to be out of scope to deliberate policy or with which education policy does not directly engage. For example, historical legacy in terms of societal commitment to education can profoundly affect the way in which education arrangements operate – affecting the status of teachers, parental and pupil attitudes to education, etc – but is a 'given', not something amenable to policy. Whilst it can be supported, adversely or beneficially affected by current educational policy, the historical legacy affects the 'zeitgeist' – the base of educational commitment. In Finland, while teachers are, and have been, paid approximately the OECD average, the societal status of teachers was consolidated by their role in the processes establishing Finnish independence. In Hong Kong, a strong, traditional parental commitment to education impacts significantly on the patterns of learning outside school contact time.

Likewise, there are powerful legacy issues from the state of economic development of a specific jurisdiction – Singapore development policy is heavily predicated on the importance of human capital, given the 'city state' nature of the jurisdiction and the absence of high levels of natural resources.

In other jurisdictions, contingent factors – chance – affect the possession of assets such as oil or mineral wealth, which drives patterns of expenditure in education provision. Downturns in globally derived revenues from such assets can cause hot or cold winds to blow through education expenditure, on a cycle which is not in the control of the jurisdiction itself.

Explanatory factors

1 **global economy**	trade patterns; upturn/downturn; labour movements; regulation; exchange rates
2 **domestic economy**	fiscal and monetary policy; upturn/downturn; industrial policy; inward investment; labour market policy; higher education research strategy; migration/immigration policy
3 **culture**	social structure; family culture; societal attitudes to education; gender identities and life expectations; signalling into schools from society and the economy
4 **political structures and commitments**	status of educational policy in domestic policy apparatus; nature of social consents
5 **historical contingencies**	natural resources; distribution of wealth and opportunity; regional political and territorial relations/conflict; domestic security and conflict; political developments impacting on education (eg post-conflict reconstruction; independence movements; demographic trends
6 **natural environment**	natural disasters and calamitous events (earthquake; fire; flood); climate change

The explanatory factors which we include here are part 'historical legacy' – the legacy which affects the context in which policy makers must form education policy – and part contemporary events outside education (world economic events, natural disasters) which can impact on education systems. They are amenable to policy responses – national culture can and does change, responses to natural disasters can relieve their effects – but they more frequently are the context in which education policy must be formed than factors which can be changed readily through education policy itself.

For example, history has played its part in the social status of teachers in Finland. Teachers are highly respected in Finland due in part to the highly selective nature of admission to teacher training and the high level of qualification obtained in the training (note that they are paid around the OECD average and paid less than in the US)[24] but also because of things which occurred in the past: the teacher as a pivotal member of small rural communities; the role they played in establishing Finnish national identity during the journey to independence, and asserting Finnish language over Russian. For countries struggling to increase the status of teaching, this fuel of status is not available – it is a 'given' in Finland, an explanatory factor – but it is not available for policy makers in other nations – it is not a control factor which they can utilise in policy formation and in actions leading to educational improvement.

The explanatory factors are important, and they explain the shape and operation of education arrangements in a specific country setting. They help us to understand and explain the operation of arrangements. They need to be taken into account; but in many instances they are 'givens' which need to be accommodated in education policy formation, a hard reality.

Some nations link education policy to other areas of public policy – health policy, social policy, economic development strategy – in more sophisticated ways than other nations. In such instances – such as Singapore's careful link between skill supply and sectoral economic development – greater coherence across some aspects of explanatory factors and control factors in education has been achieved.

10 The importance of 'culture' – and the possibility of making 'culture' an object of educational policy

Transnational analyses frequently emphasise the role which culture plays in determining the form and performance of a national education system (Robin Alexander, others). It is entirely right that it is not ignored. Societal attitude to the role of education; youth attitudes to learning; parental commitment; signals from the labour market; historical

24 NCEE http://ncee.org/what-we-do/center-on-international-education-benchmarking/top-performingcountries/ finland-overview/finland-teacher-and-principal-quality. Accessed 01/06/17.

events associated with education; all impact on education – from behaviour in the classroom to learning in the home. With South Korean pupils stating that '…the future of Korea rests on my shoulders alongside my own future…', and Finnish 40 year olds speaking of the great respect they had for their teachers, it is clear that culture resides in and grossly affects pupil motivation and engagement, pupil behaviour, pupil decision-making, and so on. In turn this plays a role in class size, teacher workloads, institutional forms, routes in the system and other key features of national arrangements.

Putative shifts in youth culture have been identified as one of the factors at play in the decline of standards in the Finnish system since 2000, and constructing a curriculum response to this has played a role in the development of pedagogic approaches intended to arrest the decline.

Differences in ideas about ability and human learning – 'models of ability and progression' – are evident in different systems: contrast 'Confucian' ideas with models associated with 'liberal individualism'. Although it is important to remain sensitive to the subtle variation in learning models within national settings as well as variations between them, it is still the case that there are important, fundamental variations in assumptions and ideas about learners and learning[25].

Confucian-derived models are characterised by 'every child capable of learning anything, depending on how it is presented to them, and the effort which they put into learning it…' while 'individualist' models are characterised by different sets of assumptions focusing on 'ability', 'individual rights' and 'identity formation'. The key issue for the analysis here is that these differences are not only detectable in the values and ideas held by educators, pupils, parents and others, but give rise to very different learner behaviours and educational practices. Key aspects of education vary in different national settings as a result of these contrasting underpinning systems of ideas[26].

However, although such ideas have a concrete impact on and expression in day-to-day practices, culture is mutable – it is capable of change. While some commentators have argued that differences in culture prevent the ready migration and export of specific educational practices (a legitimate reaction against naïve 'policy borrowing'), this does not mean that educational policy in a specific country should ignore the opportunity to change 'learning culture' or to impact on youth culture. The fact that 'learning culture' can be the object of policy within a school relates to the qualities of the school as an institution (with 'rules', an 'ethos' and so on); the qualities of the school as a 'community' (which, whilst not hermetically sealed from wider community pressures and connections, is relatively 'self-contained'); and the capacity for the school to determine policies and professional development which can adjust the ideas about ability, progression and so on, which are held by its education professionals.

The 2010–13 revisions to the National Curriculum in England explicitly confronted the implicit models of 'ability' and 'progression' which had built up in the education system 5–16[27]. Specific expressions of concepts such as 'differentiation' and 'individualised learning' were challenged, and a 'levels-based' system of national

25 Crehan, L. (2017) *Cleverlands*. Unbound.

26 Hirsch, E. D. (2015) *Equity, attainment and core curriculum*. Lecture, Cambridge, 21/09/15.

27 Oates, T. (2010) *Could do better: using international comparisons to refine the National Curriculum in England*. Cambridge Assessment.

assessment removed from being a requirement of schools. In the place of these, ideas aligned to Confucian models of ability and progression were promoted, alongside specific models of 'pace' and 'mastery' – these being designed to increase both attainment and equity. National inspection was aligned to the revised model of ability and progression, and certain learning resources became eligible for funding on the basis of conformance to this revised model (specific learning materials to support a national system of schools developing maths education based on principles and practices extant in Singapore and Shanghai).

Further evidence that 'culture' in schools – and specifically in the classroom – can be the object of policy – and thus susceptible of innovation and reform – comes from the period of intensive improvement in Finland. The reforms of the 1980s were based on a complete re-alignment of the system to fully comprehensive education with new models of 'ability', 'inclusive education' and 'progression', away from a tracked system with different models of ability, differentiation and progression. A discussion regarding revised aims and models in education took place in the media and in society at large, and preceded the practical changes in schools. A massive programme of professional development was put in place, and intensive inspection in the first years of the changes was designed to scrutinise teachers' commitments to the new ideas, as well as examine practical implementation of requirements. Ideas and culture matter. But whilst being influential, they are not a 'given' which cannot be influenced – approached in the right way, they can be the object of deliberate change policy.

11 Case studies of educational improvement

These short descriptions of four nations are not intended to be comprehensive descriptions of the causes of high performance, or a full analysis of policy responses. Instead, they are designed to illustrate the validity and importance of seeing systems through a 'control factors' lens, in order to enhance the effectiveness of co-ordinated policy.

Again, it is important to emphasise that 'curriculum coherence' is not obtained through simple patterns of 'top down' control, but through subtle forms of interplay between social consents, professional practice, co-ordinated central action, and formal restrictions such as legal measures. The cases show the importance of constant monitoring and 'fine tuning' of policy and action around the balance of different factors in each system, combined with periods of carefully managed larger scale reform of certain aspects of arrangements. Notably, our analysis differs from some others in that it emphasises the extent to which successful systems going through periods of successful change have discriminated between those things which require innovation and change, contrasted with aspects of arrangements which should be kept stable, being either essential to quality, or providing important 'reference' and 'anchor' points in arrangements.

Nor do the descriptions assume there is an 'ideal type' of progress – a set pattern of development through which nations go. There is a poor correlation between State

expenditure and system performance, which suggests that bad decisions by policy makers are just as possible as good ones. Economic and social conditions are subject to constant change, necessitating constant policy decisions and practical responses. These decisions on 'fine-tuning' and re-direction do appear to matter – transnational comparisons over time show that education arrangements do not follow 'iron laws' of economic determinism, nor are they immune from being adversely affected by events outside them. Without due care; even high-performing systems can deteriorate. In the case of Finland, while economic development may have stimulated social and individual aspirations and thus increased the demand for quality schooling, it was not inevitable that performance would rise. The initial evaluation through national sample assessments showed that decisions during the 1970s and 1980s regarding educational policy and practice were coincident with a substantial rise in equity and attainment. Looking back, the decisions to adopt comprehensive schooling and implement it in a specific way appear sound. But even as scores continued to increase to their peak in 2000 PISA, it appears that certain features of earlier arrangements in Finland may have been neglected and/or that changes in youth culture may have been overtaking the system[28]. Asserting that 'Finland seems to have made the right decisions in the 1970s and 1980s' contains a logical commitment to the idea that 'the alternatives may have been worse' – in other words, the nation could have been subject to bad decision-making – by policy makers, by school principals, and so on.

Examining 'what might have been' is methodologically challenging in historical analysis[29]. However, transnational comparison does help us gain some insights into 'what might have been' regarding decisions taken. The issue of selective schooling produces some useful insights. Around the globe there are examples of periods of improvement and high performance in non-selective, comprehensive systems (Finland, Japan, Shanghai, Massachusetts) just as there are periods of improvement and high performance in selective systems (Netherlands, Switzerland, Singapore). Likewise, there is a wide range of examples of poor-performing selective and non-selective systems around the globe[30]. This appears to be a confusing mess of contrary directions and conflicting evidence, but the core propositions of this document help us make some sense of it. Selective systems appear to have enhanced overall educational attainment, only when certain conditions are in place, such as avoidance of poorer quality pedagogy or inadequate resource allocation in some educational 'tracks'. Despite their different forms, the successful systems possess coherence and alignment, and are carefully monitored in respect of attainment and equity. Analysis through the lens of 'control factors' and 'curriculum control (alignment)' point to the importance of careful management of the alignment of the different elements of each system, effected through constant monitoring and 'fine tuning' – and abandonment or revision of those innovations and changes which initially appeared to hold promise but which monitoring suggests are not working.

This highlights two essential observations: firstly, that while contextual and contingent issues impinge on education, decisions and action matter – there is no inevitability in educational development, no set pattern of improvement. Secondly, that evidence-based policy formation is necessary but not sufficient; active implementation to secure alignment of key elements of arrangements, and a high level of monitoring and self-critical fine tuning is essential.

---

28 Heller Sahlgren, G. and Oates, T. (2015) *Real Finnish Lessons: the true story of an education miracle*. Lecture, Cambridge, 09/11/15.

29 Bunzl, M. (2004) Counterfactual history: a user's guide. *American Historical Review* vol 109 no 3, pp845–858.

30 OECD (2012) *Education at a glance 2012*. OECD.

Finland

Innovation overview

Population 5.5 million[31]

Major innovation in the system following 1968 Act and widespread social discussion on education 1950s–60s.

Increase in system equity and attainment 1970s onwards, to peak in 2000; then decline.

New national curriculum in 2014, implemented in 2016.

Finland's high score in the first PISA survey in 2000 attracted global attention, and the form of the education system in 2002 onwards became a focus of analysis and commentary. But the Finnish system measured in 2000 arose through reform activity enacted over 40 years earlier. Study of the history of reform in Finland is clear: looking at the shape of the school system in 2000 gives an entirely misleading picture of the way in which control factors were managed during the time of dramatic and sustained improvement in Finland. Many current analyses of 'Finnish success' describe the system as being one where 'high autonomy' is experienced by schools and education professionals. The analyses justify this label on the basis of high levels of intensive professional training for all teachers, the absence of blanket national assessment other than school-leaving examinations, the absence of high-frequency school-visit-based inspection, and lack of published league tables. 'High autonomy' is presented as a *cause* of high performance. However, such analyses are seriously misleading on two fundamental counts – they both misrepresent the period of improvement and reform in Finland, and they misrepresent the current situation in the country[32].

Firstly, the period of rapid and sustained improvement in both attainment and equity occurred during the 1970s and 1980s, culminating in a peak which was caught in PISA 2000 – triggering the intense international interest in Finland. Children assessed in PISA 2000 were 15 years old. The system peaked in 2000, and scores have declined since. This points interest to the system as it was when these children started their education – the 1990s. But it is unlikely that the system suddenly fell into perfect shape in the year they started their education, so it is important to examine the system of teacher training, school governance etc which was in place in the 1970s and 1980s. To understand *educational reform activity*, attention should be directed at how the system was configured and managed during its time of improvement, not to the time of its 'peak performance' in 2000. The fact that scores have declined since 2000 reinforces the concern that looking at the system in 2000 is extremely misleading[33].

A more sensitive examination of the historical trajectory is needed. The 1968 Act replacing selective education and establishing a fully comprehensive system followed nearly two decades of wide social and political debate. Two commissions – 1945 and '46 – explored the shape of comprehensive education, but vocal opposition from universities and grammar school teachers played a key role in preventing Government from creating sufficient social consensus to move on overarching system reform. In the 1950s, the issues were revisited, and further wide-ranging public debate – which included discussion

31 Statistics Finland, 2015.

32 Oates, T. (2014) The qualifications sledgehammer: why assessment-led reform has dominated the education Landscape. In Heller Sahlgren, G. (Ed) *Tests worth teaching to*. The Centre for Market Reform of Education; Heller Sahlgren, G. (2015) *Real Finnish lessons*. Centre for Policy Studies.

33 OECD (2014) *PISA 2012 Results in Focus*. OECD.

of what kinds of schools were needed to support Finland's transformation to a more modern economy – tipped the balance towards fundamental change. This commitment was underpinned by the 1968 Act, with a recognition that while the majority of the population were wholly committed to the fundamental reform, not all were.

There are two important issues which need to be highlighted. Firstly, the reforms were not solely structural and administrative. They included a fundamental shift in ***ideas*** about education – namely, that all children were capable of high attainment, and that both individual and social good would derive from a radical shift of pedagogic practice and pastoral support in schools, as well as changes in the institutional forms of schooling and 'routes' through the system[34].

Secondly, some of the conditions which supported the transformation of the system were in place prior to the reforms. The idea of a national school system with common curriculum objectives had been established with the 'folk schools' development in the late 19th century. The National Curriculum following the 1968 Act was not the first common framework, but the latest in a long line. Alongside this, consolidated and underpinned by Education Acts in the 1920s, teacher training had become highly regulated, with a very high level of demand on the teachers regarding subject knowledge and pedagogic approaches. Gabriel Heller Sahlgren's incisive analysis of the chronology of improvement[35] shows that improvements in student attainment started ***before*** the establishment of the fully comprehensive system in the 1970s, and emphasises that general improvements in society and economy may have been highly instrumental to the gains seen in education – Finnish society was engaged in conscious development of human capital, continued development of national identity after Russian occupation, and a move away from a mostly agrarian economy. But this does not make irrelevant the decisions which they took on the specifics of education policy – to move the system from a relatively moribund selective system to a high-expectation comprehensive system. That is, the decisions which they took to establish national standards, to strengthen teacher training, and so on. When we look at a range of nations, there are societies which also have been through periods of economic improvement – the USA, Germany, for example – which have not enjoyed the rapid rise of attainment experienced by Finland. This suggests that, although not solely the result of educational policy and reform established from 1968 onwards, the increase in attainment was well supported by the rational transformation of educational arrangements.

By looking at the increase in attainment from the 1960s onwards, and the decline from 2000, we can see that it is vital to look at what the system looked like, and what improvement strategy was in place, during the three decades before 2000. And the system indeed looked very different. Far from 'high autonomy', Finnish historians of education emphasise that the reform was driven from the centre, with tight control over roll-out and arrangements. This is perhaps entirely unsurprising in a country familiar with Soviet-style administration of State systems. The 'control factors' framework helps identify the different components of implementation which were used: new national standards (National Curriculum); frequent and detailed school inspection to ensure that teachers were abiding by the principles of fully comprehensive education; grade tests in each year of schooling, on a sample basis, to monitor the impact of the reforms; State-approved textbooks, with content and didactics aligned to the National Curriculum

34 OECD (2010) Finland: slow and steady reform for consistently high results. In *Strong performers and successful reformers in education: Lessons from PISA for the United States*. OECD.

35 Heller Sahlgren, G. (2015) *Real Finnish lessons*. Centre for Policy Studies.

and to comprehensive aims; elaborate and extensive data submission from schools to local governance and the State. In addition, these various measures and requirements were highly aligned; the advantageous 'curriculum coherence' identified by Schmidt and Prawat as being characteristic of high-performing systems.

These factors continue to be referred to in Finland as the 'steering mechanisms' of education.

This is an entirely different picture from that of 'high autonomy'. Restriction and control, following social debate and agreement regarding the aims of education, characterised the phase of adoption and implementation of practices which appear to be linked to the substantial rise in performance in Finland. With this more accurate analysis of the trajectory of improvement, we can see that the phase characterised by 'high autonomy' – the late 1990s – is perhaps an ***outcome*** of the rise in performance of the system, not the ***cause*** of the fundamental improvement in the system.

But there is additional analysis which further calls the 'high autonomy' characterisation of the current Finnish system into question. Many of those visiting the country and exploring the characteristics of arrangements appear to have suffered from observation bias – looking only at those aspects of the system in which they have prior interest. The resulting accounts focus on the acknowledged absence in the Finnish context of forms of restriction which are present in some other nations: frequent and imposing national school inspection; published school performance data; national pupil assessments; and so on. This kind of analysis fails to recognise the forms of restriction which ***do*** exist in Finland – more penetrating analysis suggest that this restriction is extensive and influential. The extended and detailed National Curriculum specifies the number of hours which should be devoted to particular subjects, something which would be regarded in some systems as a gross imposition on schools. Teacher training places very tight 'front-end restriction' on pedagogy, strongly steering teachers to specific professional practices, lessening the need for 'end-point restriction' in the form of targets and published data. Teachers feel that it is a requirement to administer a relatively high density of formal tests in primary schooling, in order to identify children who are at risk of falling behind. Schools are required to submit data to local and national educational governance, with the knowledge that poor trend data will trigger inspection and scrutiny. The 'control factors' analysis helps to reveal and illuminate the various forms of restriction which more superficial analyses have failed to recognise.

One interesting feature of the transformation of the Finnish system has been the stability of one key aspect of arrangements – summative assessment. In the early phases of the reforms, new national tests were introduced – grade tests assessing a national sample of pupils – in order to examine the impact of the reforms. These were administered by a university team on behalf of Government and withdrawn when it was felt that reform was yielding the expected gains. While increased levels of formal testing in primary education were established and persist, this is focused on ensuring that pupils at risk of falling behind are readily and accurately identified. But it is important to note that the high-stakes assessment at the end of higher level secondary education (Abitur) remained fundamentally unchanged throughout the reforms. As with the sample-based grade testing, the upper secondary assessment was held stable as a 'fixed point' against which to evaluate the impact of changes. Rather than use high-stakes assessment to drive reform – a tendency in a number of other countries – Finland held stable a key final assessment – a notable strategic decision.

Our analysis shows that Finnish educational policy associated with the period of transformation and elevated attainment carefully harnessed a wide range of key control factors. It was based on coherent and systematic management of these factors, and was committed to ensuring that the new ideas of pupils' ability and potential were clearly understood by society and educational professionals, and that all practices corresponded to these ideas: elevated attainment and elevated equity, with all children carefully supported. In addition, even with its commitment to tight management of the transformation of the system, the State did not assume that its massive reform programme would be simple and easy – it put in place intensive monitoring processes to allow early warning of emerging problems, and the facility for 'fine-tuning' policy and implementation strategy, and for reviewing overall progress.

Where now?

Finland has experienced a long programme of closure of small rural schools[36], rising social concern at variation in school performance in urban areas[37] and a post-2000 slide in performance in international surveys. The principal response has been to examine curriculum balance, recognising that shifts in youth culture have had an impact on older children's attitude to learning. It has introduced an entitlement for all pupils to at least one 'multidisciplinary learning module' each year, with initial implementation in the urban capital. In pre-school, projects on 'playful learning' have been implemented, which explore not only the nature of pupils' engagement with learning but the way in which new technology can be incorporated into pedagogy. A new national curriculum was drawn up in 2014, implemented in 2016. The State continues to stipulate by decree the distribution in schools of lesson hours per subject area.

What remains of continuing interest to outside observers is the fact that policy and change management models being used in this period of innovation appear different in so many respects from the policy and change model used during the earlier period of substantial, sustained improvement in Finland.

Singapore

Innovation overview

Population 5.6 million[38]

1965 onwards: creation of single national education system; 1970s: successive measures to enhance quality including tighter central specification of curriculum content for routes in the system in 1979, with the Curriculum Development Institute of Singapore established to develop high-quality materials. Singapore is characterised by continuous revisions to its system, including inspection arrangements.

Rapid increases in participation through system reform; sustained improvements in quality in all routes; continuing very high standing in PISA surveys.

36 Autti, O. and Hyry-Beihammer, E. K. (2014) School closures in rural Finnish communities. *Journal of Research in Rural Education* vol 29 no 1.

37 Berisha, A. K. and Seppanen, P. (2017) Pupil selection segments urban comprehensive schooling in Finland: composition of school classes in pupils' school performance, gender, and ethnicity. *Scandinavian Journal of Educational Research* vol 61 no 2, pp240–254.

38 Department of Statistics Singapore, 2016.

Initial steps to improvement of education in Singapore started with the 1947 Ten Year Programme for Education, and accelerated following independence in 1965. During the 1950s and 1960s, educational development was strongly linked to economic development, a pattern continuing to the present day. Deliberate integration of educational, economic and social policy has been a conspicuous feature of educational reform policy in Singapore and represents conscious pursuit of 'coherence' across control factors. This approach has been especially evident in respect of elevated participation, enhanced quality, and improved attainment in post-compulsory vocational education, particularly in the last decade. Comparativist comment has used the term 'authoritarian' to describe the nature of control over curriculum which is exercised in Singapore[39]. However, such views grossly neglect social and historical context, misrepresent the subtlety of the consents by which Government has persisted in Singapore, and ignore the patterns of development and control which government has deployed in respect of educational improvements. These processes may appear highly 'top down' to some external observers, but involve very interesting patterns of interplay between schools, society and the State. This plays out very publicly – for example, through press coverage of the Ministry of Education's 'annual reshuffle' of school principals, constantly reinforcing the State's commitment to improving educational quality and securing social consent for the shape and development of educational arrangements.

Educational improvement in Singapore has been characterised by a number of key features:

1 Sophisticated management of different, but linked areas of public policy: educational strategy, economic development policy, social policy. For example, Government has attracted inward investment by brokering supply of private capital to stimulate economic development in specific sectors, and then has committed to ensuring skills supply for those sectors through the education system[40]. Similar commitments to 'coherence' can be seen within specific elements of educational reform – this is explored below in detail in respect of 'Singapore Maths'. This deliberate management of 'coherence' across control factors is an important feature of Government innovation strategy.

2 Global observation, constant innovation, combined with close evaluation[41]. The pace and level of innovation and experiment in Singapore is extremely high. These innovations frequently are developed following careful observation of developments in systems around the world, and the emerging findings of surveys such as PISA and TIMSS. There is a very strong commitment to ongoing evaluation of initiatives, with a capacity on the part of Government to resist 'political over-investment' in specific initiatives – retaining the capacity to halt or withdraw things which demonstrably are not working. This has been evidenced in recent development work such as that on e-learning versions of established curriculum materials.

3 As in Finland, Government in Singapore has maintained stability in the form and content of high-stakes assessment at 16 and 18, recognising the important role which standards play in 'signalling' in the system and in providing benchmarks to measure the impact of innovations in areas such as pedagogy, grouping in schools and so on.

---

39 Alexander, R. (2011) Could do even better? Making the most of international comparisons as a tool of policy. http://www.robinalexander.org.uk/wp-content/uploads/2012/05/Could-do-even-better-2.pdf. Accessed 01/06/17.

40 Huff, W. G. (1995) What is the Singapore model of economic development? *Cambridge Journal of Economics* vol 19 no 6, pp735–759.

41 Tan, J. and Gopinathan, S. (2000) Education reform in Singapore: towards greater creativity and innovation? *NIRA Review* vol 7 no 3.

4 High levels of interaction and 'interplay' between the central bureaucracy and schools. This is an important feature of the system, with central appointment of school principals, and with officials in the Ministry of Education and allied agencies having close experience of teaching and management. The rotation of principals carries complex functions: dissemination of good practice; aligning school leaders' aims and objectives with nationally stated curriculum objectives; signalling the State's commitment to specific communities regarding educational quality.

These approaches make Singapore a very interesting case in respect of educational improvement. Although the development of the system frequently is described in terms of 'phases'[42], it is characterised by continual trialling, evaluation and innovation. Pressures on teachers and school principals is recognised to be high by both the profession and officials, but that pressure does not derive wholly from high expectations from the administration, but from high societal expectations regarding the importance of educational outcomes for all pupils.

The approach to reform used in Singapore is exemplified by the development of 'Singapore Maths'; an approach which is now of global interest, with the textbooks and associated professional development being implemented in a number of settings outside Singapore. The management of the department anticipated Schmidt's analysis of 'curriculum coherence' by setting out to ensure a tight relationship between the content of the maths curriculum, curriculum materials, and pedagogic approaches. Following extensive use of imported maths textbooks, in the early 1980s – under the auspices of the newly founded CDIS (Curriculum Development Institute of Singapore) – a group of teachers and officials were nominated to travel to other nations, including Japan and the USA, to examine approaches to the teaching of mathematics. This was not viewed just as a pragmatic shift in didactics, but as a reform which required full theorisation of approach, including consideration of psychological assumptions, age sequencing, concept specification and classification, instructional theory, and sophisticated roll-out and implementation. Based on their transnational comparative work, the team drew heavily on Jerome Bruner's model of pupils moving through specific phases of conceptualisation and consolidation, although it is important to note that wide-ranging review across many areas of pedagogic theory was undertaken. What emerged from this intensive development work was a fully articulated model of maths didactics, with a new generation of textbooks being considered to be an important focus of the innovation. But it is naïve to consider it solely as an innovation focused on textbooks alone. New aims were set for maths education by the Ministry of Education, and widely disseminated. Initially developing State-published materials, from 2001 policy moved to more competitive provision of primary textbooks, with the aim of maintaining quality but reducing cost and allowing choice to be exercised by schools. A specific relationship was set up with a range of commercial publishers, where the State produced guidelines for textbooks. These were implemented rapidly by publishers, and an approval process using the criteria was run by the State. The criteria ensured that textbooks were highly consistent with the aims, theory and practices of the new model, but allowed some variation between publishers' formats and structures. Crucially, it was seen as essential that intensive staff development should accompany the materials, to ensure that all teachers understood the underpinning theory associated with the new model. The 'messages' around maths education included not only specific approaches to the presentation of maths concepts, but emphasis on the importance of maths as a

42 1947 Ten Year Programme; 'survival-driven' phase 1959; the 'Five Year Plan'; 1997 'Thinking schools, learning nation'.

language, enjoyment and positive attitudes towards maths, and an understanding of its economic value. TIMSS data reveal an important balance of pupil outcomes: pupils are not only extremely good at maths, but they express high enjoyment; something not typical of all high-performing systems.

The explicit and intensive focus on the pedagogic model and on attitudes to maths has remained in place in teacher training and continuing professional development, rather than simply being seen as a short-term 'transitionary' need. The approach attained a substantial rise in scores in international surveys (Second International Science Study (SSIS) and then TIMSS) and a continuing high score in PISA. With a significantly increased underlying score, Singapore moved from 16th place of 26 nations in SISS 1984 to first place in TIMSS 1995.

The approach has been subject to continuous review and refinement, and has been accompanied by an important movement of personnel between teaching, ministries and State agencies, and publishers – a feature which has further consolidated the coherence of approach between actors in the system. A review in the late 90s introduced a revision to the sequencing and scope of material; tightening the focus and increasing the 'fewer things in greater depth' approach which had been observed in primary education by Reynolds and Farrell[43] in their transnational analysis. The refinement increased focus on problem-solving, and on higher order and conceptual thinking. In 2008, further refinement focused on ensuring fluency in computation, greater realisation of 'fewer things in greater depth' and problem-solving of verbally stated problems.

The approach has distinct parallels with Finland in respect of the deliberate focus on, and common understanding amongst all teachers of the underpinning principles of 'Singapore Maths'. The deliberate alignment of model, materials, professional development and practice is an exemplary case of Schmidt's 'curriculum coherence' and involves carefully managed relations between State, schools, private publishers and parents. This is reproduced in other areas of public policy and education policy, such as the major strategic alignment of sectoral economic stimulus, brokerage of supply of capital to new enterprises, with matched supply of skill and knowledge through the general and technical education systems.

Where now?

Singapore continues with its high commitment to educational quality for all, retaining the approach to management of improvement which has served it well to date: retention of elements of the system which are functioning well, constant domestic and international monitoring, and high levels of monitored innovation – maintaining the facility to adopt widely only those innovations in curriculum and pedagogy which yield genuine benefit. There is a recognition that constant 'fine-tuning' of arrangements is necessary. This is in part realised in 'normal tuning' through administrative processes such as annual State re-allocation of principles, and in part realised through new policy to tackle emergent issues such as pupil stress in primary education[44]. Alongside this responsive 'fine-tuning', maintaining progress in the administration's long-term policy to address curriculum balance (ensuring acquisition of higher order critical thinking as well as core knowledge), variation in teacher quality, and equity in outcomes across social groups remains a deliberate and managed part of the policy mix.

43 Reynolds, D. and Farrell, S. (1996) *Worlds Apart? A review of international surveys of educational achievement involving England.* Her Majesty's Stationery Office.

44 Ang, K. (2012) Singapore aims to curb stress on students. *The New York Times.* 07/12/12.

England

Innovation overview

Population 54.8 million[45]

Curriculum control located in 100+ local authorities from 1944; increasing central engagement 1970s, culminating in National Curriculum 1988.

Steadily increasing participation in 1970s; increased male/female equity post 1988; some subject-based improvements following specific national initiatives 1990s; 2001 onwards major shifts in school governance (academy programme); new national curriculum designed 2010–13, implemented in 2014.

From 1944 to 1988, England had no national curriculum; a curriculum was determined by 100+ local authorities within a broad set of national legal aims and objectives for education. In practice, 'steering at a distance' by the State resulted in a fairly consistent curriculum across the nation, but with very high levels of variation in outcomes. In 1944 a tripartite system of selective schools was established – grammar schools, secondary modern schools and technical schools. The last did not become a widely or well-established feature of the system. In the mid 1960s, there was a substantial shift to comprehensive, non-selective secondary schooling, although some 160 grammar schools (of around 3,000 secondary schools) remain in place. Around 7% of pupils attend private (independent) schools, a figure which has shifted little in the last five decades. In 1995, a substantial revision of national qualifications for 16 year olds was implemented, establishing a single class of examinations (GCSE) to replace the previous O levels and CSE qualifications. A national curriculum was established in 1998, for the first time in England, with attendant assessment at 7, 11, and 14 years of age. National qualifications continue to be used used for assessment at 16 (typically 10 subjects) and at 18 (typically three subjects).

Research on the impact of the National Curriculum showed positive impact on science provision in primary and secondary education; on female attainment in maths; on pupils prone to transferring schools; and on expectations of all pupils[46]. Implementation of the new GCSE examinations supported the National Curriculum aim of 'entitlement' for all pupils (equity in access to education and attainment). Despite the National Curriculum, the structural and administrative forms of education (size of schools, types of schools, transfer ages, etc) remain extremely diverse across the nation – considerably more diverse than in other comparable nations.

Throughout, the national inspection service – Her Majesty's Inspectorate of Schools (HMI), and then its successor the Office for Standards in Education (Ofsted) – remained crucial to State monitoring and evaluation of schools and schooling.

With the advent of national testing, the data from national assessment and examinations were used to establish an increasingly elaborated system of school accountability. During the 1990s, a full database of all students 5–16 was established. While the allocation of unique pupil identity numbers and other administrative aspects of these arrangements took time to embed, the comprehensive nature of student and school data, and its high dependability, remains a considerable achievement of public administration. The data are used extensively in system monitoring, examination standards-setting and standards-maintenance, school inspection and school performance measurement (accountability arrangements); and research.

45 Office of National Statistics, 2015.

46 Oates, T. (2010) *Could do better: using international comparisons to refine the National Curriculum in England*. Cambridge Assessment.

In 1997, Government funded the innovative EPPE project (Effective Provision of Pre-School Education), a 6-year longitudinal study, which established clearly the importance of high-quality pre-school experience in conveying educational advantage and improved equity in education. The project detected enhancement of equity in centres which integrated education, health and social care in areas of deprivation.

A wave of reform was put in place in 2000[47]. The National Curriculum had been revised in 1995 and 1999, with the reforms well received by schools, as each simplified and made clearer the expectations of the National Curriculum. In 2000, a major change in qualifications at 18 introduced AS qualifications (essentially to be attained after one year of advanced secondary education), encouraging greater breadth (four specialised, in-depth examinations at 17 followed by a focus on three higher level examinations at 18). The detailed marks available from these AS qualifications proved of particular value for university admissions decisions, since applications typically are made, and offers given, prior to the availability of final A level results at age 18. However, provision of AS qualifications was later discouraged (2010) as a result of perceived 'gaming' of results through retakes of AS, and difficulties in maintaining A level standards. The balance of advantages and disadvantages of AS provision remain strongly contested. However, the majority of universities throughout 2000 to present, while using AS results in admissions offers, continued to use 'three good A level grades' as the 'leading' outcome for university admission.

Accumulating social concerns about the load being placed on teachers by national testing at the ages of 7, 11 and 14 led to successive reductions in the amount of testing. However, evidence continues to suggest that reducing the subjects assessed has tended to encourage schools to narrow the curriculum, essentially by spending more teaching time on those subject areas which remain tested – essentially maths and reading. In response to instrumentalist responses to targets expressed as desirable assessment outcomes and exam grades, in 2010 Government reviewed accountability targets and measures, introducing more elaborated targets, particularly at age 16. Government targets continue to be potent in affecting the priorities of schools, demonstrated by the immediate impact of the non-statutory 'English Baccalaureate' measure – a specification of a specific combination of GCSE subjects in which higher grades should be obtained. Despite its non-statutory status, on its introduction a high proportion of schools immediately began to reconfigure their focus and provision in order to align outcomes with the requirements of the measure. This instrumental character of schools in the face of accountability measures and inspection requirements was also seen as the reason why the large-scale trialling of Assessment for Learning approaches (structured and targeted formative assessment 1999–2001) failed to realise the gains predicted in prior research synthesis on the approaches[48].

Also in 2010, Government reversed the recommendations of the 2007 review of the National Curriculum – which had resulted in massive reduction in the detail of the National Curriculum subject specifications. From 2010, a new review of the National Curriculum returned the specifications to detailed descriptions of 'core' elements of subject disciplines, essentially overturning the direction of the 2007 review and returned to the 'trajectory' of refinement of the previous 1995 and 1999 reviews. The 2010 review did not simply return to the 1999 specifications, but updated the curriculum using wide-ranging transnational review of the curriculum specifications of high-performing jurisdictions, combined with domestic evidence[49]. The new curriculum was implemented in

47 Hodgson, A. and Spours, K. (2003) *Beyond A Levels: Curriculum 2000 and the reform of 14–19 qualifications*. Kogan Page.

48 Black, P. et al (2003) *Assessment for Learning – putting it into practice*. Open University Press.

49 Oates, T. (2014) Progress in science education? The revised National Curriculum for 2014. *School Science Review* vol 95 no 352, pp21–29.

September 2014, along with revised GCSE and A level qualifications and new assessment arrangements. A very explicit part of the design and implementation theory associated with the 2010 Review was the concept of 'curriculum coherence' and deliberate alignment of key factors across educational arrangements. While the Secretary of State 2010–14 was characterised by some as being 'hyperactive', explicitly he emphasised that the rationale for wide-sweeping changes was the alignment of curriculum, assessment, the structure and governance of schools, funding, inspection and other key factors.

But unlike the drive to national standards in Massachusetts, British Columbia, and other states, in England the National Curriculum is not a requirement of all State-funded schools. School reform has proceeded not only by changes to curriculum and assessment, but also through the revised status of schools. Two new categories of school have been introduced to arrangements. 'Academies' were first announced in 2000, as part of 'School improvement' strategy. By 2010 around 200 schools' 'forced conversion' was required where by 2016, over 2,000 of the 3,400 State secondary schools had changed status to 'academy' schools – either by self-election or by State requirement. Academies do not have to follow the National Curriculum, although they do need to follow a legal requirement for a 'broad and balanced curriculum' and must meet requirements regarding national testing and national examinations. The second category of schools, 'Free Schools', were given approval in the Academies Act of 2010, and are similar in concept to Swedish Free Schools and US Charter Schools; and are set up by community groups and/or organisations such as universities, charities and faith groups. Contrasting with Finland, much reform of the past two decades has focused on successive waves of qualification and assessment reform – associated with 'system steering' through accountability measures. Research shows strong impact of this approach[50], but without substantial impact on results in international surveys. A rise in maths outcomes in TIMSS can be attributed to the 'numeracy strategy' of the late 90s, which combined with a 'literacy strategy' in schools. This was a highly prescriptive set of interventions in school curricula, and whilst not statutory, they were associated by schools with notions of compliance with State inspection of schools by Ofsted. The high degree of prescription of the content of teaching – the Strategy guidance documents to primary and secondary schools were essentially a system of didactics for maths and literacy teaching – remains controversial. While some commentators associated it with an undermining of teacher autonomy, others saw the National Strategies as massive programmes of intensive staff development.

School funding has not been neglected as a policy instrument. From 1997, successive changes in the detail of funding entitlements were oriented towards directing funding to more deprived localities or schools with higher proportions of pupils affected by particular needs – such as speaking English as a second language. By 2017, this had resulted in very significant discrepancies between schools' funding entitlement – with some schools funded to a level almost twice that of schools with the lowest level of funding. In 2017, an initial policy attempt to introduce greater parity and 'fairness' in funding was abandoned, after it was clear that the complex revised formula (which included a criterion to direct funding to areas of social deprivation) would result in substantial reductions in funding to schools already receiving low levels of State financial provision. Unable in a time of financial discipline regarding Government expenditure to 'level up' all schools to the level of funding enjoyed by the best-funded State schools, pressure from community groups, schools and others led to a Government decision to withdraw proposed changes and re-examine the detail of a revised funding formula.

50 Hamilton, L. S. et al (2002) *Making sense of test-based accountability in education*. Rand; Hooge, E. et al (2012) *Looking beyond the numbers: stakeholders and multiple school accountability*. OECD.

Where now?

What characterises the last three decades of reform and improvement effort in England is the emphasis on qualifications reform and accountability measures, based on an assumption of the strong impact, on school behaviour, of both of these. Secondly, all reform in England takes place in the context of extraordinary variation across the entire system in respect of the form of schools, the enacted curriculum, didactic and pedagogic models and assumptions, and almost all dimensions of educational arrangements. This presents policy makers with significant challenge in respect of policy formation and system management. The new National Curriculum (implemented in September 2014) has 'landed' well in schools; assessment at age 11 of the curriculum content in maths and English continues to cause reaction, while policy on the reform of textbooks has had very rapid beneficial impact on the form of support materials. New, more demanding examinations at 16 and 19 were being taken for the first time in 2017.

Massachusetts

Innovation overview

Population 3.6 million[51]

Massachusetts was an early adopter of statements of educational 'standards' – these were at the heart of the 1993 Massachusetts Education Reform Act, which also included funding for experimental 'charter schools'. The Act thus included curriculum reform encouraging conformance to common standards, and structural reform which increased the putative autonomy of schools. This tension between restriction and autonomy has played out in complex ways in the last two decades, both in Massachusetts and the rest of the USA.

Massachusetts has long been regarded as a high performer in the US context, and if it were treated as a separate system in the manner of Singapore, Finland etc, it would feature in PISA 2015 amongst the top eight jurisdictions in Reading (with no significant difference in scores in this top group), be second only to Singapore in science, and ranked 12th in mathematics.

For the last three decades, standards-based reform has been a principal strategy for educational improvement in the USA, with test-based accountability allied to those standards. Reform in Massachusetts was initiated in 1993, with the enactment of the Massachusetts Education Reform Act (MERA). From 1996, a framework of subject-based curriculum standards was developed (Curriculum Frameworks), initially in five main subject areas, with frameworks for technical and vocational education implemented in 2006, and 2008 for Kindergarten provision. Revisions in the first set of standards occurred in 2000–04, and 2008 onwards.

The MERA included structural as well as curriculum reform, with the creation of Massachusetts Charter Schools. These changes are interesting in terms of patterns of control and governance: a simultaneous combination of increased restriction through common curriculum requirements and related assessment on the one hand, and on the other hand, relaxation of control through changes to governance. For example, although aligned to the State standards, teacher training has remained located in 75 separate institutions, with allocation devolved to schools – there is a disproportionate number of newly trained teachers in underperforming schools. It is important to note that there

51 US Census Bureau, 2014.

remains national controversy and discussion regarding the characteristics of teachers, which mostly support high equity and high attainment.

The specific mix of restriction in some elements of arrangements and loosening in others was not undertaken inadvertently. Standards-based reform in the USA is underpinned by assumptions that a mix of 'firm central direction and maximum individual autonomy' is both possible and an effective overall strategy. Current evaluations, such as that by Paul Peterson[52] (Harvard and Stanford), suggest that if the aim is a step-change in attainment combined with higher equity in outcomes, these assumptions may need to be questioned. Peterson's well-designed and largely philanthropically funded evaluation of school choice and voucher policy highlights the curriculum variation (pedagogic quality; curriculum focus; expectations of pupils) which appears to explain differences in performance between charter schools. This cuts through the debate regarding 'charter schools good/bad', since charter schools are not universally an improvement over the schools which they replace. The issue of 'curriculum coherence' and 'curriculum control' becomes important in the light of this variation, and throws light on the particular mix of restriction and control present in the US context.

Since 1993, whilst average gains in Massachusetts have been impressive, there remain significant attainment gaps across the state, with Massachusetts continuing to possess one of the highest attainment gaps in the US (3rd highest in 2015). This indeed points to some interesting 'control factor' issues. Governance arrangements in the US and in Massachusetts continue to be controversial – spending variations result principally from governance structures: six districts exceed 25,000 USD per pupil funding, while seven spend below 11,000 USD, leaving Massachusetts amongst the 10 US states with the highest per-pupil district spending variation.

While the Curriculum Frameworks developed under 1993 MERA enjoyed widespread social confidence and support, the focus of all institutions and professionals on the State standards has been seriously disrupted by discussions of the alignment of the Massachusetts standards with the Common Core standards, with Partnership for Assessment of Readiness for College and Careers (PARCC) standards, and with the requirements for accessing Federal funds which were associated with the 2001 NCLB (No Child Left Behind) Federal improvement strategy. The issue of alignment between more recent Federal standards and the State standards set originally in the 1990s, and which enjoyed widespread support in the education sector and throughout the State, has invoked widespread controversy, and has pre-occupied many key political and educational institutions. Many parental groups have raised protracted objections regarding dilution of standards and quality, breaking the consensus which had been created around the original standards. This uncertainty in respect of curriculum has been accompanied by a long period of shifting patterns of control, incentives, and access to funding at school level, as the school competition policy has rolled out. The structural changes – the use of competition as an improvement strategy – remains controversial. Considerable attention continues to be paid to the composition of district and school governance, with wide social and media discussion of the influence of different groups and individuals on the direction of policy in respect of funding levels, and funding flows to different types of schools.

Interestingly, in contrast to Hong Kong, Shanghai, Singapore and reform-period Finland, Massachusetts does not have State-written or State-approved textbooks. But this is deceptive. Textbook purchasing is subject to regulatory restriction, with State law stating that while school principals are responsible for purchasing decisions regarding 'textbooks and other educational materials and supplies', these decisions are taken

52 Peterson, P. (2016) Post-regulatory School Reform. *Harvard Magazine*. September-October 2016.

under the supervision of the District Superintendent and must be 'consistent with the educational goals and policies established by the [District] school committee'. This process means that textbook purchasing and approval exists in joint decision with a level above that of the individual school, and the goals are heavily determined by the pervasive State Curriculum Frameworks.

One element of the restrictions placed on schools by MERA appears to have had in turn particular potency: high school diplomas were linked to pupils passing State tests linked to the curriculum frameworks. Standards in the tests initially were set after careful balancing of the educational goods deriving from setting high expectations and the problems of denying graduation diplomas to too many pupils. Initially setting the pass standards at the lower rate from a range of alternatives has nonetheless appeared to result in a significant elevation in outcome standards in high schools, across the State. Although there has been some improvement below high school level, gains are not as evident in earlier phases of education as they are in the later phases, where the impact of the test requirements is particularly prominent.

Where now?

Reform in Massachusetts focused on standards-based strategy (curriculum frameworks with allied assessment requirements), combined with structural reform based on school competition. The reform process was initiated at a time of considerable crisis in school funding, an issue which has persisted throughout the reform period. The central aspects of the curriculum reform have been dogged by slow initial development of the frameworks, and repeated unsettling by external (Federal) policy developments on alternative sets of standards. Curriculum standards have thus not been stable, and clarity and social consensus have not been sustained. The structural changes have invoked protracted tensions over governance and curriculum control. Nonetheless, considerable gains in attainment have been made, albeit with persistent issues of equity across schools and across society. With these still extant, the original aims of the reforms remain far from fully realised. The 1995 call to consensus present in the Commissioner for Education's open letter continues to appear to be a desirable goal, but not yet realised in the State:

> **'Education Reform will be successful if all parties involved in local school governance – school committees, superintendents, principals, and school councils – communicate and work collaboratively with each other and with the wider community to achieve the common goal of improving educational opportunities and outcomes for students…those who work in and with schools and students "must share a vision, a clear purpose, and the ability and courage to lead". We hope this advisory helps to clarify the roles and relationships in the local school governance structure. If we are to assure that our students realize the promise that Education Reform holds, we must continue to work together'.**[53]

53 Letter from Robert V Antonucci, Commissioner of Education 01/11/95. http://www.doe.mass.edu/lawsregs/advisory/cm1115gov.html. Accessed 01/06/17.

12 Concluding remarks

We said the following at the front of this document:

This document does not give precise steps to formulating policy or managing implementation. We think that to do so would be quite wrong. Different nations, at different times, face different challenges, have different resources available and are presented with contrasting opportunities to effect change. Sometimes urgent action is required, sometimes the long view needs to be taken. In recognition of this, we do not here recommend a fixed approach to using the insights and approaches outlined in this document. Instead, the text asserts some strong principles and models, underpinned by research, to support effective policy formation and implementation strategy. This is intended to guide thinking on policy formation, making sure that policy formation takes a more comprehensive view of the forces and factors at work in education systems.

This document offers 'high level organising principles' – they are no less useful for being high level. They have extremely practical applications.

We believe that the analysis provides new insights and valuable perspectives to policy makers, researchers and all those interested in attainment and equity in education. We believe that the country descriptions add contrasting contexts and perspectives, allowing the 'control factors' and 'explanatory' factors to be seen as a valuable lens through which to view educational arrangements and improvement strategy.

Every year eight million learners take our exams in 170 countries, and with 160 years' experience and the largest research facility of its kind in Europe, the Cambridge Assessment Group is chosen to be a key education partner to governments across the globe.

We are committed to education excellence and are trusted by universities, colleges, employers and governments, and are proud of the relationships we build and partnerships we form around the world.

At Cambridge Assessment, the reliability of our assessments stems from evidence-based and research-led approaches to all products, services and new developments. We have the largest research capability of its kind in Europe with more than 100 researchers across the Group.

Our researchers conduct and publish authoritative research in order to validate, improve and develop our assessments and services, and to influence thinking and policy on educational assessment nationally and internationally.

cambridgeassessment.org.uk
@Cam_Assessment

University of Cambridge
Local Examinations Syndicate

Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate, a department of the University of Cambridge. Cambridge Assessment is a not-for-profit organisation.

5799909010