Flexibility within rigour – achieving more learning for more learners through reporting against aspirational standards

R. Allen
August 2008
“Insanity is doing the same thing, over and over again, but expecting different results”
(Rita Mae Brown 1983, also attributed to Einstein and Franklin)

“We can’t solve problems by using the same kind of thinking we used when we created them”
(attributed to Einstein)

*The Ti-Twwadi Effect (this is the way we always do it) is very resilient – it sneaks back in when you least expect it* (Anon)

*To train and educate the rising generation will at all times be the first object of society, to which every other will be subordinate.* (Robert Owen, The Social System, 1826)

*Better an imprecise measure of something important than a precise measure of something unimportant.* (David Byar)
The Commonwealth government has described three priorities for education and training:

- building social capital
- boosting productivity
- social inclusion.

Prime Minister Kevin Rudd, echoing Robert Owen (cited above) nearly two centuries ago, has said: *We want education to be about lifelong learning. From the cradle to the classroom, from the living room to the workplace, we need to keep investing in ourselves, in our skills and therefore in our future.*

We face a challenge: more learning for more learners.

I won’t repeat the facts and figures that delineate the details of this challenge. It is enough to note that it is clear that there is a big gap between where we are and where we want to be.

This gap has three dimensions. First, there are many people who learn less than they might. Secondly, there are shortages of people in many areas of skills and knowledge that are important to our future. We don’t have surpluses of people with technical skills, with knowledge and skills in mathematics and science. Thirdly, we don’t do well enough in terms of equity. Whatever the causes, there are clear relationships between educational outcomes and students’ socio-economic background and gender.

We can sidestep this challenge, especially if we think that’s the way it has to be, that the gap in all its dimensions is fixed by nature. Differences in the patterns of participation and achievement in different places or different times suggest that nature isn’t a good culprit. A more subtle sidestep says that the gap is fixed by external circumstances, ones we can’t change. Obstacles of culture, history, privation or poverty prevent us, we say, from meeting the challenge.

We’ve seen this challenge before. Past responses fall into five categories:

- resources
- curriculum change
- system-wide documentation
- measuring
- accountability.

Responses based around resources include:

- improved buildings and infra-structure
- more teachers
- more pay for teachers
- more professional development
- more curriculum support materials.
Resource-based responses have the advantage that they come complete with a built-in let-out clause for any lack of effectiveness. If the answer is ‘more resources’ then lack of impact of any increase is taken to be evidence that the increase should be bigger.

Thirdly, responses to this challenge in the last 100 years based around curriculum change have included:

- more diversity (localised variations) in curriculum
- more choice for students of courses, units, modules
- multi-disciplinary studies
- integrated studies
- individualised learning
- subject disciplines
- concentration on the basics
- less diversity in curriculum.

This list illustrates that for each curriculum change ‘particle’ it seems there is an ‘anti-particle’. The swirl around curriculum, hot and contested as it is, has had no clear and sustained positive impact on meeting our challenge – more learning for more learners.

Responses based around system-wide documentation are based on an assumption that the lack of learning is caused by lack of knowing just what is supposed to be learned. That is, we need to set out very clearly just what teachers are supposed to be teaching and students are supposed to be learning.

Some years ago I acquired a teacher’s manual that includes a copy of England’s Elementary School Code of 1903. This sets out in clear detail what students were supposed to learn (and lots of other good advice as well). It seems so obvious that we should start here, that this is the foundation for future success. It has been and is popular in many countries. The many different versions of this response that have been used have in common a lack of clear success. Should we start asking why these approaches have not led to the desired increases in learning for all students? And how we might approach the problem differently?

The fifth category of responses to our challenge assumes that what is needed is better measurement of student learning. With more and finer detail based on better and better tests teachers will at last really know what they have to do. Something they don’t know without these test results. And we will be able to find evidence-based methods, methods whose success is demonstrated beyond doubt. I won’t go through the arguments for and against this here. There’s a useful overview at http://www.edutopia.org/assessment-standardized-tests-flaws. What is worth noting is something I will return to later. These measurements do not capture much of what we see as really important in education and training. Their use focuses our attention on a few aspects. And it moves our eyes away from other, equally vital, equally important, learning.

I’ve heard an argument that literacy and numeracy in the way they appear in many discussions of the basics are nineteenth century skills. I’ve heard the reverse – that
literacy and numeracy are the only basics, that nothing else matters. I’ve also heard the case made that most people don’t need or use mathematics, that they can be and are successful without it. As we think about learning for the future – it is hard to see how one can have learning for the past - there is a different case to be made. Literacy and numeracy are like the wheel-bearings on a car – without them, the wheels fall off. But without an engine and gearbox the car doesn’t go forward. If we focus our attention only on the wheel-bearings we can be sure that the wheels won’t fall off. That’s not enough.

The last category of responses to our challenge is to drive change through improvements to system and school accountability. This means providing indicators of systems’ (and schools’) performance, their outputs and outcomes, in ways that let us reach judgments about how well they function and where improvements should be made.

Sheerens, Glas and Thomas ¹ (2003, p. 216) describe the systematic model for system-level indicators used in the OECD-INES project. This has five categories of indicators: context, input, process, outputs or outcomes, impact or long-term outcomes. Output (outcome) indicators means things like participation rates, progression rates, drop-out rates, average achievement in the ‘basics’ and in generic skills. Impact indicators are things like labour market outcomes and changes in other areas of society that ‘can be seen as the effects of education’. These are measurable indicators and so are thought to provide a sound basis for holding systems accountable for their performance and hence for increasing their success. In this view, as Beadle (in Sirotnik 2004, p. 46) puts it, there is the assumption ‘that differences in rate of academic achievement across schools [and systems] are essentially a technical problem, and that the failure to resolve this problem is due primarily to a lack of will and expertise on the part of educators’². This extends easily into the punitive version of accountability systems – the assumption that punishment or the threat of punishment will lead to improvements.

The United States has been exploring the idea of system, school, teacher and student accountability in depth for many years. We know that the US:

- spends more money on education than any other country
- obtains mediocre outcomes (see www.2mminutes.com)
- has shown no improvement in high school graduation rates since 1965
- has tightly specified content and linked testing regimes that do not achieve focus on the knowledge and skills said to be needed for a global economy
- has not demonstrated improvements in outcomes (except for some modest improvements for minorities) (see www.skillscommission.org and Baker 2007 The Ends of Testing, Educational Research vol. 36 No. 6).

---


By
And its *No Child Left Behind* regime has a strongly punitive element.

England has twenty years experience with ‘increasing government control of the curriculum, the assessment of pupil attainment and mechanisms for assuring system effectiveness’ (Wyse, McCreery and Torrance 2008 ³).

England

- has detailed statutory specification of core content and achievement standards
- defines achievement standards by reference to a narrow range of educational outcomes
- publishes school reports and tables
- compares schools by ‘value-added’ measures ⁴
- has an inspection system with consequences, one that ‘inspects and regulates to achieve excellence’. (see [http://www.ofsted.gov.uk/](http://www.ofsted.gov.uk/))

The results, however, do not provide “evidence to support the claim that testing drives up standards of student achievement.” (see [www.primaryreview.org.uk](http://www.primaryreview.org.uk))

While the official rhetoric is that the system drives real and important improvement, recent developments and official actions suggest otherwise. England has now taken a more ‘flexible’ approach to curriculum to offer ‘increased opportunity to design learning that develops the wider skills for life and learning as well as making links to the major ideas and challenges that face society and have significance for individuals’. (see [http://curriculum.qca.org.uk/key-stages-3-and-4/developing-your-curriculum/what_has_changed_and_why/index.aspx](http://curriculum.qca.org.uk/key-stages-3-and-4/developing-your-curriculum/what_has_changed_and_why/index.aspx))

What implications should we take from the lack of major, sustained and significant impact of each of these categories of responses?


*The idea that education is just a system of schooling invites the idea that the best way to improve it is through the techniques of mass customisation, efficiency and quality improvement, driven on by central targets, national strategies and inspection regimes.* (Leadbeater 2008 ⁵)

---


Let’s try another way of thinking about system improvement. Here’s an example.
Prime Minister Kevin Rudd wants Australia to become “the most Asia-literate country in the world.”

To do this requires
• that our students learn, not that we adopt statements about giving them an opportunity to learn
• significant, systematic and sustainable changes in practices in schools, since right now very few students are Asia-literate.

Three key elements

1. A high performance standard written in terms of outcomes, not processes.
2. Education systems report regularly against this standard.
3. An external body uses a variety of evidence gathering techniques to evaluate and verify different aspects of these reports on a cyclical basis.

1. Performance standard

The standard will focus on each system’s performance in developing students’ Asia literacy, not the system’s processes or the opportunities it provides.

The standard describes an outstanding level of performance and so can be used to both measure existing performance and drive continuous improvement.

An evaluation (internal with external validation) against the standard will assess the extent to which a system’s output meets it, not yes/no compliance. The standard is therefore written to require an assessment of the extent to which it is met and is aspirational rather than regulatory.

Here is the standard:

By the time they leave school, young Australians understand the importance of our neighbours in Asia in their future success and how to live and work with people of many cultures and backgrounds.

2. Each system reports against these standards

At some suitable interval (every three years), after an initial base-line assessment, each system provides a self-assessment report of how well it meets the standard, setting out the
extent to which it is achieving the required student learning and the evidence it has for this.

The reports are written to help the system to do better and tell its community how well it is doing, rather than as a compliance or bureaucratic accountability task.

3. External evaluation/audit

An external evaluation/audit both gives people confidence in the accuracy of each systems’ self-assessment report and supports the system in making improvements.

The external evaluation/audit will look at the adequacy of the evidence given by systems in their self-assessment, asking if the evidence supports the claims about students’ development of the knowledge, skills and understanding in the performance standard.

The evaluation/audit will be done by an independent evaluator and at a low cost – its role is as a cross-check of systems’ self-assessment. It will focus on different aspects of the reports at different times.

The ABS will do surveys of young Australians’ understanding and attitudes in ways that will give a useful perspective on systems’ self-assessment about their success in developing Asia-literacy.
Such an approach has a lot in common with outputs and outcomes aspects of the OECD-INES ideas about system accountability I described earlier. The differences are, however, important.

First, there is a standard. And it is linked to something thought to be very important, at least by prime ministers.

Secondly, the standard links directly with a quality of what is achieved by the educational system, not with opportunities to be provided or access to be given. Note that the standard is about what happens, the achievement of the system is to contribute to this. It’s important to be careful about going from ‘students have these results’ to ‘schools achieve these results’.

Thirdly, it is an example of something that is assessable, something that can be judged, but not something measured by multiple-choice tests or given by aggregated data on educational achievement.

There’s no mention of processes, indeed none of the old favourites listed in the first part of this paper.

And educators do like to focus on inputs and processes, don’t they. It’s easier to be accountable for these, of course, especially processes. For processes, the question is ‘do we have them’, not ‘do they work’. Because to answer the second question is challenging – challenging to measurement and challenging to our self-esteem. If we found that our processes didn’t work, we’d have to change our practices wouldn’t we? Or we’d have to find explanations in other circumstances, in context, in deficient inputs or in ambiguities in the standards or in the assessment of outputs and outcomes.

The standard is simply stated, easy to understand, clearly worthwhile and, I think, easy to recognise. Because it is aspirational and not compliance-based it can be stated without splitting it into components.

The processes, however, needed to reach towards this standard are almost certainly many, varied and complex. A lot of different factors, some of them alterable immediately, some alterable over time and some outside the influence of systems, schools and teachers, will affect how well young Australians understand how to live and work with people from many different cultures and backgrounds.

The existence of factors outside the immediate influence of systems, schools and teachers will lead to the immediate objection that “we can’t be accountable for something we can’t directly control”. In turn, this usually leads in education to the reduction of performance expectations to statements of intentions and provision of access and opportunities.

It doesn’t have to.

---

6 When I’ve talked with people about this standard, most think they already know how (not at all) well we do right now. A more refined analysis is unlikely to do more than provide nuances.
Instead, we can recognise that we can report against worthwhile, important, standards in the following terms:

- this is how well what happens matches up to the standard
- this is how what we’ve done contributes to this
- these are the factors we think we can’t influence and this is how much influence on the outcomes we think these factors have.

The adequacy of this report is open to public scrutiny. The public can examine its claims about its effectiveness.

If we really believe that social and cultural factors prevent us from educating young people we should be prepared to say so, simply and clearly. Reluctance to do so directly and personally to individual students is the objection I have to those who would use ‘socio-economic status’ or gender to explain away differences in students’ reading, writing and mathematics.

Suppose that the education systems of this country were able to say to the community:

- these are the standards for our work
- the standards focus on the most important aspects of what we do
- this is how well we do in terms of these standards (our quality)
- this is why you can believe what we say about how well we do
- you can therefore have confidence in our quality.

The standards will describe (very) good practice levels of performance and so can be used to both measure existing performance and drive continuous improvement. A system’s self assessment will include an identification of the particular organisational processes and practices that contribute to meeting the standards and identify opportunities for improvement. These processes and practices may well vary from one system to another.

In summary, the key elements of this approach are

- a small number of high-level aspirational standards describing the important outcomes we want education systems to achieve
- systems provide public reports of what they achieve in terms of these standards
- confidence in the value and accuracy of these reports through external evaluation.

To further illustrate the ideas, here are some draft aspirational standards for three areas of system performance, covering

- the match of student learning and their destinations
- equity of outcomes
- the qualities of the knowledge and skills students learn.
1. Learners are well prepared for their post-senior secondary destinations.
2. Patterns of learner participation and achievement are unrelated to socio-economic background and gender.
3. Learners have the knowledge and skills required for them to prosper in the information-rich world of the twenty-first century and participate in society as active and informed Australian citizens.

It is easy to see how evidence can be gathered to show how actual performance measures up against the first and second of these standards.

Studies that look at where students go to and how well their skills and knowledge match up with what they need there are not difficult to do, have been done and are not expensive. The critical element is to use these studies systematically: as evidence of the system’s performance and as the basis for continuous improvement.

There are two obvious objections to the third of these standards, which is about twenty-first century learning: Learners have the knowledge and skills required for them to prosper ...

First, it might be tricky to gather evidence about to extent to which learners have the required knowledge and skills, since what they are might not be sufficiently clear and the learners’ prosperity is in the future, not right now.

Of course, if we can’t say what twenty-first century learning means and we can’t find out if it is happening then we should perhaps wonder about the place of a statement like this in our national goals for schooling.

On the other hand, having such a statement and not reporting against it seems a recipe for paying it at best the lip-service of its inclusion in curriculum statements without any discernable impact on what actually happens and what learning is actually achieved.

The second objection is that the statement is ‘really’ a goal, not a standard. One definition of a goal (or an objective) is that it is a situation that a system intends to achieve. Implicit in the objection that this is a goal, rather than a standard, is an assumption that goals are high-level guides to intentions, ones that we don’t expect to achieve and that don’t have sufficient clarity or focus to them that we can find out how far we’ve got towards them. In this sense, goals provide a guide to designing inputs – such as curriculum documents – and not as a statement of where we intend to be against which we can describe our achievements, our outputs and outcomes.

How well has this approach served us in the last thirty or more years? Has it helped us achieve better or has it served as a convenient evasion, along with its helper words like ‘opportunity’, ‘access’, ‘maximise’, ‘offer’.

It is at least arguable that what are called ‘employability’ skills are part of those needed to live and prosper in the twenty-first century. So are literacy and numeracy. So is use of computers
and the internet. Equally, so also is the Asia literacy described in the example I gave before. It can also be argued that these do not capture everything important about living and prospering in the future.

There will be other areas of performance that merit standards, provided we can be clear enough about what they mean. Is what is called ‘student well-being’ – social and emotional health – something to which education systems contribute? If so, should there be some standard for the outcome, a standard that systems report against? What happens to things that we like to say are important but are not a fundamental part of how systems describe what they achieve and how these descriptions are evaluated? How effective are declarations?

The sorts of processes described by Leadbetter for ‘21st century schools’ – small schools, building sustainable relationships for learning, mentoring, personalised plans, peer support, community-based teachers, short holidays, local education compacts – are the sorts of ideas about processes that will be needed to reach towards worthwhile standards, such as these, for the quality of learning and its equitable attainment. Under the model proposed here, these would be natural ways in which systems sought to improve their performance.

This model for system improvement does not specify how things are to be done nor how achievement against the standard is to be reported. It replaces the belief, sketched in the first part of this talk, that requiring or encouraging change in processes is sufficient to drive changes in outputs and outcomes. It replaces the belief that a narrow focus on a limited range of outputs is a sufficient driver of change in the full range of important outputs. Instead, it is based on the assumption that publicly reporting against major matters is a key aspect of creating a continuous improvement approach in education – the drive to do better.

Each system will gather evidence that will show the extent of its success. A system that wants to do better will seek advice from experts about improvements to its processes. It will examine very carefully and self-critically what it does and how it knows how well it does. It will, with a standard like the first of the three in my example, need to show that it has identified what is actually needed for future prosperity (as opposed what is thought or asserted to be needed).

Curriculum, materials and other resources, pedagogy and assessment, like school organisation, are critically important processes, of course. But they are a means to an end, not an end in themselves. And so, in this approach, they are matters on which a system that wishes to improve its performance in terms of its achievement against these standards – one that seeks to make students more Asia-literate, one that seeks to achieve high levels of equity in its outcomes – will seek expert advice on how it should change its processes, will seek the best curriculum it can find, the best professional development it can afford and the most informative assessment that contributes to improvements in learning.

---

Otherwise the system will find itself reporting in public that it hasn’t done any better since the last time or that the external evaluation/audit of its report will show that its claims are unfounded.

In this approach, the rigour is in the standards, the flexibility is in the means to the ends, in the assessment against these standards, in evidence that is produced and in the public reports – their content, not their function.

It is very important to be careful about a word like ‘flexibility’. It can be used as a side-step, a way to avoid our challenge. Let’s not do that.

Flexibility too often has a covert meaning: applying lower standards for student achievement, leaving out the difficult bits in courses, not asking schools in difficult circumstances to achieve as much as others, avoiding the sharp questions about what is achieved and at what cost.

Flexibility is sometimes argued for as something nice to have or, in a federal system, as a way for state systems to eat their funding cake but feel in control.

In some parts of the world, flexibility is a consequence of the importance in democratic terms given to local control.

The idea of responsiveness to parent control, something argued for by many parts of the non-government sector in Australia, surely entails flexibility in curriculum, in provision, in methods and in assessment.

In the context of meeting the challenge for more learning for more learners, this paper argues none of these senses of flexibility or the reasons given for it.

The record of approaches to improvement through locking down what happens in schools and classrooms is poor. They haven’t succeeded, despite lots of attempts, beyond at best reaching a basic level, a level too low, I believe, to meet our needs for future economic and social prosperity for all. Have they been tried enough to justify looking for a different approach or should we tighten down the screws still further?

Worse, should we instead be satisfied as a nation with some more curriculum documentation and some simple counting:

- proportions of 19 year-olds attaining year 12 or equivalent (whatever that means)
- proportions reaching literacy and numeracy benchmarks
- attendance rates.

It is not that counting or curriculum documentation are bad things, just that they are only a small part of the response we should make. They haven’t been effective drivers of the sorts of improvements we need.
Why should other factors not get a guernsey in national accountability, in how systems account for what they do? It should not be beyond us to judge the quality of learning, the relevance of learning to post-school destinations, Asia-literacy and levels of equity.

This is a time when as a nation we could develop an approach to accountability that aligns what is reported with what is important and that develops, promotes and rewards the implementation of a continuous improvement approach within education.

A lot of money goes into education and training. Is there an opportunity for us to evaluate and report on the quality of what we get for this money in terms of what is achieved, of all the major outputs and outcomes we want, not just a few of them? What sort of vision and leadership is needed to do this?