

NATIONAL STANDARDS AND LEAGUE TABLES: POLITICAL AGENDA

NOT SCHOOL ASSESSMENT

Cedric Croft EDUCATIONAL RESEARCHER



FROM THE TIME the Education Standards Act was passed under urgency after the 2008 election, it's been evident that the so called 'National Standards' are more to do with politics and the three-year election cycle, than student learning, or primary school assessment. Few will forget the claims from the present Minister that her 'standards' would help promote plain language reporting, identify the bottom 20 per cent of students, improve achievement generally, and that somehow, by some undisclosed means, they would ensure that all students were prepared to pass NCEA Level Two. Any questioning of the 'standards' was responded to in political terms by a Minister unable or unwilling to address the underlying assessment problems that continue to bedevil this ill conceived, overly expensive and poorly implemented assessment policy. However, this political agenda need not deflect professionals from scrutiny of the policy.

If political capital was to be gained from National Standards they had to be implemented rapidly so that claims could be made to the electorate about their success. The restricted time available is a major reason for the poor quality 'National Standards' (which are little more than hastily adapted curriculum statements) that primary teachers are now required to use as a basis for reporting aspects of reading, writing and mathematics twice a year for every primary school student. The necessity for speed inhibited the development of sound and valid standards using an appropriate methodology. No trial of the Standards was undertaken and there is no published evidence on how consistently they are applied. Assured validity and reliability are imperative if the Standards are to operate credibly or be used as a basis for league tables, which are discussed later on.

A point of confusion with the 'Standards' is that politicians and officials who promote them are not clear what they are talking about, as the 'standards' are variously referred to as 'targets', 'signposts', benchmarks, 'reference points', 'broad descriptions of expectation', or 'aspirational standards' and are too often confused with 'norms'. 'Broad descriptions of expectation' and 'aspirational standards' at least capture some elements of an educational standard, but they add to the proliferation of terminology which does not help clear communication. But 'target' and 'signpost' have no established meaning within educational assessment and may give false impressions about the underlying nature of assessment standards. Benchmarks are often taken erroneously as a synonym for norms, but norms are not standards.

The confusion between norms and standards is potentially the more serious, as the distinction should be clear to officials or politicians who are influential decision makers. But clearly this is

not so. In their educational assessment sense norms show what students do achieve. A norm is readily recognisable. Examples of common forms of norm are a student score on a standardised test, a readability estimate of school text, the frequency of 'essential words' in writing from national samples, or the number of students nationally at 'advanced counting' in junior numeracy. A standard is more subtle, but in essence will describe a desirable outcome for students. Norms describe 'what is', standards describe 'what should be'. For a standard to become a basis for student assessment as the current 'standards' purport, there needs to be a clear and unambiguous description of an achievement dimension, accompanied by a means of assessing the extent to which the dimension has been achieved. These requirements are not evident in most of the 'standards' as published. Supporters of the 'standards' may point to the four-point scale used for the assessments, but these do not compensate for the lack of clear assessment criteria for the abundance of outcomes covered by each 'standard'.

A distinction between norms and standards may be further confused for some by the manner in which scores from some common standardised tests may be incorporated in the Overall Teacher Judgment. It is noted in the *Blueprint for National Standards*, for example, published in the *Education Gazette*, 11 October 2010, that a STAR total score, an asTTle reading score, and a PAT comprehension score may help determine how well a standard is met. It seems to be ignored that each test will cover a small portion only of the broad reading standard, and that each score point quoted is given without noting the error range for each score point. It is questionable whether the methodology used to map test scores to the 'Standards' was sufficiently robust for these recommendations to be made. The essential point is that mixing norms and standards is not a valid means of assisting teachers to assign 'Standards' to individual learners, unless of course the standard is more like a norm! The reading Standards are mostly norms anyway, as readability estimates are central, but they have rather confused the issue and given many teachers a lot of unnecessary work. It is to be hoped that existing assessment tools and procedures are not abandoned in favour of the reading Standards.

There was ample opportunity in the Ministry article noted above to clarify the distinction between norms and standards and to discuss the relative contributions each may make to primary school assessment. However, these matters were left out. Instead there is a discursive discussion of the model used to develop National Standards, which traverses very broadly



over a number of the hallmarks of school assessments without ever getting to grips with the so called 'model'. But usefully it is stated that, 'National Standards are broad descriptions of expectation . . . which . . . use both descriptions and exemplars. The description states what is required at the standard, while the exemplar shows what this looks like in practice.'

This statement is very useful at its face value but where it fails is that the curriculum statements masquerading as standards do not, in most cases, describe in any clear or unambiguous manner what the achievement to be assessed actually is.

In most cases the Standards cover too much to ensure that all teachers are assessing their students on the same outcomes. They abound with subjective terms, require reference to multiple sources before an assessment may be completed, are likely to require more time to complete than the quality justifies, and provide little basis for reporting to parents, or for teachers to plan further learning.

The article noted previously states that New Zealand, ' . . . has taken a different approach to the rest of the world'. This is certainly true but it is not necessarily a cause for self-congratulation, as to the best of my knowledge no other modern education system of our type has made so many fundamental errors when introducing a national policy on student assessment. Within the 35 years or so that I have worked in primary school assessment, no new assessment policy has been so poorly conceived or introduced in a manner which puts so much responsibility on teachers and principals to make it appear to be working. In the short term it may seem as though they are working but as research data are collected and parents begin to take more interest in what 'National Standards' mean for them, this may not be the case.

Already some early data show that the mathematics standards are well out of line with PAT data which have their reliability and validity established.

With the advent of 'National Standards' as the only form of high stakes testing in primary schools, and results for every school to be reported to the Ministry in 2012, the inevitable league tables may come next. This would be a major and unwelcome change to the present low stakes assessment environment where Assessment for Learning (AFL) / formative assessment is the major focus of current policy. Going by international comparisons this present policy has served New Zealand well.

Low stakes and high stakes have little to do with the qualities of the assessment. They reflect the idea of the gambling 'stake'. High stakes assessment is no better than low stakes assessment. Low stakes assessment is no easier than high stakes assessment. There may be little riding on a student getting a mathematics item wrong, if there is the chance to relearn the material (low stakes). But there may be more riding on a school being reported as below standard if this prompts parents to remove their children, or if funding is linked to this outcome (high stakes). High stakes assessment is mainly for reporting individual or school-wide performance, giving entry to courses, making awards and so on. High stakes assessment is at odds with the most agreed aims of primary education. It is the form of assessment which contributes least to student progress, as generally there is little opportunity to use high stakes information for formative purposes.

Research is clear about the beneficial outcomes of low stakes /AFL procedures for primary students, particularly students who are struggling. Research is also clear that when high stakes and low stakes assessment are in competition the high stakes

If you wanted to spend your days putting out fires, you would have become a firefighter instead of a school principal.

Learn How to Work Less, Produce More, and Still Get the Job Done in a Sensible School Week with Malachi Pancoast, President, The Breakthrough Coach. It's one of the most practical – and liberating – programs you will ever attend.

UPCOMING PROGRAMS IN YOUR AREA:

- Tasmania, AU - August 2011
- Wellington, NZ - August 2011

the
breakthrough
coach

Management Development for Instructional Leaders
(904) 280-3052

To register, visit <http://www.the-breakthrough-coach.com/pages/calendar.php>

data tend to predominate. A consequence may be that reduced emphasis on AFL leads to a diminution of achievement for all, but particularly the students in the 'tail' of the distribution, or the very students the policy of National Standards is supposed to benefit most.

League tables for reporting National Standards are costly, time-consuming and largely redundant as National Standards will do little more than demonstrate the status quo. They will show that schools with many socially and economically advantaged students and better resources, achieve higher levels than schools with many socially and economically disadvantaged students and inferior resources. The same outcome would be evident if the present decile system was reported nationally. Increasingly, league tables are associated with the central control of schools by governments. They are inconsistent with a devolved system that invests more responsibility for governance at a community level.

If league tables report some literacy and numeracy assessments only, they show a restricted picture of what schools and students are achieving. No account is taken of other important curriculum areas such as spoken language, languages other than English, science, technology, health, or social studies. There is no recognition of the arts, physical skills, education outside the classroom, or special programmes schools may develop for their particular communities. A typical league table reports about one-third of curriculum outcomes at best. On this ground alone they are invalid measures of school outcomes.

If league tables are to do more than demonstrate the status quo they need to include sophisticated 'added value' measures to take some account of the attributes of the schools' differing

student cohorts and resources. Measures of 'added value' have met with limited success in the US, UK, and Australia and in these countries they have been applied to more reliable data than the likely New Zealand Standards data. There is little realistic chance that 'added value' indicators can be developed for proposed New Zealand data. The data for National Standards as presently reported on a four-point scale are entirely unsuitable for inclusion in national league tables because of the inferior quality of the 'Standards' themselves. League tables based on these will provide information of little utility to parents, teachers, students or the New Zealand government.

Providing their limitations are acknowledged, and they can be achieved in all curriculum areas, valid and reliable national standards may contribute useful additional information alongside the range of assessments currently used in our primary school. National standards are no silver bullet though. However poor quality 'National Standards' which are promoted for political purposes in the first instance, are not worthy of support. An acid test might be the level of use they would attract if they were voluntary, like every other assessment in our primary schools!

ABOUT THE AUTHOR

Cedric Croft is an independent consultant. Previously he was Research Director, Learning, Curriculum and Assessment, NZCER. His wide experience of school assessment here and overseas includes service on numerous national committees and the Executive of the International Association for Educational Assessment. Recently he has undertaken work in large scale assessment for UNESCO and the World Bank.

You Can Copy TV for Teaching

screenrights

Just take out a Screenrights licence through the STA and you can:

- copy documentaries, drama, movies, whatever you like
- record from pay TV, free to air TV, and radio
- download programmes legally available on the internet
- store, play and access programmes using interactive whiteboards, e-cast, Clickview and Digital Video Commander



More information

Freephone: 0800 44 2348

Freefax: 0800 44 7006

licensing@screenrights.org

www.screenrights.org/nz

Image credits: (L-R) *Captain Cook: Obsession & Discovery*, Screen Australia
Voyage To The Planets, Essential Media and Entertainment Pty Ltd