

Assessment in Change: Some Reflections on the Local and International Background to the National Certificate of Educational Achievement (NCEA)

JIM STRACHAN

Abstract:

This paper backgrounds a number of factors that have influenced proposed changes in New Zealand's senior secondary assessment system. It documents various moves which have been made over the last 30 years, refers to international practices, reviews and developments, and notes the ongoing reluctance to shift from traditional systems in this country. Research about assessment, changing educational needs, and growing recognition of the pervasive influence of high stakes assessment and the opportunities offered by assessment aligned to learning, are seen as further influencing factors. The article emphasises the need for change, contrasts New Zealand and overseas practice, and observes that while the current NCEA design is an improvement over the old, compromises have led to a model with some features needing review.

Educational testing will change more in the next 10 years than it has in the last 50 years.
(Educational Testing Service Annual Report (1990), cited in Nisbet, 1993, p. 145)

In the senior secondary assessment arena, the last ten years of the twentieth century in New Zealand have been a decade of debate and deep division (Allen et al., 1997; Lee & Lee, 2001). Moves for reform that began more than 30 years ago have gained momentum, in spasms, but resistance has increased in response, forcing them to be shelved repeatedly. No proposal has managed to appease all, despite repeated compromises by respected education leaders in numerous forums.

Proposals have sought to increase validity by improving assessment coverage, matching methods more closely to what is being assessed, and closing the gap between qualifications and their purposes. These have received rhetorical support, but in practice New Zealand has been reluctant to move from a system that limits curriculum coverage by being heavily examination oriented; that assesses students for national qualifications in each of the three final secondary years; and that is structured at each level to sort and label students as successes or failures, largely on the basis of end-point, time-bound examinations. Reasons include cost, the power of tradition, selective valuing of some skills over others, little understanding of the flaws and consequences of past practices, and the influence of those advantaged by the system in the past, who have little wish to see this pathway to privilege broadened. Until now, major reform has proved elusive; internal assessment has remained subordinate to examinations; and School Certificate, a qualification which outlived its original purposes 25 years ago, has continued its high stakes position.

But by the middle of the 1990s, the need for reform became too strong for such proposals for change to be shelved yet again. Towards the end of the decade, the design for a National Certificate of Educational Achievement (NCEA) emerged, and development began.

Progressive implementation is to see level 1 in place at year 11 in 2002, level 2 at year 12 in 2003, and level 3 and scholarship¹ in the final secondary year, year 13, in 2004.

There has been considerable argument, both for and against the proposed model. Although existing qualifications gave employers less information than they needed, the

Business Round Table, for example, has been active in its opposition (Irwin, 1999). To this end, it has commissioned reviews from overseas consultants, Smithers (1997) and Donnelly (2000). Concerns have been raised about:

- reduction in the amount of external assessment
- the move to assess against standards rather than ranking
- reduced discrimination between candidates
- the consistency of teacher judgements
- the mixing of vocational with academic studies
- fragmentation of learning
- international acceptability
- over-assessment and teacher workload.

This article reflects on the foundations that underpin the change in senior secondary qualifications in New Zealand. (For reasons of space, and to avoid repeating the work of other writers, it does not attempt to detail changes during the last 25 years. These can be found in the contributions of Openshaw et al. (1993); Lennox (1995, 2001); Allen et al. (1997); and Lee & Lee (2001), for example.) The article then considers the international scene, focussing on certification, reviews, and the use of standards, referring particularly to other countries that New Zealand aims to emulate. The accumulating findings of research affect qualifications reform, particularly the influence of assessment, consequences for teaching and learning, assessment for better learning, and the limited validity and generalisability of examinations. Finally, changing educational needs are seen as worldwide drivers of assessment change that have impelled New Zealand past the barriers that previously impeded reform.

A second article, yet to be published, will call for a better, more balanced approach to evaluating the NCEA. It will provide a comparison on the basis of *purpose* and *quality* indicators, to question whether there is a “balance of advantage” (Smithers, 1997; see also Hall, 2000). It concludes that while few see NCEA as the perfect solution, spawned as it is from political appeasement as well as primary purpose, it has a clear balance of advantage over its forerunners.

New Zealand Background

How recent are the reforms that caused the debates of the 1990s? These extracts are revealing:

“School Certificate – the final exam“

”Exam ladder being shaken”

“Goodbye school cert. – but not just yet”

“External examinations ... are founded upon distrust of teachers.”

Headlines like these in the 1990s would surprise no-one, as the ability of teachers to assess to national standards is questioned, as the assessment documents associated with the National Qualifications Framework (NQF) highlight the past dominance of external examining, and as NCEA policies signal that the nearly 70-year reign of the largely externally-examined School Certificate might be nearing its end.

But only the first quote is recent (QANews, 2000, November). The last expresses the view of the *Reichel-Tate Report* of 1925, cited in the *Thomas Report* (Department of Education, 1944). The other two are headlines from main centre dailies in the mid-‘70s, the first from the *Auckland Star* (Tully, 1974, January 19), and the second from *The Dominion* (Goodbye School Cert., 1975, February 10) as the reviews of the 1970s proposed changes to senior secondary assessment.

These excerpts highlight an often-overlooked fact – the issues and changes in

assessment practice now occurring did not begin with the emergence of the New Zealand Qualifications Authority (NZQA) and the NQF in the 1990s, or even with the Committee of Inquiry into Curriculum, Assessment and Qualifications (CICAQ) in 1986 (Lennox, 1995, 2001; Allen et al., 1997). Concerns about external examinations have been expressed for over a century: “[Students] work to pass, not to know. And outraged science takes her revenge – they do pass, and they don’t know” (Huxley, 1874, cited in Satterley, 1994). In New Zealand, the *Thomas Report* (Department of Education, 1944) and *Currie Report* (Department of Education, 1962) both record views about the negative consequences of external examinations.

By 1971, removal of School Certificate was official policy for both the New Zealand Post-Primary Teachers’ Association (NZPPTA) and the Secondary Principals Association.² As early as 1972-75, nationally respected authorities³ were recommending major revision to assessment practices in the senior secondary school, including removal of external examinations within five years (Education Development Conference, 1974) and replacement of School Certificate (Department of Education, 1976). Through the 1970s, the School Certificate Examination Board (SCEB) consulted and made proposals about the shape, assessment practices and future of School Certificate (Department of Education, 1972, 1974, 1975, 1978, 1979). These groups were not left-wing revisionists, not ivory-tower academics, not doctrinaire educationists, not barrow-pushing lobby groups, but respected education professionals of considerable standing (including NZPPTA representatives), consulting widely and working in informed groups which reflected a wide range of views. Views on examinations were not restricted to the education sector. Other reviews have shown that the public prefer standards to ranking. A survey of employers in 1981 (Lennox, 1995) noted the limitations of School Certificate, favoured standards over ranking, and supported teacher assessments using a range of methods; and a New Zealand Employees Federation submission to CICAQ proposed “abandonment of ranking or norm-referenced systems of assessment.” The advocacy for change was quite firmly established by the mid-1970s.

This answers two oft-repeated questions, posed as assessment proposals surfaced during the 1990s:

- Why do we need to change?
- Why is the change so large?

I would argue that we need to change for the reasons that persuaded education professionals in the 1970s – that existing qualifications did not fit the curriculum, and did not meet the needs of users or of the full range of students. And the change is large – not for the *procrustean* reasons referred to by Irwin (1999) as he criticised the Achievement 2001 initiative, but for the *procrastean* (i.e., procrastinatory) reasons that avoided the evolutionary change that should have begun in the mid-1970s and continued thereafter. A delay of 30 years resulting from inertia, politics, and lobby group influence has led to a situation so far behind that in other developed countries, that incremental change can no longer bridge the gap to where we ought to be. (See note 4 for the moves made by some of our economic competitors). Indeed, New Zealand must be one of the best international examples of the following syndromes:

- *redundancy* – “the extraordinary capacity of assessment procedures to linger long after their educational relevance has passed” (Piper & McGaw, 1991);
- *resistance to change* – “assessment systems show a remarkable resistance to change.... Each country has its own ‘assessment culture’, a body of practices ... which ... are strongly and often irrationally resistant to change” (Nisbet, 1993, p. 26);
- *the influence of personal capital* – “Educational qualifications are a currency, possession

of which bestows social status and wealth ... the contrary arguments (against qualification changes) ...were driven by the belief that those who are currently advantaged by the ways in which certificates are allocated should continue to be so” (Capper, 1997; see also Black, 2001; Little, 1996).

The wish to maintain advantage applies to not only individuals and their families, but also institutions such as schools, which for generations have used qualifications results to advertise their prowess.

Despite the strength of the 1970s advocacy, a desire to avoid major change was still alive and well in the late 1990s. The School Certificate award (NZSC), largely externally examined in year 11 (form 5), is a good example. The 1962 Currie Commission’s perception of “the emergence of a cultural norm” in the form of NZSC helps to explain its tenacious hold on form 5. In 1997, a review of senior secondary qualifications, *Te Tiro Hou* (Allen et al., 1997), found it politically expedient to add an alternative to its preferred qualifications proposal, retaining School Certificate, while at the same time echoing earlier reviews that it “serves no particularly useful purpose.” Smithers (1997) takes a similar position as he skirts delicately around the issue of School Certificate, referring to “the New Zealand psyche” and suggesting that “a well-respected qualification at the end of year 12 might render the School Certificate unnecessary.” Irwin (1999) cites a report in the *Education Weekly* of October 11, 1999 which claims “87% of teachers who responded to a national survey wished to retain School Certificate.”

The earlier struggles of the school-assessed (and potentially more valid) Sixth Form Certificate to gain status in the face of the longer established and exam-reinforced University Entrance is another example of resistance to needed change. Indeed, there was some pressure to reinstate a nationally available form 6 examination, after the University Entrance award was removed from form 6 in 1986. This strong opposition to significant change for such a long period, while other countries altered their own examination systems, speaks of reluctant and selective listeners, great complacency, and considerable inertia. One can only speculate about the social and economic effects in New Zealand of a senior secondary education system so constrained by the limited range of attributes assessed by examinations, and so wedded to practices that labelled a majority as failures.

International Practices

In arguing against proposed changes, and preferring refinement to redesign, international practices have been called in rebuttal of the NCEA. I question whether they support the contentions of those resisting change. The certification practices, secondary education reviews and moves towards standards and criteria in countries comparable in culture and development to New Zealand are illuminating.

Certificates

In most developed Western countries, exit level certificates use wide-ranging external assessment administered by a central authority (usually in combination with a school-assessed component, which may be of the order of 50%). In these countries, there is not a great deal of nationally-controlled assessment before exit levels – certification in the penultimate year, such as Sixth Form Certificate (SFC) in New Zealand, is rare. (See Strachan, 2001 for a comparison of international practices at middle and upper secondary levels in the late 1990s.)

Other post-colonial countries, Canada and Australia, ceased awarding mid-secondary certificates some years ago.⁴ Instrumental reasons would appear to be a wish to:

- encourage further education
- diminish the negative effects of external assessment
- remove the expense of a practice whose purpose was past, and
- cease sending signals that the end of compulsory education was an appropriate school leaving point.

The few developed Western countries still externally examining and grading students across the curriculum at the end of lower or compulsory secondary education are now questioning the practice (e.g. Ireland: Department of Education and Science, 1999).

At the New Zealand year 11 level, the only certification regime that can be argued on the basis of current international practice is school-based assessment directed towards a completion certificate, perhaps accompanied by some form of national skills assessment. This has been the recommendation of review committees since the 1970s, and obviously questions the persistence of external examination and national certification at NCEA level 1 – and even possibly the extent of external examination at level 2!

Reviews

Many countries are currently engaged in revising their secondary assessment systems. "In virtually every country ...the assessment system is undergoing change" (Council of Europe Secretariat, 1993). The following excerpts from a Council of Europe newsletter describe the situation in Europe in the early 1990s. They come from a secondary assessment and examinations workshop attended by 31 countries (Tynjälä, 1993).

Some countries have national standardised upper secondary school final examinations, while others leave schools free to organise their own examinations. Some countries have quite broad requirements with 7-8 subjects, while only 2-3 subjects are required in others...

"Backwash" effects of assessment can result in defined learning objectives and actual teaching being in contradiction with each other, teachers tending to stress academic contents required in final examinations rather than other desirable goals...

In most European countries there is a demand to change education systems including examinations and assessment procedures. (Sweden and Finland) are moving from centralised administration to a decentralised system with more local and school autonomy. This is also the case in other countries....

In assessment systems there is a European trend to pass from norm-referenced to criterion-referenced testing and to combine elements of internal assessment with elements of external evaluation....

The USA, England, Scotland, Ireland, Israel, Ontario, NSW, Victoria, South Australia and most recently Western Australia have had reviews that are at various stages of action. Few reviews have much to say on assessment and certification before the leaving certificate year – most certification at the end of compulsory schooling (or end of lower school) is of low stakes and is already internal. An exception is the Republic of Ireland, which now wishes to remove its lower secondary examination (equivalent to New Zealand's School Certificate) because of its restrictive effect on curriculum and its toll on resources that are needed at the exit level, while in England, the purpose of the GCSE and the extent of external assessment are also being questioned (Strachan, 2001, pp. 3-4).

Close to home, the first year of the new New South Wales Higher School Certificate has concluded, while Western Australia has published its very thorough review of post-compulsory education *Our Youth, Our Future* (Curriculum Council of Western Australia, 2002). In New South Wales, there was "strong support for the decision to introduce standards-referenced assessments ... nobody called for a return to an earlier system" (Masters, 2002). Usage by employers or tertiary institutions of the greater

information provided by the new course reports has yet to develop. A review of international best practice in outcomes-based assessment (Tognolini, Andrich, & Ball, 2001), carried out for the Western Australia post-compulsory review, notes international movement in the direction of outcomes assessment – easier in New Zealand than in countries that lack a common national curriculum. The review observes overseas acceptance of a mix of vocational and general education, and supports the mixed school/external assessment proposals in Western Australia – “the combining of school-based assessment and external examinations adds reliability and validity to assessments and tends to engender support both from educators and the lay public.” As in New South Wales, the proposals for moderation invoke statistical moderation at first, moving later to peer review, suggesting “both approaches to moderation (statistical and peer review) at least as an interim measure in order to ensure fairness in the short term while the standards become better understood by teachers.”

Reviewing international practices and initiatives makes it quite clear that New Zealand is far from alone in initiating change. Nor is it alone in the directions it seeks to follow. While it may have been in the vanguard of change in its achievement-based assessment initiative in the late ‘80s and early ‘90s (Steer & Davidson, 1989; Codd et al., 1990; Crooks, 1990; Gilmore, 1990), it is now one of many countries seeking to match its assessment systems more closely to its educational targets.

Some commentators (e.g., Irwin, 1999) have argued that New Zealand should follow established international best practice in its move to revise senior secondary qualifications. This might be commendable if international secondary qualification regimes showed uniformity, but the international scene seems characterised by diversity rather than conformity, and the conclusion from reviewing international practices is that no system is so compellingly successful that many others are persuaded to copy it – “there is no one best system of assessment” (Noah, 1996; see also Nisbet, 1993, p. 137). Classic illustrations of diversity occur in Canada and Australia, where adjacent jurisdictions in Ontario/Quebec and Queensland/NSW differ fundamentally, with the former member of each pair having decades ago adopted full internal assessment, while external examinations have dominated its next-door neighbour. *Learning and Achieving* (Department of Education, 1986) indicated, “In the final analysis, the Committee believes that New Zealand would need to develop its own system.”

While it is hardly a counter-argument, New Zealand seems to be blissfully blinkered from international qualifications practice. In having widespread external examination at form 5, in assessing externally at three successive secondary levels (pre-1986 and post-2002), and in failing a significant proportion of its students (and asking them to pay for the privilege!), it is virtually unique amongst developed Western countries. Strangely, the extensive arguments *for* change available from overseas countries seem to carry less weight among those arguing international practice *against* NCEA, and wishing to retain the character of previous systems.

Standards and criteria

In education as in other fields, *standards* have been a catch cry of the 1980s and 1990s. Educational reform initiatives for improved achievement, including demands for accountability and reporting of what students know and can do, increased interest in assessment against standards during the 1990s. At the same time, there was growing recognition that the discrimination, ranking and selection approaches adopted for secondary qualifications in the mid-20th century had become less appropriate for the broader range of students now retained, and fitted neither the principles of valid reporting (Gifford, 1992, p. 5), nor the needs of a 21st century economy (Ball, 1991, pp. 66-73).

The move to assess against criteria rather than rank students against one another is

widespread and growing. The Ontario Royal Commission on Learning (Ministry of Education, Ontario, 1994) supported outcome-based education – “schools describe what students are expected to know when they graduate, and then rigorously assess their success in doing so. As one long-time Canadian educator commented, “This seems so eminently sensible that there must be something wrong with it.” The American Association for Higher Education (Angelo, 1995) defines assessment as “making our expectations explicit ... setting appropriate *criteria and high standards*; systematically gathering evidence to determine how well performance matches those standards” [italics added]. The combined Australasian senior secondary certification authorities have published formal quality guidelines that state: “High-quality curriculum documents ... provide explicit *criteria and standards* [italics added] for the assessment of achievement.” (ACACA, 1999, p. 18). Publications from jurisdictions as diverse as Denmark: “students’ level of attainment must not be assessed in relation to each other, but in relation to a number of fixed criteria”; (Ministry of Education, Denmark, 1993, p. 16), the states of Australia (e.g., Queensland Board of Senior Secondary Studies, n.d.; Curriculum Council of Western Australia, 2002, p. 68); plus others in Ontario, the USA, Scotland, Sweden and Thailand, all refer to assessment against standards or criteria.

The most recent senior secondary initiatives in Australia – New South Wales and Western Australia – are both firmly fixed on standards-based assessment for their final-year certificates, with a substantial internally-assessed component (around 50%). “The new Higher School Certificate will use a standards-referenced approach to assessment and reporting” (Stanley, 1999, p. 1).

A move to assess against standards and criteria is hardly surprising. The policies that support a curriculum based on outcomes also support an assessment system based on outcomes. Indeed, if curriculum and assessment are not in harmony in this way, then research evidence is adamant that curriculum and learning will be damaged by the dissonance. The ranking (cohort-referenced) mark-based approach used in the past does not provide the information about student achievement sought by many users. One can only wonder whether the observation from industry that people can do more than their educators report, has something to do with the mindsets generated in learners by the ranking assessment process and the failure tags it awards (Ball, 1991, p. 81; Picot, 1991, p. 78). Industry has less freedom than education to take the view that “nobody is any good unless they are better than someone else” (Ball, 1991, p. 72). Perhaps educators should also have less freedom to take this stance!

Qualification users, who want to know what students can *do*, find little use for aggregated scores that hide evidence of achievements in areas of interest. “A 1991 survey by the Employers’ Federation of 400 New Zealand companies on the usefulness of School Certificate found that the qualification was unable to provide information on whether the student had the generic skills and competences required in the workplace” (*The Employer*, 2001, March, p. 6). Employers seek information on specific capabilities, and standards-based assessment is better placed to supply that.

Assessment against standards (in the form of criterion referenced assessment) has been part of assessment methodology for many years. Despite practical difficulties, the idea is too attractive to set aside. Indeed, the public has long believed that senior school assessments do represent standards: surveys repeatedly reject ranking and scaling practices, in favour of standards (Department of Education, 1979). Standards are clearly part of public expectations in national assessments. If the examination warts hidden by veils of secrecy and tradition had been exposed earlier, the public would probably be more receptive to – perhaps demanding of – the changes proposed (and politically side-stepped) for decades.

The Israel profile in the UNESCO Education database encapsulates worldwide

trends:

The Ministry is currently reforming its assessment policies and procedures. This reform is taking place at various levels of assessment and is based on current worldwide trends. It involves moving from centralized regulation to a more autonomous form of assessment. It also involves: moving towards exams that are a more integral part of learning and teaching; moving from assessing basic skills to higher-order skills such as reasoning and exploratory thought; moving from pencil-and-paper assessment towards a wide range of assessment techniques; reducing the proportion of test items that assess limited and restricted knowledge while enlarging the proportion of items stressing contextual understanding; and using criteria rather than norm tests for the matriculation examinations. Based on these trends, the matriculation exam format has changed, a national assessment test has been introduced, and performance assessment is being promoted and supported within the smaller unit of the school. (UNESCO, 2000, Israel Profiles)

Influence of Research

Accumulating research evidence and associated changes in thinking, coupled to changes in the needs of learners, economies and societies, has added to the pressure for change. This is illustrated by the frequency of change among European countries noted above. Gipps (1993, p. 1), among others, has observed: "assessment is undergoing a paradigm shift ... from a testing and examination culture to an assessment culture." A growing body of research provides information about:

- the influence of assessment
- the consequences of assessment, and
- how well traditional assessment systems perform the functions we expect of them.

The NCEA design has sought to recognise the powerful influence of assessment. The next section shows that that influence can have a beneficial or a negative effect on teaching and learning, and on student attitudes towards lifelong learning.

Influence of assessment

The influence of assessment on the level and direction of teachers' and learners' time and effort is one of the best-established findings of educational research (Eraut & Cole, 1993; Crooks, 1988). Where high-stakes assessment is involved (and sometimes the stakes are more perceived than real), "teachers and pupils alike will tend to work to the reality of requirements as outlined by the form of assessment rather than to the rhetoric of a statement of intent" (Nisbet, 1993, p. 26). The consequence is a narrowing of curriculum to focus on those parts most easily measurable by examination. Cohen (1990, p. 37) argues that coverage by "current assessment strategies is totally inadequate, reflecting at best about 10 per cent of learner development." This is particularly alarming in the face of evidence that generalisability of assessment results is much more limited than we once thought. It is no surprise that Cohen concludes that "currently used assessment strategies constitute a corruption of education, curriculum and of student learning processes" (p. 38).

Educators have long been aware of this, although the awareness did not produce a great deal of change in New Zealand. Further, it is not well appreciated amongst the public, including some employers, that existing assessment results tell little about many of the attributes desired for employment and further education. The consequence is that assessment results have limited validity for the purposes of their users. This fact tends to be hidden by aggregation and ranking, as opposed to specification of what the result represents, and is part of the reason for the widespread move to standards and criteria, illustrated by the NCEA.

Recognition of the influence of assessment has led to extension of the key criterion of assessment quality (validity) to include consequences of assessment – for instance, *consequential* validity (Messick, 1989), and *systemic* validity – the extent to which assessment has desired effects on education settings and systems (Fredericksen & Collins, 1989). This leads to the conclusion that it is no longer sufficient that assessment should cover its intended topics, nor even match its intended purposes. Its other consequences (intended or not) must also be considered. Positive or negative effects on teaching and learning and influence on attitudes towards further learning are examples of other consequences, as illustrated by Broadfoot (1994) and McGaw (1996): “anticipation of what the external examiner might do diverts students and teachers from the real process of teaching and learning” (p. 66).

One of the dominant themes at the 1996 International Association for Educational Assessment (IAEA) Conference was the attention being given by one country after another to the problem of how to modify assessment in order that it should facilitate rather than inhibit the learning needs of each country. Raising student achievement through assessment has prompted a strong statement from the Chief Executive of the UK Qualifications and Curriculum Authority: “The most important development in assessment is undoubtedly what is now called assessment for learning ... (which) ... has huge potential ... to raise student achievement” (Hargreaves, 2001a, pp. 3-4). This echoes the assessment for better learning theme of *Tomorrow's Standards* (Ministry of Education, 1990), and recognizes the work of Black and Wiliam (1998). A further comment from Hargreaves states “Assessment for learning marks the beginning of a revolution” (2001b, p. 1).

Another 1990s illustration of the influence of assessment is assessment-driven reform, which uses assessment as a powerful lever to drive the curriculum in desired directions (Resnick & Resnick, 1992; Noah, 1996). This movement recognises the inevitability of assessment influence, and aims to harness it by selecting assessment methods and content that will contribute to curriculum reform. “Assessment ... is driving this beneficial convergence (between curriculum, assessment and pedagogy)” (Hargreaves, 2001b, p. 1).

The negative effects of high stakes ranking assessment voiced by many researchers and reviews (e.g., Department of Education, 1986) are covered in Gipps (1993, Chapter 3), and summed up by Crooks in a 1988 review paper – “normative grading, with the social comparison and competition between students which accompanies it, results in undesirable consequences for most students” (Crooks, 1988, pp. 468), and by Taylor (1994, pp. 255) as she observes: “The real consequences of these tests is that they have affected what we believe about children’s capacities to learn.” Research tells us that not only “we”, but also children themselves, infer their ability and capacity to learn from these results, and continue with or disengage from education as a consequence.

Arguments later in this paper make it clear that we can no longer afford this disengagement, nor overlook the limiting effects of past assessment practices.

Examination effectiveness

Examinations have long had standing as the pre-eminent form of assessment. However, informed opinion continues to point out that current practices can neither justify this standing, nor achieve fairness in relation to their intended uses.

Examination results should be treated with more reality and less veneration. A number of reviews since the early part of the 20th century (e.g., *Thomas Report*, *Currie Report*, Education Development Conference) showed awareness of examination weaknesses, and the Committee of Inquiry into Curriculum, Assessment and Qualifications in forms 5 to 7 (Department of Education, 1986, p. 73) was “of the opinion

that many of the claims for national public examinations are overstated and difficult to sustain." Elley's reference to "our blunt instruments"⁵; the central finding of the report of the USA National Committee on Testing and Public Policy (NCTPP), cited in Gifford and O'Connor (1992, pp. 3-4, and 5), "that many current practices in educational and employment testing stand in the way of efforts to identify and develop talent"; and comments such as that from Satterley, "clearly, external assessment is too inflexible to provide the best estimate of a student's performance" (1994, p. 59); and Black, "the limited reliability and even more limited validity of externally set, timed and formal written examinations are not understood" (1994, pp. 191-203), all testify to expert concerns over the effectiveness of external assessment. However, the need for public acceptance has delayed change (Broadfoot, 1994; Black 1994). Nisbet claims "the main obstacles ... are connected with public attitudes and expectations" (1993, p. 143). But the need for assessment that better matches its purposes, clients and the needs of education, is now causing most Western countries to give increased weight to internal assessment (close to home, most notably in New South Wales and Western Australia).

There is not much evidence that examinations can justify the view that they outdo internal assessment – the Reichel-Tate Report of 1925 (cited in Thomas, 1944) argued that the case for external examinations was located more in a distrust of teachers than in their measurement of attributes desired by universities, and its forward-looking proposals for accrediting of University Entrance were echoed by others and finally introduced 20 years later. The common belief that examination marks are more effective in predicting future performance probably lies less in reality and more in tradition, marketing, convenience, and/or the mindset that since we have long presumed this, therefore it must be true. Or perhaps it lies in the well-attested (and potentially self-fulfilling) correlation with success in later written, time-bound, end-of-course examinations.

Several pieces of evidence suggest that internal assessment is at least as effective. Forty years ago, the *Currie Report* (Department of Education, 1962) noted the arguments around school-based accrediting for university entrance, and cited Parkyn's *Success and Failure at the University* studies (1959) in concluding that "there is no evidence that a change, for example, to a single external examination would produce a student body much more successful" (p. 211). Support for this view is presented in the USA by Gifford (1990), while a Royal Commission in Canada's largest province, Ontario, notes, "in the late '60s, the exams were discontinued and teachers' marks became the only basis for university entrance. That change was made in part because it was learned that teachers' marks predicted university achievement as well as the exams" (1993, Vol II, Chapter 11, p. 3). Hall, McMurray and Capper (1985) noted that the internally assessed Sixth Form Certificate was effective as a predictor of university performance, while Hall (1999, pp. 173-196) finds "strong evidence to suggest that teachers' assessments are at least the equal of external examinations." Satterley notes a study which concluded, "internal assessments, moderated externally, appear to be no less reliable than the traditional External Examination" (1994, p. 62).

A further factor encouraging increased internal assessment is the lack of generalisability of examination results. Smithers argues we should "avoid putting more weight on to a qualifications structure than it can bear" (1997, p. 80). Satterley (1994, p. 54) cites Nuttall (1987) as follows: "Generalisability of external assessments is very limited since controlled conditions and timed papers are about as unlike real life as you can get", and then goes on to argue, "They also fail to represent pretty well every other context in which what has been learnt can be demonstrated or applied." The National Committee on Testing and Public Policy in the USA (cited by Gifford, 1992) emphasises the point: "Test scores should be used only when they differentiate on the basis of

characteristics relevant to the opportunities being allocated” (p. 5). Recognition that assessments gain from being related to the context of learning, and that selection should use assessments having some relation to the tasks that will be required of successful applicants, has led to the *authentic assessment* movement (using assessment tasks that reflect the learning context, and link to future destinations).

As so often in the past, debates have paid considerable attention to *reliability* – often linked to national consistency of teacher judgements. However, it needs to be remembered that *validity* is the primary criterion – results must bear a significant relation to their intended uses before reliability becomes meaningful (Hall, 1999). And this apparently uncontroversial statement contains a key message – there must be *intended uses*. Despite the motives that seem to drive many teachers, pupils, parents and schools, the magpie syndrome of the accumulation of qualifications for their own sake has little to commend it.

In summary, research evidence – the large influence of assessment, its capacity to induce change, consequences for teaching and learning, the gains to be made from assessment that promotes desired learning, the limitations of examinations, and the limited generalisability of assessment results (particularly those obtained under unrepresentative conditions) – has contributed significantly to moves for qualifications reform.

The NCEA design has sought to take account of these effects by specifying standards, increasing curriculum coverage, giving weight to both internal and external assessment, and reporting achievement directly (as well as by rank). The qualification has the potential to overcome a number of negative effects from the past; it will be important to take steps to ensure that the potential is realised.

Changing Educational Needs

The last two decades of the 20th century have seen debates about education, qualifications and assessment around the world. The underlying themes in the debate include:

- changes in the nature of work
- technological change and the need for lifelong learning
- international jockeying for competitive economic advantage through education
- deeper understanding of learning
- recognition of the impact of assessment
- dissatisfaction with present assessment practices, and
- ongoing attempts to improve the quality of assessment.

The theme of learning as a basis for competitive advantage has driven international agendas. In an OECD paper, Sir Christopher Ball (1993) observes: “In the post-modern world of today and tomorrow it is learning that provides the key to prosperity.” An OECD Education Ministers communiqué (2001, p. 3) states “our goal is competencies for all – basic competencies on which other learning depends, and the high-level intellectual and social competencies on which full engagement in the knowledge society depends.” The United States national education goals require that “all students learn to use their minds well”, while the USA Secretary’s Commission for Achieving Necessary Skills, cited by Fredericksen and Collins (1996, p. 193) seeks “people who can put knowledge to work.” Related papers (Resnick & Wirt, 1996, p. 6) recognise “the need for a developed talent pool that is broader ... and deeper, in terms of the capabilities for thinking, problem solving and team work that are suited to the demands of high performance workplaces.” In the 1990s, the goals of the New Zealand Government for education and

training (Maharey, 2001) are consistent with this theme, and the NQF Assessment Principles include the statement:

Economic and social needs require a well-educated population and an increasingly skilled workforce motivated to continue learning. Employers and tertiary institutions look for a range of knowledge and skills. Assessment can serve these needs if it recognises the achievements of all New Zealanders, samples the full range of skills and knowledge required of learners, embodies description of performance, contains incentive for further learning and discards the assumption that a substantial proportion of the population cannot reach an acceptable standard. (NZQA, 1993, p. 1)

In their 1995 publication *Education for Europeans – Towards a Learning Society* (Cornélis, 1995), the European Round Table of Industrialists (ERT) sound “a cry of alarm” over education. They express concern (p. 6) over “an ever-widening gap between the education that people need for today’s complex world and the education they receive. Too many disillusioned young students drop out of educational systems through failure or rebellion, or come through with only minimal skills. This is a major economic and social concern.” They argue for “education directed towards the needs of the future, not the legacy of the past” (p. 11), which produces basic and interpersonal skills quite different from those that dominate traditional assessment strategies and the learning that they encourage and reward (pp. 9, 12-13, 22).

Such thinking in Australia, England and New Zealand has led to developments that highlight key skills, and aim to deal with the imbalance in

systems of education and training, typically designed in another age on the basis of three false premises; the idea that they should serve only an elite of fast learners; the belief that initial education can be sufficient; the prejudice that education and training are different in kind. (Ball, 1993)

The very thorough Western Australian review of post-compulsory education recognises such thinking as it proposes a secondary assessment and certification system built on principles which include explicit standards, flexibility that includes all students, vocational education and training as an integral part of certification, and encouragement of learning throughout life (Curriculum Council of Western Australia, 2002).

There is clear expectation, given this world-wide recognition of skill needs and ongoing learning, that people and countries which respond will find themselves advantaged over those which do not – though this expectation is not without debate (see Capper in Bowen-Clewley & Strachan, 1997). The ERT argues that the lack of such skills puts “at risk (economic) competitiveness and democratic ideals, and foster(s) unemployment and social marginalisation” (Cornélis, 1995, p. 14).

This theme of changed purpose and direction for education requires matching change in New Zealand secondary assessment and qualifications. The change is encapsulated in two themes. First, assessment must articulate with the wider educational menu needed to provide for the broader range of students now retained to the senior secondary school, and the broader and deeper range of skills that economy and society need them to develop – including team as well as competitive skills. And secondly, assessment must move from “gatekeeper to gateway” (Gifford, 1992, p. 5), to encourage further learning rather than to signal its end. This theme of using assessments more “for purposes of ‘selecting in’ than for ‘selecting out’” is one of four lessons for successful systems identified by Noah (1996) in comparative studies of assessment in different countries.

Conclusion

Synthesizing international trends, local initiatives and research findings leads to an

inevitable conclusion – substantial change in New Zealand senior secondary qualifications is essential and well overdue. Trends in countries that New Zealand seeks to emulate suggest that change should be towards increased validity, involving a wider range of assessment methods, more weight on school-based assessment, assessment referenced to standards rather than ranking, assessment better linked to learning, and results used in an inclusive rather than exclusive manner.

Internationally, the need for change in secondary assessment systems is widely recognised. “These issues ... are fundamental, and appear in various forms in all countries” (Black, 2001, p. 2). Moves to lessen the external examination load in the middle/senior school, and increase school-based assessment began decades ago, both here and elsewhere. They have steadily spread. But extensive change in New Zealand has been elusive, despite the single national curriculum that makes it easier to introduce change here than in federal countries with multiple jurisdictions responsible for multiple curricula (e.g., USA).

International research has strongly advocated changed assessment systems. It has recognised the mismatch between assessments and assessment purposes, between curriculum and assessment, and between the aims of learning and the effects of assessment. A wider range of assessments; assessments better integrated with learning; reduced emphasis on year-end, time-bound, written examinations; assessments and reporting better matched to their purposes; certificates awarded only where there is a purpose for doing so; have all been advocated. These are directions reflected in reviews established by one country after another as each seeks to improve its education system.

Both local and international research and reviews have also questioned the pre-eminent position traditionally accorded to external examinations. Their lack of coverage, their measurement error, the limited generalisability and misuse of their results, and their negative “backwash” effects argue the need for substantial augmentation by other information to achieve meaningful certificates. Use of widespread external assessment before exit levels is questionable at best; the practice is not widely adopted in comparable jurisdictions. At exit levels, a mix of school-based and external assessment adds validity, and is not inferior in terms of its quality.

Referencing assessment to standards and criteria is not new (though application to all studies is relatively recent). Criterion-referenced assessment has been with us for decades, and New Zealand began achievement-based assessment trials in 1987. The public expect and believe that nationally certificated assessment is based on standards. The ability to track a change in national achievement levels requires some form of standardisation: accountability initiatives have sought this sort of information about key achievement areas. Internationally, the rhetoric in this direction is strong, though the practice has some distance to go.

The gap between what the economy and society need, and what secondary assessment delivers, has widened. There is international agreement on this, and on what those needs are. All should continue their education, rather than a select group; the retention of students to senior secondary levels must be expected to change accordingly; and those students need learning and certification matched to their aspirations and talents. (A likely shortfall in available talent, and the need for measures to address this are not new – the Royal Commission (Currie, p. 209) called attention to the issue in 1962!) An assessment and certification system that encourages rather than inhibits is essential to upskill a much larger proportion of the population. If the widespread rhetoric that links economic well-being to levels of education is to be believed, then assessment practices that fail and discourage will slow the country’s progress towards its social and economic goals. Indeed, they have been doing so for some time.

Contrary to the views of some, proposals for certification change are not recent, not

isolated to New Zealand, and not based on the ideologies of NZQA. Rather, moves for reform in this country began over 30 years ago. In New Zealand and Australia, curriculum has adjusted to changing needs, but in New Zealand, assessment and certification have marked time. The momentum gained since the late 1980s has finally overcome the resistance and apathy that thus far has inhibited needed change (and in so inhibiting it, prevented the progressive and evolutionary approach that might have avoided the present turbulent upheaval).

All of these factors argue for a substantially changed senior secondary assessment system. The NCEA is New Zealand's solution. Have we arrived (as one of my colleagues remarked) at assessment's "promised land"? Not yet – the potential is there; the task now is to turn the potential into practice. Also, the compromises made to achieve acceptance of the NCEA model have retained some questionable features. An example is the continuation of a formal nationally assessed qualification at level 1, long after its purpose has gone. This has substantial teacher workload implications. (Indeed, one might say that we now have the "son of School Certificate"). And engaging our most expert teachers as moderators and examiners has significant opportunity cost. Can we afford this at three successive levels? Likewise, should the designs for the qualifications at levels 1, 2 and 3 be so similar when their purposes are quite different? However, there is no perfect solution, and NCEA has a clear balance of advantage over its predecessors.

The foundation for that claim will have to be taken up in another paper, but it is worth noting that the NCEA matches most of the principles listed by Smithers (1997, p. 75). And Tognolini, who recently co-authored a report on international best practice in outcomes-based assessment, has said: "When you go through the best practice indicators I think the model put up by New Zealand has all those features" (McCarthy, 2001).

What must be ensured is that the potential for improvement offered by the NCEA does achieve the aims sought by the change, and does not repeat the negatives of the past. That is the challenge for New Zealand. It will require evaluation and evolution – evaluation to identify areas needing improvement, and evolution to make improvements – and to make sure that we never again require a 30 year lead time to achieve needed change.

Notes

1. A scholarship award is proposed to recognise the achievement of the most able. It will be based on standards, assess level 3 material, demand higher level skills than those required in level 3 achievement standards, and be totally externally assessed.
2. Tate, O. (1971). *Some considerations of the School Certificate and University Entrance Examinations in New Zealand secondary schools*. Unpublished paper.
3. NZPPTA (1969), researchers at NZCER (Elley & Livingstone, 1972), Ministerial working parties (EDC, 1974; McCombs, 1976), National Advisory Council (ACEP, 1975), the New Zealand Planning Council (1979), and Department of Education, (1972, 1974, 1975, 1978, 1979). See Lennox (1995), Allen et al. (1997) for more detailed information.
4. In Australia, for instance, Queensland eliminated external assessment in 1972, Victoria discontinued its Intermediate Certificate in 1967, South Australia in 1968, NSW reduced its School Certificate to four externally examined subjects (with remaining studies school-assessed) in 1974, and Western Australia moved to a standards-based approach with teacher assessment in 1985, and issued its last Certificate of Lower Secondary Studies in 1993.

Elley (1984), visiting Canada on a Claude McCarthy Fellowship to review policy changes in high school examinations, observed that most Canadian provinces had abolished external examinations before the final level of schooling some years previously, and showed "no signs of reinstatement".

Most countries in Europe restrict widespread external examination to exit level certificates (Strachan, 2001).
5. From the title of an unpublished paper by Elley.

Glossary

ACACA	Australian Curriculum, Assessment and Certification Authorities
ACEP	Advisory Council on Educational Planning
BERA	British Educational Research Association
CICAQ	Committee of Inquiry into Curriculum, Assessment and Qualifications
DTEC	Department of Training and Education Coordination, New South Wales
EDC	Education Development Conference
ERT	European Round Table of Industrialists
GCSE	General Certificate of Secondary Education
IAEA	International Association for Educational Assessment
NCEA	National Certificate of Educational Achievement
NCTPP	National Committee on Testing and Public Policy
NQF	National Qualifications Framework
NSW	New South Wales
NZCER	New Zealand Council for Educational Research
NZPPTA	New Zealand Post Primary Teachers' Association
NZQA	New Zealand Qualifications Authority
NZSC	New Zealand School Certificate
OECD	Organisation for Economic Co-operation and Development
QBSSSS	Queensland Board of Senior Secondary School Studies
QCA	Qualifications and Curriculum Authority
QDG	Qualifications Development Group
SCEB	School Certificate Examination Board
SFC	Sixth Form Certificate
UNESCO	United Nations Educational, Cultural and Social Organization
Y11	Year 11 (Y10, 11, 12 in Australia refer to the same levels as Y11, 12, 13) in New Zealand

References

- Advisory Council on Educational Planning. (1975). *Directions for educational development*. Wellington: Government Printer.
- Allen, P., Crooks, T., Hearn, S., & Irwin, K. (1997). *Te Tiro Hou: Report of the Qualifications Framework Inquiry* (Commissioned by the New Zealand Post-Primary Teachers' Association). Wellington: NZPPTA.
- Angelo, T. (1995, November). What is assessment? *Assessment Forum (American Association for Higher Education Bulletin)*, 48(2), 7-9.
- Australian Curriculum, Assessment and Certification Authorities (ACACA). (1999). *A set of guidelines for the integrity, quality and long-term credibility of certificates of achievement*. Brisbane: Queensland Board for Senior Secondary School Studies.
- Ball, C. (1991). Content and process. In G. Hawke (Ed.), *Sharks and Splashes* (pp. 63-77). Wellington: The Printing Press.
- Ball, C. (1993). *Lifelong learning and the curriculum*. Paper prepared for OECD Project.
- Black, P. (1994). Performance assessment and accountability. *Educational Evaluation and Policy Analysis* 16(2), 191-203.
- Black, P. (2001). *Report to the Qualifications Development Group (QDG), Ministry of Education, New Zealand on the proposals for development of the NCEA*.
- Black, P., & Wiliam, D. (1998). *Inside the black box*. London: Kings College, School of Education.
- Broadfoot, P. (1994). Approaches to quality assurance in six countries. In W. Harlen (Ed.), *Enhancing quality in assessment* (pp. 26-52). London: BERA, Paul Chapman.
- Capper, P. (1997). Commentary on NZQA assessment stocktake. In E. Bowen-Clewley & J. Strachan (Eds.), *Report on an assessment stocktake* (pp. 71-86). Wellington: NZQA.
- Codd, J., McAlpine, D., & Hansen, S. (1990). *Achievement, assessment, and consensus: An evaluation of the Sixth Form Certificate Practical Art Moderation trial*. Palmerston North: Massey University.
- Cohen, D. (1990). Reshaping the standards agenda. In P. Broadfoot, R. Murphy, & H. Torrance (Eds.), *Changing educational assessment: International perspectives and trends* (pp. 32-51). London: Routledge.
- Cornélis, F. (1995). *Education for Europeans - Towards the learning society*. Brussels: European Round Table of Industrialists.
<www.ert.be/pdf/edu2.pdf>

- Council of Europe Secretariat. (1993, June). *Report on Workshop: Research into pupil assessment and the role of final examinations in secondary education*. Council of Europe, Strasbourg.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Crooks, T. (1990). *Evaluation of Sixth Form Certificate training in assessment developments*. Dunedin: University of Otago, Higher Education Development Centre.
- Curriculum Council of Western Australia. (2002). *Our youth, our future: Post-compulsory education*. Perth: Curriculum Council of Western Australia.
- Department of Education. (1944). *The post-primary school curriculum. Report of the committee appointed by the Minister of Education in November 1942* (Thomas Report). Wellington: Government Printer.
- Department of Education. (1962). *Report of the Commission on Education in New Zealand*. (Currie Report). Wellington: Government Printer.
- Department of Education. (1972). *Internal assessment for School Certificate*. Wellington: School Certificate Examination Board, Department of Education.
- Department of Education. (1974). *Internal assessment for School Certificate: No. 2*. Wellington: School Certificate Examination Board, Department of Education.
- Department of Education. (1975). *Internal assessment for School Certificate*. Wellington: Department of Education.
- Department of Education. (1976). *Towards partnership* (McCombs Report). Wellington: Department of Education.
- Department of Education. (1978). *Steering Committee report on the future of School Certificate*. Wellington: School Certificate Examination Board, Department of Education.
- Department of Education. (1979). *The future of ... School Certificate: What is your opinion?* SCEB Discussion Paper. Wellington: School Certificate Examination Board, Department of Education.
- Department of Education. (1986). *Learning and Achieving: Second report of the Committee of Inquiry into Curriculum, Assessment and Qualifications in Forms 5 to 7 (CICAQ)*. Wellington: Department of Education.
- Department of Education and Science. (1999). *Junior Certificate: Issues for discussion*. Dublin: Department of Education and Science, National Council for Curriculum and Assessment.
- Donnelly, K. (2000). *New Zealand's National Certificate of Educational Achievement*. Auckland: Education Forum.
- Education Development Conference. (1974). *Improving learning and teaching*. Working Party Two. (Chairman P. Lawrence). Wellington: Government Printer.
- Elley, W., & Livingstone, I. (1972). *External examinations and internal assessments*. Wellington: New Zealand Council for Educational Research.
- Employers and the NCEA. (2001, March). *The Employer*, pp. 6-7.
- Eraut, M. & Cole, G. (1993). *Assessing competence in the professions*. Brighton: University of Sussex, Institute of Continuing and Professional Education.
- Fredericksen, J., & Collins, A. (1996). Designing an assessment system for the future workplace. In L. Resnick & J. Wirt (Eds.), *Linking school and work: Roles for standards and assessment* (pp. 193-221). San Francisco: Jossey-Bass.
- Gifford, B. (1992). Introduction. In B. Gifford & M. O'Connor (Eds.), *Changing assessments* (pp. 1-7). Boston: Kluwer.
- Gilmore, A. (1990). *An evaluation of group and skill sampling for moderating Sixth Form Certificate English*. Christchurch: University of Canterbury.
- Gipps, C. (1994). *Beyond testing? Towards a theory of educational assessment*. Lewes: Falmer Press.
- Goodbye School Cert. (1975, February 10). *The Dominion*.
- Hall, C. (2000). National Certificate of Educational Achievement: Issues of reliability, validity and manageability. *New Zealand Annual Review of Education*, 9, 173-196.
- Hall, C., McMurray, S., & Capper, P. (1985). Sixth Form Certificate grade allocation. *PPTA Journal*, Term 1, 34-37.
- Hargreaves, D. (2001a). *QCA Annual Conference Speech 2001*. London: Qualifications and Curriculum Authority.
<[www.qca.org.uk/annual conference/david_hargreaves.asp](http://www.qca.org.uk/annual%20conference/david_hargreaves.asp)>
- Hargreaves, D. (2001b). *A future for the school curriculum*. London: Qualifications and Curriculum Authority.
<www.qca.org.uk/ca/14-19/dh_speech.asp>
- Irwin, M. (1999). *Achievement 2001*. Paper presented to Examining Assessment Conference, New Zealand

- Council for Educational Research, Wellington.
- Lee, H., & Lee, G. (2001). The National Certificate of Educational Achievement (NCEA): "Fragile – Handle with care". *New Zealand Annual Review of Education*, 10, 5-38.
- Lennox, W. (1995). *Advocacy, evolution and learning*. Paper presented to the New Zealand Association for Research in Education (NZARE) Conference, Palmerston North.
- Lennox, W. (2001, June). Where did the NCEA come from? *QANews*, 10-11.
- Little, A. (1996). Contexts and Histories. In A. Little & A. Wolf (Eds.), *Assessment in transition* (pp. 3-27). Oxford: Elsevier.
- McCarthy, C. (2001, October 12). NCEA opponents to be heard. *NZ Education Review*.
- McGaw, B. (1996). *Their future: Options for reform of the Higher School Certificate*. Sydney: New South Wales Department of Training and Education Coordination (DTEC).
- Maharey, S. (2001, December 10). Speech at the launch of Youth Training/Training Opportunities Review consultation document *Building Futures*, Wellington.
<www.executive.govt.nz/minister/maharey/futures/sp.htm>
- Masters, G. (2002). *Fair and meaningful measures?* Camberwell, Victoria: Australian Council for Educational Research.
- Messick, (1989). Validity. In R. Linn (Ed.), *Educational Measurement* (3rd ed.) (pp. 13-103). Washington: American Council on Education, McMillan.
- Ministry of Education. (1990). *Tomorrow's Standards: The report of the Ministerial Working Party on assessment for better learning*. Wellington: Ministry of Education.
- Ministry of Education, Denmark. (1993). *Good practice: Signs of quality in Upper-Secondary schools*. Copenhagen: Department of Upper-Secondary Education, Denmark.
<<http://eng.uvm.dk/publications/12good%20prac/good.htm#assess>>
[Marking and continuous assessment]
- Ministry of Education, Ontario. (1994). *For the love of learning: Report of the Royal Commission on Learning*. Queen's Printer for Ontario.
- New Zealand Planning Council. (1979). *The welfare state? Social policy in the 1980s*. Wellington: New Zealand Planning Council.
- New Zealand Post-Primary Teachers' Association. (1969). *Education in change*. Auckland: Longman Paul.
- New Zealand Qualifications Authority. (1993). *National Qualifications Framework assessment principles* (Unpublished paper). Wellington: NZQA.
- Nisbet, J. (Ed.) (1993). *Curriculum reform: Assessment in question*. Paris: OECD.
- Noah, H. (1996). One best system? In A. Little & A. Wolf (Eds.), *Assessment in transition* (pp. 88-95). Oxford: Elsevier.
- OECD Education Ministers. (2001, April). *Investing in competencies for all communiqué*. Paris: OECD Media Relations PAC/COM/NEWS(2001) 32.
- Openshaw, R., Lee, G., & Lee, H. (1993). *Challenging the myths*. Palmerston North: Dunmore Press.
- Parkyn, G. W. (1959). *Success and failure at the university*. Wellington: New Zealand Council for Educational Research.
- Picot, B. (1991). Commentary. In G. Hawke (Ed.), *Sharks and Splashes* (pp. 78-80). Wellington: The Printing Press.
- Piper, K., & McGaw, B. (1991, May). *Educational assessment for educational policy and strategy*. Unesco/IAEA Round Table, Nairobi.
- QANews*, (2001, November), pp. 4-5.
- Queensland Board of Senior Secondary Studies.
<www.qbssss.edu.au/moderation/Review.html#Verification>
- Resnick, L., & Resnick, D. (1992). Assessing the thinking curriculum. In B. Gifford & M. O'Connor (Eds.), *Changing assessments* (pp. 37-75). Boston: Kluwer.
- Resnick, L., & Wirt, J. (1996). The changing workplace. In L. Resnick & J. Wirt (Eds.), *Linking school and work: Roles for standards and assessment* (pp. 1-19). San Francisco: Jossey-Bass.
- Satterley, D. (1994). Quality in external assessment. In W. Harlen (Ed.), *Enhancing quality in assessment* (pp. 53-70). London: BERA, Paul Chapman.
- Smithers, A. (1997). *The New Zealand Qualifications Framework*. Auckland: Education Forum.
- Stanley, G. (1999, May). Securing their future. *Newsletter No. 14*. Sydney: New South Wales Board of Studies.
<www.boardofstudies.nsw.edu.au/archives/stfreview/stf_14.html>
- Steer, M. & Davidson, C. (1989). *An investigation into teacher panel moderation and the use of grade-related criteria associated with Sixth Form Certificate Geography*. Wellington: Department of Education.

- Strachan, J. (2001). *Certification systems in the senior secondary school – Some international perspectives*. Wellington: NZQA.
<www.nzqa.govt.nz/ncea/e_prof/pdf/certsys.pdf>
- Taylor, C. (1994). Assessment for measurement of standards. *American Educational Research Journal*, 31(2), 231-262.
- Tynjälä, P. (1993). Assessment and examinations. *Council for Cultural Cooperation of the Council of Europe Newsletter/Faits nouveaux* 3(93), 3-4.
- Tognolini, J., Andrich, D., & Ball, S. (2001). *Report of the consultancy to research international best practice in outcome-based assessment related to post-compulsory education*. Sydney: Educational Testing Centre, University of New South Wales.
- Tully, J. (1974, January 19). Exam ladder being shaken. *Auckland Star*.
- UNESCO. (2000). *World data on education: Israel 2000. Assessing learning achievement nationwide*.
<www.ibe.unesco.org/International/Databanks/Dossiers/mainfram.htm>
Dossier: Israel (Profiles).

Acknowledgement

Support provided by the New Zealand Qualifications Authority in the preparation of this paper is gratefully acknowledged. The opinions expressed and inferences drawn are those of the author. No claim is made that they represent the official position of the NZQA.

The author

Jim Strachan was a member of the Department of Education's Examinations and Testing Unit when it was formed in 1974, and he remained there through its various changes until 1990. During that time, he was involved in developing, operating, analysing, reviewing and improving a range of aspects of secondary assessment – mainly in School Certificate and Sixth Form Certificate. He was a founding member of the New Zealand Qualifications Authority, first managing the team responsible for assessment and moderation of the existing senior secondary and post-secondary national qualifications, then moving to an assessment coordination and advisory role, and finally to a policy development position. Since his (partial!) retirement in 1998, he has undertaken contract work in educational assessment, both for NZQA and elsewhere.