

ACHIEVEMENT 2001

MICHAEL IRWIN
POLICY ANALYST
NEW ZEALAND BUSINESS ROUNDTABLE

WELLINGTON
28 OCTOBER 1999

ACHIEVEMENT 2001 CONTENTS

| | |
|---|-----------|
| Introduction | 1 |
| Two initial observations | 3 |
| <i>Lack of consultation</i> | 3 |
| <i>The speed of development and implementation</i> | 4 |
| Two underlying problems | 5 |
| <i>The concept of a framework</i> | 5 |
| <i>The search for both uniformity and inclusiveness</i> | 7 |
| The assessment methodology | 8 |
| Non-state examinations | 15 |
| Outcomes | 16 |
| <i>An elastic concept</i> | 16 |
| <i>Can the concept apply to all learning?</i> | 18 |
| <i>The relationship of Achievement Standards to the curricular outcomes</i> | 18 |
| <i>The effect of outcome specification on pedagogy</i> | 19 |
| <i>Size of outcomes</i> | 20 |
| <i>Setting outcomes of 'equal' difficulty</i> | 21 |
| <i>The elusive goal of precision</i> | 21 |
| Costs and benefits | 22 |
| Conclusions on Achievement 2001 | 23 |

| | |
|--|-----------|
| Where to from here? | 25 |
| <i>First, do no harm</i> | 25 |
| <i>The purpose of school qualifications</i> | 25 |
| <i>What is to be assessed?</i> | 27 |
| <i>Which assessment methodologies?</i> | 28 |
| <i>Reporting results</i> | 29 |
| <i>Separation of the curriculum from assessment and qualification design</i> | 30 |
| <i>Can qualifications be combined within a framework?</i> | 31 |
| Conclusion | 32 |

ACHIEVEMENT 2001

Introduction

Achievement 2001, the policy for school qualifications, was announced by the former minister of education, the Hon Wyatt Creech, on 5 November 1998. There had been no prior public consultation on school qualifications except in the context of the submission process on the government's Green Paper on the National Qualifications Framework¹ (NQF) which was released in June 1997.

There was only a page and a half specifically on school qualifications in the Green Paper. It declined to be drawn into a debate about the merits or otherwise of school examinations and said that criteria for the registration of qualifications should not favour one assessment practice over another. The announced policy would establish an "inclusive" and "comprehensive" framework which would be "settled and stable". However, registration criteria for qualifications would include the use of "explicit learning outcomes" and "credit values" towards the National Certificate of Educational Achievement.

¹ Creech, W, The Hon (1997), *A Future Qualifications Policy for New Zealand: A Plan for the National Qualifications Framework*, Ministry of Education, Wellington, June.

The policy decisions as announced in November 1998 followed the general lines of the Green Paper. According to the announcement:

- *every* student will study towards the NCEA, and *all* learning in the senior secondary school will generate credits towards the NCEA;
- there will be four levels to the NCEA (typically 5th, 6th and 7th forms and scholarship);
- in conventional curriculum subjects credits can be gained by a mix of internal and external assessment which would be against "Achievement Standards" that will be developed by expert groups and will set the standards to be met. Achievement Standards will also recognise merit and excellence and at least half are to be assessed externally;
- School Certificate (SC) and Bursary would remain and form the external assessment component of the NCEA;
- unit standards will be employed for non-conventional subjects at levels 1 to 3 with the notion of 'merit' not being ruled out; and
- credits would be attached to Achievement Standards and some Unit Standards towards the award of the NCEA.

In announcing this policy² the minister described much of previous debate about the NQF as "very arid and rather pointless" and urged that it was time to "bury" it. He saw his new policy as a "well constructed middle path" and, while some "extremists" would still argue about it, one which the "great bulk of the sector [would] back".

At the risk of being seen as "extreme", I have problems with important aspects of the announced policy and propose to raise a number of difficulties. I would also point out that the majority in any cause can be wrong - if it is true that the "great bulk" of the school sector support Achievement 2001 then that fact alone doesn't make the

case for it. But I have no idea what measures if any have been taken by the ministry to ascertain the depth of support³. I would doubt if, in any case, many in the sector have been sufficiently well informed about it to enable them to make a responsible decision one way or the other. There is no official paper that I know that has been made publicly available and draws on local and international research, analyses the problems, identifies and evaluates the various ways of addressing them, and explains why certain options were chosen and not others and how the inevitable trade-offs were resolved.

I am all for a robust and enduring schools qualifications system but consider that it is far too important a matter to be decided by some sort of majority vote. Of course teachers must be actively involved and their knowledge and experience drawn upon. They must be persuaded of the merits of whatever the final decisions are to be if a policy which depends on their cooperation is to work. But teachers and others will want to be satisfied that whatever is proposed has been developed from first principles, from a sound theoretical understanding of assessment and of the pragmatics of the classroom including the many pressures on teachers and schools, and as a result of acknowledgment and credible resolution of the tensions and trade-offs that are inevitably involved in the construction of an assessment system.

I have a strong sense of déjà vu. Not long ago a totally unit standards-based NQF was being promoted with evangelical fervour as the indubitably correct solution to our educational problems. I also recall being told by a ministerial speech writer that I was the object of a reference in a speech by a previous minister of education to a 'Custer's last stand' being made against his NQF policy. I had never seen myself as a swash-buckling military figure before! In April 1996 it was announced that the NQF would be "broadened" to include whole qualifications. This and other changes effectively meant that the central unifying idea of unit standards had been abandoned, though they were misleadingly portrayed as part of the "natural

² Creech, W, The Hon (1998), "Launch of Achievement 2001 – School Qualifications for 16 to 19 Year-olds", Speech at Rangiora High School, Rangiora, 5 November.

³ A nationwide survey of secondary school teachers by a group calling themselves Concerned Teachers, reported in *The Education Weekly* of 11 October 1999, indicated that 87 percent of secondary teachers who responded to a national survey wished to retain School Certificate. A spokesperson for the group said that the survey was carried out "to verify earlier feedback showing major concerns about 'Achievement 2001'".

evolution" of the NQF.⁴ And here we are again being told that a system based entirely on achievement standards and unit standards is indubitably the best school qualifications system and only "extremists" might dare to differ.

The initial ministerial announcement and accompanying documents contained a number of ambiguities. But I think at the time most of us were too punch drunk by all the policy developments and debates, including assessment debates, to pay close attention. Perhaps we saw in the announcement whatever we wanted to see – norm-referenced exams for those who wanted them retained and standards-based assessment for those who preferred that form of assessment. 'Feel-good' statements that Achievement 2001 was a "middle path", "comprehensive", and employed the "best of what we have now" might have helped to anaesthetize our critical faculties. In any case it was close to Christmas, and there were other things to think about!

Two initial observations

Lack of consultation

My first observation is that the scheme as a whole has not been subject to public consultation. I understand that there are various advisory groups and, of course, expert groups for particular subject areas, but they are involved in implementation issues: the basic design had already been set.⁵ Certainly in some, but not all, respects the scheme is similar to that outlined in the Green Paper, but the discussion in that Paper was extremely brief and I am not aware that problems raised about it⁶ have been answered. I think the lack of consultation on the scheme is bringing its own difficulties, though I have to say that my experience with the consultation process used in the curricular area does not encourage me to think that consultation over

⁴ For discussion see Irwin, M D R (1997), "The National Qualifications Framework: Where to Now?", *ACCESS*, Volume 16, Number 2, pp. 14-26.

⁵ A questionnaire, to be returned by 28 October 1999 and sent to all schools by the New Zealand Post Primary Teachers' Association (PPTA) and the Ministry of Education, seeks comments on aspects of the proposed scheme – not on the scheme as a whole.

⁶ See Smithers, A (1997), *The New Zealand Qualifications Framework*, The Education Forum, Auckland, November; Irwin, 1997, *op cit*.

Achievement 2001 would have led to any significant change.⁷ Presumably the calculation was made that the scheme as announced would appeal to all reasonable people and further consultation would simply be a waste of time. After all, only "extremists" would oppose it. As it is, serious concerns remain and important details have to be worked out, and the devil is, as has so often been said in other contexts, to be found in the detail.

The speed of development and implementation

But now development is happening at great speed and schools must be preparing to introduce the new level 1 Achievement Standards. So my second observation is the incredible speed proposed for the development of the Achievement Standards and their introduction. No trialling is proposed; in effect, the whole 1999 3rd Form cohort is to be an experimental group for each of the three years of the scheme's introduction, 2001-2003. It suggests quite astonishing confidence in the new system – a confidence which is certainly not justified by the analysis in any of the papers I have seen. The proposed national testing for primary children is to be trialled, so why not Achievement 2001 which is surely at least as significant?

Quite apart from any design defects, the sheer extent of the development required is such that some inconsistency in policy application, unevenness in the interpretation of what is required and frequent revisions would seem to be inevitable given the compressed timeframe and lack of trialling. I have on my desk about a pile four inches high of Achievement Standards, but they are only for 15 subjects and only for level 1. With standards required at three grades for each of the four levels⁸ for each Achievement Standard and from 5 to 8 Achievement Standards per subject, it is easy to see how large the task is. If there are about 30 'conventional' subjects we are talking about 2500 sets of standards. Perhaps a third of these will be required at the beginning of the 2000 school year "so that teachers would have a year to consider

⁷ For discussion of the consultation process on the draft curricula see Irwin, M D R (1999a), "The Curricular Reforms – Are they Taking us Forward or Backwards?", paper presented to a course on contemporary education policy, Victoria University of Wellington, 17 July.

⁸ There is some debate whether 'scholarship' will be a separate level or will be awarded for superior performance at level 3.

programmes that would prepare students for assessment against the standards in 2001."⁹

Two underlying problems

The concept of a framework

Achievement 2001 represents yet another attempt to put flesh on the concept of a qualifications 'framework'. In 1991, in the very early days of the construction of the aborted NQF, John Codd, Don McAlpine and Jenny Poskitt wrote as follows¹⁰:

The concept of a 'framework' for national qualifications ... is one which has obvious appeal to policy-makers. It represents a unified, logically planned and systematic approach In theory, it is a concept with considerable plausibility; in practice, however, it may be extremely difficult, if not impossible, to achieve.

They went on to point out a number of tensions and conflicts, including those between:

- the standardised outcomes required for a framework and institutional sources of energy and diversity which standardisation might restrict or suppress;
- the inevitable constraints of any framework and the desirability of flexibility and diversity;
- the adoption of a single assessment methodology which is desirable for framework purposes and the fact that one assessment methodology cannot meet all certification requirements; and

⁹ National Certificate of Educational Achievement – Background for Consideration in Consultation, undated Ministry of Education paper, p 3.

¹⁰ Codd, J, McAlpine, D and Poskitt, J (1991), *Designing the Framework – a Position Paper on Problems and Issues in the Restructuring of National Qualifications*, Education Department, Massey University, Palmerston North.

- the requirement for a common structure of levels and credits which form a necessary basis for a framework and the very different types of knowledge and skills to be fitted within it.

To my mind this was very perceptive and much of what they said back in 1991 is as true today for Achievement 2001 as it was for the then emerging NQF. The paper listed some of the tensions and conflicts which need to be addressed openly and squarely in designing a framework, whether one for all qualifications, as with the NQF, or only for particular group of qualifications, as with Achievement 2001. They reflect epistemological constraints, the limits of assessment technology and the diverse purposes which qualifications have to meet. They have been commented on in many papers on the NQF and in submissions on the Green Paper. The basic argument for uniformity is simplicity. But if we seek to apply a uniform approach to assessment over a wide range of learning we end up with all sorts of distortions. Certainly what we were offered in the NQF and now in Achievement 2001 is far from simple.

In the final analysis a judgment has to be made about where the balance should be struck between the perceived benefits of a framework in terms of uniformity and the costs of trying to force diverse material and assessment purposes into one structure which cannot accommodate everything without the costs of at least some distortion. And some of those costs arise from the implicit denial that there is a valid distinction between academic and vocational learning – that learning can be an end in itself¹¹. I find very little hint of the existence this fundamental problem in the official papers available to me.

We seem to be going through a stage of our educational history characterised by a search for frameworks, clear learning outcomes, transparent standards and the like. I have to say I am bothered by it: we are collectively suffering from a severe case of *hubris* – the mentality that places us over nature and seeks to reduce all knowledge, understandings and skills to numbers, to levels, to credit values and so on, and which appears to deny the mystery, the uncertainty and the qualitative aspects of our human existence and of our search for the truth about ourselves and the rest of

the physical world. We are moving, dangerously in my view, from a view of education as open with unpredictable results to a reductionist view of education as closed and with predeterminable results.¹² If education is seen as closed and predictable, then we place it at the mercy of the social engineers; we have moved from education to indoctrination.

Of course, the pragmatics of the educational enterprise require some, often arbitrary, ordering of material. But we should always be aware of epistemological constraints and the limitations of assessment technology. This involves facing up to the tensions and trade-offs involved, being explicit about our judgments and being prepared to defend them. I anticipate that sometime in the future we will find that education is really much more messy than we thought and resistant to the nice neat categories and distinctions which our search for frameworks assumes exist. The trouble is that we can, in the meantime, do a lot of damage as we lop and chop to make education fit preconceived procrustean beds. In my view, our framework fixation has led to some poor judgments and is still doing so.

The search for both uniformity and inclusiveness

There is one particular dilemma which arose in the case of the NQF and which arises just as starkly with Achievement 2001. The dilemma¹³, in brief, lies in the tension between:

- a framework emphasising uniformity and thus interchangeability of component parts; and
- a framework emphasising comprehensiveness and inclusiveness.

Both types of framework have their problems and limitations. The original NQF was of the former type: the rigorous specification of unit standards as the common

¹¹ A point made by John Codd in "Higher Education and the Qualifications Framework: A Question of Standards", *Delta*, Volume 48, Issue 1, 1996, p 63.

¹² Irwin, M D R (1999b), "From Virtues and Vices to Passionate Values", paper presented to the Independent Christian Schools' Fellowship, Auckland, 19 August, pp 9-11.

¹³ See Irwin, 1997, *op cit*.

building block was aimed at maximising comparability, portability, 'seamlessness', and the like. The 'broadening' policy announced in April 1996 and elaborated in the 1997 Green Paper proposes the latter type, though misrepresented as a natural evolution of the former and one that maintains all its assumed advantages. The problem is that there are trade-offs to be taken into account. The more uniform the system, in terms of its structure and methodology, the less inclusive it can be in terms of its ability to embrace different methodologies.

Achievement 2001 also wants to have it both ways. The Green Paper, as we have seen, proposed a "comprehensive", "inclusive" framework and one which can accommodate a variety of assessment methodologies, and yet it immediately went on to talk about "explicit learning outcomes" which clearly points to a particular type of assessment. The minister, in his speech announcing Achievement 2001, claimed that the new policy was "a well constructed middle path" which, like Tony Blair's impenetrable 'Third Way' in politics, steers a successful course between the divides set up by what are portrayed as the "arid" arguments of the past. But the chosen "middle path" aims at uniformity and will necessarily exclude other 'paths', thus limiting comprehensiveness and inclusiveness.

The speech also claimed that the new "system brings together the best of what we have now". We are not told what is the "best" of what we have now, the basis on which the evaluation and selection were made, and how these superior components are to be assembled. But whatever combination is in mind will necessarily exclude other combinations. Moreover, the "best" regime for one purpose may not be the best for another: the best assessment methodology for selecting entrants to a medical school is not necessarily the best one for certifying competence in typing. In short, while it is claimed that Achievement 2001 has resolved the tension between uniformity on the one hand and inclusiveness on the other, the words actually used in making this claim make it clear that the dilemma remains unresolved.¹⁴

The assessment methodology

¹⁴ A consultative forum established by the Secretary of Education met on 7 and 8 September to consider issues that had emerged. It recommended, *inter alia*, that the NCEA be "As **Inclusive** and **Flexible** as possible ... " and that it be **Coherent** (a

But perhaps the starkest and most serious problem is that of assessment methodology. As we have seen, the 1997 Green Paper was ambiguous. It emphasised inclusiveness and said that NQF registration criteria should not "be used to decide in favour of one assessment practice over another". It went on to say that "where options are available, secondary students could choose how they wished to be assessed". This seemed to mean that various forms of standards-based assessment (including standards- and achievement-based) and norm-referenced assessment, including forms of scaling, were all open for consideration and selection according to the material to be assessed and the purposes of assessment. But then in stating that registration criteria should include "explicit learning outcomes" and "credit values" it seemed to contradict the emphasis on "inclusiveness" and point to some form of standards-based assessment within a unified system.

The minister's speech¹⁵ launching Achievement 2001 in November 1998 went some way towards resolving the ambiguity: it seems clear that "inclusiveness" had been rejected in favour of uniformity. There would be *one* "middle path" for *all* students and *one* "unified" form of assessment. *Every* student was to study towards the NCEA and *all* learning in the senior secondary school would generate credits towards it. Moreover only "skills and abilities" were to be assessed – there was no mention of knowledge, understandings and such like, but this was before we were introduced to the heady notion of the "knowledge" economy. A media statement issued with the minister's announcement advised that the minister said "that teachers would no longer have a dual system of assessment for Unit Standards and exams". One of the government's objectives in "unifying" assessment was to "remove the status distinction between the standards-based and traditional examination approaches to qualifications assessment."¹⁶

But there were still some ambiguities and unacknowledged tensions. For example, the minister stated that for "regular" subjects SC and Bursary "will remain and form part of the external assessment component of the National Certificate". But how

unified system)" (emphases in original). The forum did not appear to recognise the trade-off between inclusiveness and uniformity.

¹⁵ Creech, 1998, *op cit*.

¹⁶ Qualifications Development Group (QDG) (1999), *Update*, 15 June.

could examinations, which are norm-referenced and employ forms of scaling¹⁷, remain along with Achievement Standards and Unit Standards within a "unified" qualifications system employing a single form of assessment? I note in this regard that Achievement Standards are to be "more consolidated, less atomised than Unit Standards" which is fine as far as it goes, but the different degree of specificity may require different forms of assessment.

An official brochure issued at the same time as the announcement¹⁸ said that "Examination marks will not be scaled" and as Achievement Standards are to replace examination prescriptions it is hard to see how SC and Bursary would in fact "remain" in any recognisable form. Moreover, success in some Unit Standards is to contribute credits to the NCEA: how is the assessment methodology for them to be unified with that of Achievement Standards and examinations? Further, the brochure assures us that Achievement 2001 would provide "results that show how well students have done against each other and against national standards" which seems to imply norm-referenced as well as standards-based assessment – unless, that is, relative performance is simply crude differentiation by three grades (credit, merit and excellence). Again, how does this square with the single "unified" system that "brings together the best of what we have" and avoids a "dual system of assessment"?

A ministry memorandum¹⁹ advises that "Achievement 2001 ... [is] to develop a "unified qualification" that [*inter alia*]:

¹⁷ There are various forms of scaling including inter-marker scaling to even out differences in marker toughness and inter-subject and inter-year scaling to even out differences in subject difficulty and in the difficulty of papers between years. Inter-year scaling is said not to be undertaken for the SC. However, the NZQA prepares a provisional marking guide, marks a random set of scripts, checks the number of students who would achieve each grade, and then adjusts the marking guide as necessary to avoid major variations between years in the percentages in each grade. Thus a similar distribution of marks is maintained over time. Given the impossibility of setting questions of equal difficulty each year this practice seems entirely sensible, indeed essential.

¹⁸ Ministry of Education and New Zealand Qualifications Authority (1998), *Achievement 2001 – Qualifications for 16 to 19 year olds*, Ministry of Education, Wellington, undated but presumably issued in November.

¹⁹ Ministry of Education (1999). An internal memorandum to the Secretary of Education entitled "The National Certificate of Educational Achievement – Paper 1", undated.

- may be gained by internal assessment (performance) and external assessment (including written examinations and externally assessed portfolios and projects); and
- unifies and removes the status distinction between the standards-based and traditional examination approaches to qualifications assessment.

It is also claimed to be "simple to understand", "easily manageable by teachers, and that it "reports information that is meaningful and useable" and which is "credible with the stakeholders". We are not told quite how all these very desirable objectives are to be achieved within a single, unified system. They are, however, to be achieved within a standards-based assessment system. It is hard to avoid the conclusion that the so-called "unification" of standards-based and traditional examination approaches is to be undertaken by expanding the former and abolishing the latter. For example, the authors of the paper:

- note that the government decided that "excellence and merit [would be] recognised through grading of achievement against standards";
- consider that "Qualifications are valuable only if they are meaningful and credible – that is, that the meaning of certification is transparent, and that the holder *really* knows or can do what is attested" (emphasis in original); and
- predict that the White Paper on the NQF will "widen and strengthen the NQF" and "continue" its "principles" and require "high quality" qualifications to "give accurate information about what a student knows and can do".

All this is very much unit standards 'talk'; it's hard to detect any elements of a "traditional examination approach" that remain in the "unified" system. Perhaps the view was that the credibility of standards-based approaches could only be assured by removing the competition of norm-referenced examinations.

It seems clear that a form of standards-based assessment – achievement-based – is to be adopted. However, we are not told precisely what it means and how it differs from competency-based assessment which is the form of standards-based assessment

employed for unit standards. Is it simply that the achievement-based assessment is graded whereas competency-based assessment is not? Will there be any conceptual difference as regards assessment methodology between Achievement Standards and Unit Standards if the latter adopt a merit level as has been raised as a possibility?

Terry Locke has analysed the transformation that takes place when the mode of standards-based assessment changes from an achievement-based one to a competence-based one as in Unit Standards.²⁰ Essentially the transformation involves the replacement of qualitative judgments by technician measurements. Such a change clearly poses considerable problems for assessment in a subject like English as, indeed, in other 'conventional' subjects at the senior secondary level. Moreover, Locke shows that the experts panel for English had (as at August 1999) moved towards the development of Achievement Standards in the Unit Standards competence-based format with quantitative 'clip-ons' for the merit and excellence grades (*ibid.*). He argues that in the draft English Achievement Standards for level 1, the proposed "basis for awarding merit and excellence ... is too inadequate to be seriously countenanced" and has serious pedagogical implications.²¹ I am not aware of any ministry response to his very detailed analysis which would appear to have considerable significance for its work on policy for school qualifications.

It would be useful to know why achievement-based assessment was rejected by the NZQA in 1993²² and what subsequent research has led to its adoption now.²³ If, as Locke's work suggests (*ibid.*), the development of at least some of the Achievement Standards in fact draws on a competence-based model rather than an achievement-based one (as, of course, might be inferred from the name of the project, *Achievement 2001*) how does this square with the ministry's acknowledgement of difficulties with

²⁰ Terry Locke has set up a website in connection with his English Study Design project to make available information on Achievement 2001 especially as it relates to English (<http://www.tmc.waikato.ac.nz/support/artslang/esd/2001forum.html>). The URL for his article "Achievement Standards: What's in a Name?" is <http://www.tmc.waikato.ac.nz/support/artslang/esd/achievestand.html>. The URL for his article "Lacks of Achievement 2001" is <http://www.tmc.waikato.ac.nz/support/artslang/esd/2001/EIAArticle.html>. An article entitled "Standards-based Assessment in English: Take 3" is pending publication in the *Waikato Journal of Education*.

²¹ In "Standards-based Assessment in English: Take 3" - see previous footnote.

²² NZQA (1993), *Framework Update*, August, p. 5.

Unit Standards which also adopt the competency-based approach? However, the ministry has provided no analysis of these difficulties which, had it been undertaken in a rigorous professional way, might have led to very different policy prescriptions to those embodied in Achievement 2001: the ministry's policy papers²⁴ rely more on self-congratulatory assertion and vacuity than on analysis and research.

The ministry's cavalier attitude to analysis and research and its appointment of many Unit Standards proponents to its expert groups suggest a continuing commitment to a Unit Standards-based qualifications framework. Its chief executive, Howard Fancy, has advised other departmental chief executives that the concerns about the National Qualifications Framework "centre less around the intent and general concept of the framework but more around implementation issues and its long term management" - this is simply asserted and is not the conclusion to rigorous analysis.²⁵

²³ See Peddie, R (1995), "Competency, Standards, Merit and Excellence" in Peddie, R and Tuck, B, *Setting the Standards*, Dunmore Press, Palmerston North.

²⁴ Papers obtained under the Official Information Act include the ministry's reports to its minister dated 23/3/98 (entitled "Qualifications for 16 to 19 Year Olds"), and those of 17/7/98 and 20/8/98 (both entitled "Qualifications for Young People Aged 16 to 19 Years") and the minister's undated paper to the Cabinet Committee on Social Policy entitled "Qualifications for Young People Aged 16 to 19 Years".

²⁵ Fancy, H (1997), memorandum to OCETE Chief Executives entitled "Senior Secondary School", 9 December.

On the issue of research, Warwick Elley observed in 1993 that²⁶:

... it should be stressed that educators in many other countries have tried to develop clear stand-alone standards in general subjects at the upper secondary level, but none have succeeded.

Cedric Croft commented in 1993 and 1994²⁷ to the effect that there was then little expertise in achievement-based assessment available within New Zealand and that little research into it had been undertaken. Some of the research that had been undertaken was not encouraging.²⁸ The only recent trial and professional evaluation of an achievement-based system at the senior secondary level has been that conducted in the context of the English Study Design (ESD) project, coordinated by Terry Locke and funded by Waikato University's School of Education. The ESD project was set-up to provide a coherent, rigorous, senior English programme, using standards-based assessment and national moderation procedures, as an alternative to Unit Standards.²⁹ It is not clear what, if any, lessons the ministry has drawn from it.

In 1995 Warwick Elley, Cedric Hall and I wrote an analysis of Unit Standards which concluded that a basic fallacy was that "one type of assessment (a particular form of criterion-referenced assessment) is suitable for assessing the myriad combinations and levels of skills and knowledge to be found in education and training courses and programmes."³⁰ Much that was said about Unit Standards would seem to apply to Achievement Standards. In 1996 Warwick Elley wrote an article discussing the particular problems that arise when trying to fit academic courses into the NQF. In it

²⁶ Elley, W B (1993), "The NZQA Agenda Unfolds", *The Christchurch Press*, date unknown, cited by Croft, C (1993), in *The Conflicting World of Standards-Based Assessment*, paper delivered at the 15th Conference of the New Zealand Association for Research in Education, Hamilton, 2-5 December.

²⁷ Croft, C, 1993, op cit; (1994), *Standards-based Assessment - What do we Know? What should we Do?*, paper Delivered at the 16th National Conference of the New Zealand Association for Research in Education, Christchurch, 1-4 December.

²⁸ For example, Gilmore, A (1991), *An Evaluation of the Sixth Form Certificate Moderation Trial*, Education Department, University of Canterbury, Christchurch.

²⁹ Hall, C (1999), *Evaluation of the Year 12 English Study Design Trial 1998*, Victoria University of Wellington, Wellington, February; Locke, T and Hall, C (1999), "The 1998 Year 12 English Study Design Trial: A Standards-based Alternative to Unit Standards", *New Zealand Annual Review of Education* 8, pp. 167-187, School of Education, Victoria University of Wellington, Wellington.

³⁰ Irwin, M, Elley, W and Hall, C (1995), *Unit Standards in the National Qualifications Framework*, Education Forum, Auckland, May.

he drew attention to a number of difficulties regarding the definition, assessment and moderation of standards in academic school subjects which, he pointed out, are "multidimensional and irreducible to a 'ladder of mastery' ".³¹

Where are the ministry's responses to the problems raised in these and many other papers³²? In what respects were these and other authors mistaken in their analyses? Is there some research of which they were unaware and which is such that the ministry can now confidently recommend the introduction of a standards-based assessment approach for most, perhaps all, school-based qualifications?

Particular questions arise about how exactly the external assessments are to be conducted. There is to be no scaling; assessment is to be against standards not against the performance of other candidates³³; Achievement Standards will replace prescriptions; and results are to attest to what the candidate "really" knows and can do - the intention seems clear enough. However, a complication has been introduced by a recent announcement that "the new qualifications system will give students a percentage mark".³⁴ The consultative forum recommended that "Externally assessed standards are to be reported with raw marks and a percentile ranking for the mark". It is hard to understand how such a requirement can be met within a SBA system unless a great many more standards are to be produced to differentiate much more finely between students than the three grades (or four grades if failure is counted as a 'grade') system allows. Perhaps we have here some political maneuvering in the face of opposition to the removal of external examinations - a move away from uniformity back to inclusiveness. But if so, what does this do to the framework?

³¹ Elley, W (1996), "Unresolved Issues in Fitting Academic Courses into the Qualifications Framework, *Delta*, Volume 48, Issue 1, pp 67-76.

³² A recent issue of *ACCESS* (volume 16, Number 2, 1997) was entirely devoted to articles about the National Qualifications Framework.

³³ However, standards usually have a normative basis ie they are based on what it can reasonably be expected that a student knows or can do at the particular stage or age in question. In the joint Ministry of Education/PPTA questionnaire recently sent to all secondary schools it is proposed (at Issue 3: Level of Difficulty) that standards for Level 1 credit be set so that a majority (70-80 percent) can be expected to achieve them. This clearly accepts the normative basis for standards. However, setting the 'norm' at a very low level of achievement to ensure a high pass rate will be at the expense of making a "credit" virtually meaningless and failure totally degrading.

³⁴ Smith, N, the Hon (1999), Address to the New Zealand Post Primary Teachers' Association Annual Conference, Wellington, 21 September.

It is also hard to see how the system could provide inter-year comparability, which is essential if the NCEA is to be credible and useful, without inter-year scaling, and there is to be no scaling of any sort. Moreover, if there is, in essence, to be two assessment systems, one for internal and another for external assessment, how does that fit with the decision to do away with dual assessment which is said to be one of the problems with our present system and one which Achievement 2001 will remove? We could ask the same question about the statement that the "new National Certificate [is to] be buttressed by external examinations to ensure consistency"³⁵.

There is no recognition that removing the so-called status distinction between standards-based and traditional examination approaches might mean using unsuitable assessment methodologies for some domains and for some purposes – that there may be a trade-off between equity and efficiency. It is, in any case, unclear what is intended. Does it mean that *All Must have Prizes*³⁶ to quote the title of a recent book ie that no student should fail and there should be few if any distinctions between achievement levels? It seems to me that any status distinction there may be originates in the world outside education and there is little that can be done within education to change it except to ensure that there are rigorous, credible qualifications for a wide range of school-based learning and not just for academic subjects.

I would note that some very high status professions depend on a very high level of knowledge which can be tested on a standards-based approach. I would want to be very confident that the surgeon taking a scalpel to my person knows exactly where *all* the arteries and nerves in the relevant area are located – not just 99 percent of them. Similarly, I want to be assured that the pilots who fly me know how to land and not just how to take off. But these highly developed vocational skills, which can be tested by a pass/fail, standards-based approach, are built on a very substantial foundation of knowledge and thinking ability which is often best tested by traditional methods.

Non-state examinations

³⁵ Smith, 1999, *op cit*.

There is also some uncertainty about whether private examinations would be registered on the framework with credits towards the NCEA. The ministry project manager is reported as saying that there was no reason why the International Baccalaureate (IB) shouldn't be recognised: "If a student meets a valid standard and it's a valid standard why shouldn't they [sic] get recognised".³⁷ This seems to imply that the IB will have to employ standards-based assessment against standards deemed valid by New Zealand authorities. Elsewhere it is stated that the IB and other "registered" examinations that meet the "quality standards defined for the NQF and the NCEA could also give a student credit towards their National Certificate".³⁸ However, there has been as yet no White Paper on the NQF, so Achievement 2001 is being developed out of the context of the NQF as a whole. We therefore don't know what the "quality standards" are to be. In any case, it is questionable whether the IBO and other overseas examining bodies would change their prescriptions and assessment methodologies to fit in with New Zealand requirements. Nor is it apparent that the award of NCEA credits would in any way enhance an internationally respected qualification such as the IB.

It would seem that New Zealand schools will not have the option of using private examination bodies that do not meet ministry criteria for registration on the NQF and that are not recognised by the award of credits towards the NCEA. All students - no exceptions seem to be envisaged - are to work towards the NCEA.

Outcomes

An elastic concept

The Green Paper on the NQF³⁹ said that:

... acceptable outcome statements could include 'purpose statements', 'objectives', 'process objectives', 'standards' or 'learning outcomes'.

³⁶ Phillips, M (1996), *All Must Have Prizes*, Little, Brown and Company, London.

³⁷ Cassie, F (1999), "Overseas exams may get the nod", *NZ Education Review*, June 18.

³⁸ QDG, 15 June 1999, *op cit.*

³⁹ Creech, 1997, *op cit.*, p. 21.

This suggests that the term 'outcome' has no clear boundary. There are various descriptions of Achievement Standards. An early pamphlet simply says that they "will specify what students have to achieve in each subject to earn credit towards the NCEA."⁴⁰ The Qualifications Development Group (QDG) states that the government has resolved that they are to be "transparent outcomes that students need to meet (ie based on agreed standards of achievement)".⁴¹ Thus the concept of outcomes continues to be elastic.

The draft Achievement Standards should be the subject of separate analysis. But a brief perusal indicates that at least some are not standards at all and might best be described a 'general aims'. A few level 1 assessment criteria taken at random are:

Health, Food and Nutrition:

- Students will be able to:
 - demonstrate a satisfactory knowledge of adolescent nutritional requirements. (credit)
 - demonstrate a thorough knowledge of adolescent nutritional requirements. (merit)

History

- Some ideas(s) or point(s) of view are accurately identified. (credit)
- Most idea(s) or point(s) of view are accurately identified. (merit)
- [The learner] accurately identifies the historical facts that are consistent with the evidence selected from the sources. (excellence)

Science (complete a practical investigation)

- Scientific plan chosen to collect data would be suitable for a trial and after discussion with the teacher (of a trial) the plan becomes appropriate for the purpose. (credit)

⁴⁰ Ministry of Education and NZQA, *Achievement 2001 – Qualifications for 16 to 19 year olds*, November 1999.

⁴¹ QDG Update of 15 June 1999, *op cit*.

- After some trialling or research the scientific plan should be appropriate for the purpose of the investigation. (merit)
- Merit PLUS the plan will include some discussion on the strength and/or limitations of the methodology chosen. (excellence)

Clearly these are not "standards": words such as "satisfactory", "thorough", "some", "most", "suitable" are wide open to interpretation. The actual standards, to the extent they exist on paper and not in teachers' heads, will not be "transparent" but will be within assessment guides and other material. However, these examples show something of the attempt by the expert groups, explored by Terry Locke in his examination of the draft English Achievement Standards, to dispense with qualitative judgments and to differentiate between the grades in technicist measurement terms – in these cases by the movements from "data ... suitable for a trial" to "some trialling", from "satisfactory knowledge" to "thorough knowledge", and from "Some idea(s)" to "Most idea(s)". Again, it is hard to see how they can form an adequate basis for the award of merit and excellence. There is no clear attempt to ask markers to discriminate between candidates in terms of *how well* they have performed in a given task – which is, of course, one of the main limitations of the competency-based approach of Unit Standards.

Can the concept apply to all learning?

The concept of learning outcomes is central to Achievement 2001 as it is for the NQF and indeed for the Curriculum Framework. The issue was discussed at length in the context of the NQF by Smithers⁴² and Hall⁴³, but the concerns raised by them do not appear to have been acknowledged, let alone addressed.

Smithers challenges the basic assumption, which is one on which the NQF rests, that all learning can be expressed appropriately as outcomes. He pointed out that school subjects are different ways of making sense of the world and that not all learning is organised as subjects: academic study, applied education and vocational training

⁴² Smithers, 1997, *op cit*, pp. 39-40.

⁴³ Hall, C (1998), *The National Qualifications Framework Green Paper – What Future for the Framework?*, in Livingstone, I (ed.), *New Zealand Annual Review of Education* (7.1997), School of Education, Victoria University of Wellington, Wellington.

have different organising principles. While it might well be important in, for example, occupational training to specify what is required in terms of outcomes, other matters such as relevant experience, ie training, may well also be important. In the humanities the point might be for students to have shared their insights and it is intrinsic to the experiences that students take away different things from them. In such cases the most authentic method of assessment might be to set common tasks (eg discuss the meaning of a particular play) and "by comparisons of the responses (rather than against pre-set outcomes) judge what the candidates have made of it" (Smithers *ibid.*).

The relationship of Achievement Standards to the curricular outcomes

The Achievement Standards are to be derived from the outcomes of the curricula developed for the various "essential learning areas" established by the New Zealand Curriculum Framework. The curriculum statements are to "specify clear learning outcomes against which students' achievements can be assessed" and also assessment procedures and examples.⁴⁴ It is not clear why we need to retain the curricula from Form 5 upwards if we are to have Achievement Standards which also purport to set the desired outcomes against which students' learning can be assessed.

The new curriculum statements have been assessed in Education Forum submissions and reports. They reveal a generally poor set of curricula ranging in quality from 'could do better' to downright appalling.⁴⁵ The new curricula do not form a good basis for the construction of 'standards' for qualifications. The lack of content in many of the curricular outcomes, their vacuity, and the student-centred approach to outcome determination raise serious doubts about the suitability of the curricula as a basis for summative assessment. It is not clear to me how 'curricula' which lack

⁴⁴ Ministry of Education (1993), *The New Zealand Curriculum Framework*, Learning Media, Wellington, pp 1 and 22-23.

⁴⁵ See Education Forum submissions and reports: Education Forum, *English in the New Zealand Curriculum – A Submission on the Draft* (1994); *Technology in the New Zealand Curriculum – A Submission on the Draft* (1994); *Social Studies in the New Zealand Curriculum – A Submission on the Draft* (1995); *Social Studies in the New Zealand Curriculum – A Submission on the Revised Draft* (1996); *Health and Physical Education – A Submission on the Draft* (1998); *Art in the New Zealand Curriculum – A Submission on the Draft* (1999); Howson, G (1994), *Mathematics in the New Zealand Curriculum – A Review*, Education Forum, Auckland, August; Kelly, P (1995) *Science in the New Zealand Curriculum – A Review*, Education Forum, Auckland, November.

content, activities and process can be readily converted into standards or examination prescriptions.

The effect of outcome specification on pedagogy

The change from content to outcomes as the basis for curriculum specification will, I assume, be having significant effects not only on *what* is taught but also on *how* it is taught – on pedagogy. The change represents a decisive movement towards "the why" and the "who for" and away from "the what" and "the how" ie away from the questions of "what content should be taught?" and "by what appropriate methods and procedures shall students acquire it?".⁴⁶ The pedagogical effects of this movement – good or bad – will be greatly increased in the senior school by replacing examination prescriptions specifying content by 'outcomes' – or, perhaps more accurately, by 'general aims' – in the form of Achievement Standards. This seems to me to be an important issue which remains to be addressed.

Clearly there can be positive aspects of an outcomes-based education (OBE) in so far as it seeks to make clear what is to be learnt and the stages in which learning can best be undertaken: indeed it makes the obvious and very important point that learning is expected to take place. But its suitability surely depends on what is to be taught. Where outcomes can be readily defined then OBE would seem eminently suitable – but this assumes content, a concept which is regarded by the government as outmoded.⁴⁷ Where the outcome is less easily defined, perhaps even ineffable, then clearly OBE is problematic. How, for example, could one specify as "transparent outcomes", and by grade and by level, the qualities that, it is hoped, education in the arts will foster – qualities that defy precise specification such as:

... sensitivity, delicacy, poise, elegance, imaginative insight that can go off on its own, the ability to respond in the appropriate mode, to paraphrase, to simplify and render open for inspection some of the deeper levels of significance that the surface of a work [of art] conceals – above all, to see or present all the many-sidedness of things in one

⁴⁶ For discussion see Education Forum (1999), *Arts in the New Zealand Curriculum – A Submission on the Draft*, Education Forum, Auckland, September, pp 49-51.

⁴⁷ The Green Paper entitled *Assessment for Success in Primary Schools*, May 1998, noted at page 7 that "In New Zealand, the shift from a content-based curriculum to an outcomes-based curriculum indicates [the] change of focus [from one on inputs to one looking directly at student learning outcomes]."

occasion, presentation or object - all the informed understanding and responsive appreciation that furnish the marks of progress and success in being educated in arts subjects, what Broudy (1971)⁴⁸ called the "enlightened cherishing" that comes along with, and is promoted by, an education in the arts.⁴⁹

Size of outcomes

As noted above, the size of Achievement Standards is to be larger than that for Unit Standards to avoid the problem of atomisation. A ministry of education pamphlet of 1 June 1999 says they are to be "broad" and many who were concerned about the application of Unit Standards to 'conventional' school subjects will welcome this. But at the same time it is difficult to see how one can have "breadth" and "transparency". Isn't there a tension here? A subsequent ministry pamphlet tends away from breadth and towards transparency when it states that "A reader [of a student's NCEA] will know *exactly* what a student knows and can do within a subject, and will know which subjects and what parts of the subject the student is very good at, or otherwise"⁵⁰ (emphasis added). This would seem to imply the incorporation of an enormous amount of detail into each Achievement Standard and the loss or diminution of any sense of overall achievement: it resonates with the sort of claims that were made of the original NQF.

I would also note that the rejection of all forms of scaling and the substantial amount of internal assessment envisaged must inevitably put considerable pressure on the Achievement Standards to be very specific in the interests of consistency and comparability of assessment and thus the credibility of the qualification. Internal assessment, for example, means that many different assessors in different schools will be employing different assessment tasks, and students will be assessed under different conditions and rules. Of course, moderation will aim to reduce the inconsistencies and achieve reliability, but there will be a trade-off between central direction in the form of increasingly precise standards and local autonomy in terms of decisions about the assessment tasks to be used, and when and in what manner they are to be undertaken.

⁴⁸ Broudy, H S (1971), "Enlightened Preference and Justification" in Smith, R A (ed) *Aesthetics and Problems of Education*, University of Illinois Press, Urbana, Illinois.

⁴⁹ Education Forum, 1999, *op cit* p. 79.

⁵⁰ Ministry of Education (1999), *Policy Directions*, issue 1, July.

Setting outcomes of equal difficulty

One of the least contentious results of assessment research is, I understand, that it is very difficult and maybe impossible to set assessment tasks of equal difficulty from year to year. Many factors are involved including the wording of the tasks, the options allowed and the order in which tasks are listed. And if the tasks themselves vary not just between years but between schools and if conditions under which they are taken also vary between schools then the achievement of comparability becomes even more difficult. Inter-year scaling is designed to address this issue and in most circumstances⁵¹ its results are entirely useful: it allows my Bursary result this year to be compared with Joe's results from last year and Jill's results next year. In Achievement 2001 we are to establish between 5 and 8 Achievement Standards per subject per level and all at the same level are, presumably, to be of equal difficulty. The same, of course, applies to the sets of standards to be established for each grade (credit, merit and excellence) at each level. Scaling to even out differences in marker toughness is also ruled out. So all this will add still further pressure on the Achievement Standards to be very precise indeed.

The elusive goal of precision

In the circumstances it would not be difficult for the "spiral of specification", about which Alison Wolf⁵² wrote in the context of the United Kingdom's National Vocational Qualifications (NVQs), becoming a serious problem. It is worth quoting what she said:

The 'standards' on which NVQs are based are intended to be so precise that they convey exactly what an assessor should look for. ... [T]his goal of precision has proved elusive. ... [T]his ever-receding goal of total clarity derives not from bad luck or incompetence, but is actually inherent in the methodology adopted. The more serious and rigorous the attempts to specify the domain being assessed, the narrower and narrower the domain itself becomes, without, in fact, becoming fully

⁵¹ Circumstances when it might not be useful include a very small student group taking the examination or when there is some significant change in the composition of the student group.

⁵² Wolf, A (1995), *Competence-based Assessment*, Open University Press, Buckingham, p 55.

transparent. The attempt to map out free-standing content and standards leads, again and again, to a never-ending spiral of specification.

The government's aim has been to avoid the difficulties encountered with Unit Standards including "issues of outcome definition" and of the "credibility of internal assessment"⁵³. I applaud the intention, but I am not convinced that Achievement 2001 manages to avoid those difficulties: indeed they have been perpetuated and transparency and credibility will continue to be elusive. It looks to me as if we will be revisiting the problems that have been already written about in the context of Unit Standards within the NQF.⁵⁴

Costs and benefits

I have found no detailed assessment of the costs and benefits of the new system. This should be undertaken and the results confirmed or adjusted by trialling before a decision is made on changes to present arrangements. In this case there is no analysis to speak of and there are to be no trials. What we do have is a great many questions and issues which await resolution.

It is hard to identify the benefits. The one unified system employing the best of what we have at present looks suspiciously like a slightly modified Unit Standards system, the deficiencies of which for much senior school work have been identified in several analyses of the NQF and Unit Standards.⁵⁵ The fixation with frameworks has driven policy towards a highly restricted system which will not accommodate much of the learning that takes place in the senior school and will meet few of the information needs of the users of school qualifications.

On the costs side, there would appear to be major increases in store offset to some small extent by reductions. But it is hard to be specific given so many uncertainties. Clearly the development costs are considerable and there will be costs on all secondary schools in familiarising themselves with the new Achievement Standards

⁵³ QDG Update of 15 June 1999, *op cit*.

⁵⁴ See, for example, Irwin, M D R, Elley, W and Hall, C, 1995, *op cit*; Elley, 1996, *op cit*; Irwin 1997, *op cit*.; Smithers, 1997, *op cit*.

and assessment procedures as well as, of course, with the new curriculum statements. There would appear to be substantial increases in internal assessment at Forms 5 and 7 with commensurate increases in administration and in moderation. However, given that the 6th Form Certificate is entirely internally assessed at present, it seems likely that externally assessing at least half will reduce teacher assessment workload at that level, though the reduction may be small because the linkage with the previous year's SC results amounts to a simple moderation procedure. As regards the external assessments, presumably the Achievement Standards will replace existing marking schedules or require new marking schedules to be developed. Presumably there will be three or four external examinations per subject (for each of the externally assessed Achievement Standards) and not just one.

If the new system provides qualifications which are not widely accepted nationally and internationally there will be very major costs. For example, some schools will decide that Achievement Standards are unacceptable and will face the costs of developing alternative qualifications systems or signing up to overseas ones such as the IB – provided, of course, they are allowed to do this, and this may require a change in policy which presently requires *all* learning to be towards the NCEA. Schools offering only the NCEA may face a substantial fall in the numbers of overseas students, and consequent loss of revenue from that source, if their parents lack confidence that the NCEA will be readily recognised by New Zealand and overseas tertiary institutions. Tertiary institutions and employers may face the costs of developing and applying their own cognitive assessments to provide the information presently provided by SC, the 6th Form Certificate and Bursary.

The costs on New Zealand students who find that their school qualifications are not readily understood and accepted are incalculable. Whatever new system is introduced, there will be the costs of discontinuity if it is not possible to readily compare the new qualifications with the present SC, 6th Form Certificate and Bursary. Finally, there will be the costs of further dislocation as problems within the new system arise and lead to significant modification or its replacement.

⁵⁵ See Elley, 1996, *op cit*; also Smithers, 1997, *op cit*. which lists such material, for example at p. 22.

Conclusions on Achievement 2001

I would summarise my conclusion on Achievement 2001 as follows:

- (i) it is not supported by any detailed analysis of:
 - the government's aims as regards school qualifications and of
 - the present difficulties and obstacles to be overcome in achieving those aims.

Nor is there any identification and evaluation of the alternative solutions, taking into account the inevitable tensions and trade-offs. It is being developed at enormous speed without trialling. It is a very high risk venture with a great deal at stake;
- (ii) the lessons from our own recent experience of the NQF and from experience elsewhere with forms of standards-based assessment (including achievement- and competence-based) have not been sufficiently heeded;
- (iii) the policy is being made 'on the hoof' as problems, which were entirely predictable and which should have been identified at the policy design stage, are being encountered. It is unsurprising that Achievement 2001 is riddled with contradictions and unidentified and unresolved tensions. At times it seems as if we are expected to be capable, like the White Queen, of believing "as many as six impossible things before breakfast"⁵⁶;
- (iv) there is no evidence whatsoever that Achievement 2001 will be a better alternative to what we have at present and that the considerable costs of development and implementation and of the dislocation caused among schools and teachers are outweighed by whatever benefits there may be. In fact it is hard to identify any benefits; and
- (v) development and implementation of Achievement 2001 should be put on hold pending a thorough, professional analysis of school qualifications

drawing on available research from within New Zealand and overseas. In the meantime present arrangements should continue. The last thing anybody should want is a major policy failure in this area which is so important to the life chances of so many of our young people.

In a recent analysis of the curricular reforms⁵⁷ I said:

... my overall judgment is that the curricular reforms of the last ten years have imposed enormous costs with few discernible benefits Given the apparent lack of thought that has gone into them, this is not surprising. I have little doubt that we will spend much further time and effort over the next decade in unraveling them.

It is disappointing to have to come to a very similar judgment on yet another major educational policy initiative and one which also affects so many people.

Where to from here?

Having come to these melancholy conclusions it is incumbent on a critic such as myself to indicate, at least in outline, the lines a further enquiry into school qualifications might profitably follow. I stress that the following is 'work in progress'. It does not seek to provide precise answers, which would require much more research and well conducted trials.

First, do no harm

A great deal is at stake with Achievement 2001, for the life chances of generations of pupils, for teachers whose morale is already suffering from an onslaught of policy changes, for school boards and administrators, and for the wider society. Any significant change in qualifications will have substantial costs. Surely the first and obvious point to make is that we should be very confident that the changes are worthwhile, that they represent a significant improvement on what we have now, and that the costs are clearly outweighed by the benefits. At the very least we should adopt the medical prescription: first do no harm. I would urge a cautious,

⁵⁶ Carroll, L (1971), *Alice's Adventures in Wonderland and Through the Looking-Glass*, Oxford University Press, Oxford, p 177.

evolutionary approach with well-researched changes being thoroughly trialled before final decisions are made.

The purpose of school qualifications

We should consider very carefully the purpose of school exit qualifications. Sometimes it seems to me from the official papers that uniformity and framework construction are ends in themselves. But this is unhelpful and likely to preempt several important issues.

An obvious but necessary point to make is that qualifications can only certify certain aspects in the cognitive domain – certainly not everything about a student and not even everything he or she knows or can do. Other methods of reporting are required for the professional judgments and intuitions of teachers about their students' attributes, non-classroom performances, personalities, dispositions and likely areas of success beyond school. This can include all sorts of information, for example about participation in school leadership, sporting and cultural activities, and supplemented by portfolios, certificates and such like.

Secondly, in talking about qualifications we are primarily concerned with summative assessment, not diagnostic, formative or evaluative assessment⁵⁸. The value of school exit qualifications lies in what can be done with them – in the doors and pathways they open up at the tertiary level or in employment. This necessarily requires that qualifications discriminate in terms of cognitive abilities and achievement levels.

The importance of distinguishing between students is enormously increased by the sheer numbers of students now sitting SC and staying on to the 6th and 7th forms. Having a SC or UB no longer indicates that the holder is one of an elite group at or near the top of the age cohort in terms of academic ability. Employers and tertiary institutions need to distinguish between the tens of thousands who now take the

⁵⁷ Irwin, 1999a, *op cit.*

⁵⁸ With more and more pupils staying into Forms 6 and 7, it can reasonably be argued that assessment in the 5th Form is increasingly becoming formative rather than summative.

secondary school qualifications. For example, in 1980 there were 8400 UB candidates. This grew to 19,400 in 1990 and to 25,500 in 1996.

Parents also need to be able to monitor secondary schools' performance in terms of the qualifications achieved by their pupils. For this they too will need methods of aggregating the results of individual achievement standards so as to compare schools' results. Of course, parents take other matters into consideration when choosing and monitoring a secondary school, but for most academic achievement is a key aspect.

With these huge increases in the number of qualifications holders, the qualifications will become stratified; even those held out by education authorities to be equivalent will be ordered by users within a hierarchy. This may be a complicated and messy business, but students along with tertiary institutions and employers will quickly learn to 'read' the system⁵⁹. Success in maths may, at least for some purposes, be seen as more important than apparently equivalent achievement in social studies. Externally assessed components may be seen as more credible and therefore rated more highly than those that are internally assessed. Users appreciate that underneath a norm-referenced examination result lies the notion of standards. Over time they learn to understand what such results indicate in terms of actual ability levels.

If users of qualifications can't select with confidence on the basis of school qualifications they will either institute their own examinations or use what they know about the schools the applicants attended – which is likely to be at best a rough and ready guide and to the disadvantage of able students attending less well-favoured schools.

What is to be assessed?

The next question is, surely, "What should be assessed?". I have already queried why we should have two sets of 'outcomes' – one in the curricular statements and another in the Achievement Standards – which both purport to set transparent standards against which students can be assessed. We should, I suggest, scrap the

⁵⁹ See Alison Wolf (1996), "Education: The Tyranny of Numbers", *Prospect*, December.

curricula for Form 5 and upwards. Whatever is prescribed for the qualifications will drive the curriculum anyway.

How should the qualification requirements be specified? The new curriculum statements and the Achievement Standards purport to be outcome statements (though may be more accurately described as statements of aims) which means that curricula are, in effect, left to schools to determine. But this creates a major problem for national assessment: if there is no common curricula how can we have common assessment tasks? In any case, our current obsession with outcomes ducks one of the most important questions in education – the choice of content, which, as Peters⁶⁰ remarked, is:

... between a range of activities that are thought to be worth passing on. Science, mathematics, history, art, cooking and carpentry feature on the curriculum, not bingo, bridge and billiards.

Content cannot, as seems to be implied in some of our new curricula, simply be left to student choice. As the Education Forum said in a recent submission⁶¹:

... explicit reference to, and specification of, suitable content are ... vital if the moral agency of the school as an educating institution and the professional integrity and disciplinary responsibility of the members of its teaching staff are to have any purpose and point of application to students' learning.

Thus from both philosophical and pragmatic points of view, a return to prescriptions specifying content would seem to be essential.

Which assessment methodologies?

Determining the appropriate assessment methodologies seems to me to be the next task. Surely it is quite unhelpful to address this issue before specifying what is to be assessed. And it is quite wrong to specify one "unified" methodology as if everything that needs to be assessed in the senior school is amenable to the same methodology. To put it another way, specifying only one methodology limits what

⁶⁰ Peters, R S (1966), *Ethics and Education*, George Allen and Unwin, London, p 144.

information can be obtained and what can be assessed – starting with assessment methodology is to start at the wrong end. It seems to me that if we are to assess the diverse knowledge and skills which are now taught and learnt in the senior school we will have to accept that there has to be dual (if not triple or quadruple) assessment approaches. And there may well be a trade-off here between manageability and the number of different methodologies that should, ideally, be employed.

If there are status issues to be addressed, I suggest they will be better addressed by ensuring that there are high quality vocational qualifications and alternative 'general' papers for those whose aptitudes and aspirations are unlikely to lead them to academic work at the tertiary level. Issues of status cannot just be wished away by prescribing one "unified" form of assessment – this is fanciful and unhelpful.

Where information about relative ability is sought, it seems to me likely that there will be a strong bias towards norm-referenced examinations, which will not always require written responses, employing common assessment tasks, which are externally set and marked, and undertaken under common conditions. The problem of cheating is largely avoided thereby and thus credibility ensured. There is also a substantial benefit for schools in that it avoids the heavy costs of internal assessment and moderation.

Where it is possible and desirable to assess against a set standard it may well also be important to employ external examiners to reduce the possibility of cheating and thus to ensure credibility. As a general rule the higher the stakes the more important is credibility, and thus the more important assessment using common tasks under common conditions by independent examiners becomes. In so far as summative assessment at Form 5 is decreasingly important as an exit qualification there is more of a case for internal assessment, and for those who take Bursary this applies for the 6th Form Certificate as well. However, students become better at examinations with practice, and the sooner they start the better.

⁶¹ Education Forum (1999), *The Arts in the New Zealand Curriculum*, Education Forum, Auckland, October, p 51.

We need to rescue 'scaling' from the list of educational swear words. Of course, the various forms of scaling can be abused, but they can also be very useful and sometimes essential if the aims of assessment are to be achieved. Although officially denied, a form of inter-year scaling continues with SC and for the very good reason that, as noted earlier, it is impossible to set questions of equal difficulty each year. Because of the aggregation of subject scores, inter-subject scaling continues in Bursary as does inter-year and, as necessary, inter-marker scaling. It is high time the forms of scaling were explained and their use defended.

Reporting results

What is reported and the manner of doing so must reflect the needs of the users of qualifications. What is presently intended is unclear, but it seems that it will involve reporting on 5 to 8 Achievement Standards per subject taken for Forms 5 to 7. A student taking an average of six subjects and staying on to Form 7 is likely to end up with reports on about 100 Achievement Standards. Users will not be able to cope with such numbers and will select, perhaps discounting all but the most recent results and perhaps only looking at some subjects and perhaps only the externally assessed Achievement Standards within them. In other words it seems likely that there will be an excessive amount of reporting on current proposals. What is needed for most subjects are fewer, aggregated scores – not a vast amount of detail which might well present a confusing picture of students' actual abilities and relative levels of attainment.

As a general rule, users of qualifications will want measures of students' abilities in activities in which they, the users, are most interested. According to Wolf⁶²:

If you are trying to predict whether people will be good at higher levels of academic study, academic tests are therefore appropriate. Conversely, if you are interested in vocational or professional work, or in practical skills, then they are almost certainly *inappropriate*. The evidence ... strongly confirm[s] the limitations of academic assessment as predictors of later performance. It also showed that assessments which mirror closely the activities and environment of a workplace are good predictors of workplace success.

⁶² Wolf, 1995, *op cit*, p 129.

This points to the different information needs of the users of qualifications. Universities, for example, will seek potential to undertake academic work at high levels. They will generally seek indications of relative ability and, especially in selecting entrants to high demand courses such as medicine, need fine distinctions in performance not simply three broad grades. On the other hand, some employers and tertiary institutions running more vocational or occupational courses may well be interested in evidence of success in particular activities which may be provided by some form of standards-based assessment.

For academic courses it will generally be that case that overall judgment by experienced, well-socialised markers and based on broad criteria are appropriate. For occupational courses evidence of success in quite precisely defined activities may be needed. One form of assessment will not suffice.

Inappropriately sized criteria will have unfortunate effects on instruction as well as on assessment. As Cedric Croft has noted⁶³:

Standards that are too specific may have an adverse impact on learning; but standards that are too broad probably lead to assessment judgments that are unreliable.

Assessment methodology and the manner in which the material to be assessed is specified are bound to affect instruction and learning. Qualification design should seek where possible to ensure that assessment raises achievement levels and not simply records what levels have been reached.

Separation of curriculum from assessment and qualifications design

One issue which has not been addressed in any of the papers that I have seen is the administrative arrangements for curriculum and qualifications, though it will presumably be addressed in the White Paper on the NQF. In my view, under whatever arrangements are devised, the tasks of setting curricula and of assessment for the senior school should be separated. The curricula should be set by subject working groups under the direction of a curriculum authority, and assessment cum

⁶³ Croft, 1994, *op cit* p 18.

qualification design should be undertaken by examination bodies. At present we seem to have a very unsatisfactory situation in which curriculum and assessment decisions, for example about which standards should be externally assessed, are jumbled up in one process with predictably muddled results.

If we effect this separation there would be no prior decisions about which methodology to use and about how much of each subject should be internally and externally assessed. Such decisions would be made by examination bodies based on what methods are best suited to the material to be assessed and the purpose of the assessment. I used the plural - examination *bodies* - advisedly because I do not think we should become locked into one qualification system. Only if there are two or more bodies offering their services can we allow innovation and see which bodies offer the kinds of services schools, parents, students, employers and tertiary institutions actually want. The 'one-best' system is a 'take-it-or-leave-it' system in which innovation is difficult and which forces students to take qualifications which may not suit their post-school aspirations. It is also one in which political considerations will always tend to trump educational ones.

Can qualifications be combined within a framework?

There is an increasing range of senior school courses as schools seek to make schooling at the senior levels relevant to a wider range of student in terms of their abilities, attainments and post-school aspirations. These necessarily result in a wider range of qualifications - no longer only in conventional academic subjects - and this raises a complication for framework designers. To what extent can diverse qualifications be sensibly aggregated so that the aggregations, and therefore the students, can be readily compared? Reducing everything to a common currency of credits is a superficially attractive solution. But qualifications are likely to be based on different assessment systems to meet different purposes and to certify different types of ability. And this makes the establishment of a common currency problematic - how, for example, could a standards-based Unit Standard be aggregated in any sensible fashion with the marks from a norm-referenced examination?

Users may not accept the credit rating attached to some qualifications and may even disregard some credits entirely from their calculations. In other words, education authorities cannot expect qualifications users to rate qualifications as they think they should be rated. It is better to simply list qualifications in cases where aggregations would be forced and artificial. To do otherwise is to risk bringing the system into disrepute. The notion that a nice, neat qualifications framework can be constructed in which all qualifications have a predetermined value based on a common 'currency' of credits derived from a single "unified" assessment methodology which will be accepted by all is, I suggest, an illusion.

What are needed in the senior schools are qualifications catering for the increasingly diverse abilities, attainment levels and post-school aspirations of senior students. At present the senior school is largely geared to the more academically inclined students. Quality programmes leading to quality qualifications in the applied and occupational areas are urgently required. At present there are only unit standards. The ministry clearly accepts the shortcomings of unit standards but does not propose any changes to improve them except the possible introduction of a merit grade. Achievement 2001 sticks to a non-differentiated curriculum framework and offers a non-differentiated school qualifications system. This is an enormous opportunity missed.

Conclusion

What I have tried to do in this paper is to analyse present proposals for school qualifications and to suggest an alternative way of thinking about them. In my view Achievement 2001 is a blind alley. It has been motivated by the search for a political compromise, seeks the elusive 'one-best' system for the assessment and certification of all school-based learning, and has been encouraged by the current fixation with educational frameworks. It sought to offer all things to all people but in fact offers a very restricted system littered with contradictions and unidentified problem areas. Moreover, it ignores substantial research that has been undertaken in New Zealand and elsewhere into forms of standards-based assessment, not least research into New Zealand's very recent experience with the National Qualifications Framework.

Further work is urgently called for, and Achievement 2001 should be put 'on hold' until it has been undertaken. An enduring, respected, credible school qualifications system will need to take seriously into account epistemological constraints and the limits of assessment technology. A system built on political compromise and not on sound educational foundations may have short-term attractions but will ultimately fail. While the current arrangements could, in my view, be improved (especially in the non-'academic' areas) they are certainly not so hopeless as to require the drastic revolutionary treatment envisaged by Achievement 2001. A cautious, evolutionary approach which identifies and addressed particular problems within an overall view of school qualifications is much to be preferred.

In developing that overall view, the further work required will need to ask 'first-order' questions such as the purpose of school qualifications and the information needs of their users. It should not be burdened by a prior commitment to a single 'framework' solution to multidimensional problems. Tensions and trade-offs must be confronted and not blurred. The further work needs to go on to consider what should be assessed and only then to ask what assessment methodologies are suitable given the material to be assessed and the information needs of users. I have suggested that an important issue here is the relationship between curriculum on the one hand and summative assessment and qualifications design on the other; in my view they are best separated. The issue of the extent to which qualifications can be linked within a framework is one of the last things to address not one of the first. The approach of Achievement 2001 is back to front.

In short, a great deal more thinking needs to be done before we can proceed with any confidence to reform or replace what we have at present. I hope that what I have had to say is accepted as a useful contribution to that further work.

Acknowledgements:

I am grateful to Glennys Allen, Cedric Croft, Cedric Hall, Roger Kerr and Terry Locke for valuable observations and suggestions on an earlier version of this paper.

Comments on this paper would be welcome and should be sent to:

Michael Irwin
New Zealand Business Roundtable
P O Box 10-147
WELLINGTON
Fax: 04-471-1304; Phone: 04-499-0790; Email: nzbr@actrix.gen.nz