

THE REFORM OF STUDENT ASSESSMENT IN NEW ZEALAND:
WHOSE AGENDA? WHAT EFFECTS

John Codd, Don McAlpine and Jenny Poskitt
Department of Education
Massey University
New Zealand

Paper presented at the AARE/NZARE Joint Conference,
Deakin University, Geelong, Victoria
22-26 November 1992

THE REFORM OF STUDENT ASSESSMENT IN NEW ZEALAND:
WHOSE AGENDA? WHAT EFFECTS?

John Codd, Don McAlpine and Jenny Poskitt
Department of Education, Massey University,
New Zealand

Recent educational reform in New Zealand has followed similar general directions to that occurring elsewhere. School administration has been decentralised; former bureaucratic structures have been dismantled and schools have become independent, self-managing units competing within a deregulated market environment. In this new environment, teachers are to be held accountable for learning outcomes, and students are to be regularly assessed in

relation to nationally specified curriculum objectives. Obvious comparisons can be made with the British Education Reform Act (1988) but there are also some important differences.

Unlike the British situation, education reform in New Zealand has occurred in relatively distinct phases. The first phase, referred to as Tomorrow's Schools, has been concerned with administration and is continuing with current attempts to restructure teachers' employment conditions. The second phase, commencing in 1990, has been concerned with curriculum and assessment reform.

During 1991, the New Zealand government launched its Achievement Initiative, calling it a key educational policy aimed at raising standards of achievement in the basic subjects of English, mathematics, science and technology (Ministry of Education, 1991a). In an earlier paper (Codd, McAlpine and Poskitt, 1991) we argued that the policy is fraught with internal contradictions concerning the nature and purposes of assessment. We identified within the Achievement Initiative "a fundamental tension between assessment that aims to improve learning in relation to the needs and abilities of individual learners, and assessment that constitutes a mechanism for centralised control and accountability" (Codd, McAlpine and Poskitt, 1991: 3).

As it was articulated throughout 1991, the Achievement Initiative appeared to be a policy that pointed in opposite directions. Thus, we described it as a 'Janus-policy' shaped by conflicting political imperatives. Some elements of the policy clearly pointed towards summative forms of assessment that would increase accountability, improve standards, and be used for school comparisons and national monitoring. We called this the "marketisation imperative" (Codd, McAlpine and Poskitt, 1991: 10). Other elements, however, pointed just as clearly towards diagnostic forms of assessment that would improve the quality of teaching, guide the learning process, and cater for the needs of individual students. We called this

the "professional imperative" (ibid). What we identified, in other words, was a policy for the reform of student assessment that was being driven by two conflicting agendas. It was perhaps predictable, therefore, that over the past year this policy would become the focus of a political struggle between proponents of educational professionalism on one side, and New Right advocates of accountability, managerialism and market forces on the other.

In the discussion that follows, we present an account of the changing character of current assessment policies within New Zealand and examine their potential educational effects. It is suggested that, to a large extent, the political agenda of the New Right in this particular area of educational reform may have been subverted and that recent initiatives open the way for major qualitative changes to the assessment practices of teachers.

Assessment Policy and the Crisis of Motivation

Hargreaves (1989) argues that government-led strategies for education reform are manifestations of wider social crises that direct the state towards particular kinds of policy responses. Thus, crises that originate within the cultural, political or economic spheres "reverberate" throughout the different parts of the state and in due course produce powerful effects in the shaping of educational policies.

This thesis is strongly supported by analyses of the reforms to educational administration in New Zealand (Codd, 1990; Codd, Gordon and Harker, 1990; Codd, Harker and Nash, 1990; Grace, 1990) where it is argued that these reforms were shaped by crises of political legitimation and economic management. Given a particular set of political and economic

conditions, it is argued, the New Zealand state responded with policies for the restructuring of educational administration that were deeply contradictory. In these circumstances, the policies contained fundamental contradictions between a democratic imperative for more community participation in decision-making, and an economic imperative for

tighter controls over public expenditure. These conflicting imperatives for devolution and control, although they originated in crises that were external to education, had produced crises that were in themselves educational.

In the context of state policy-making, as Hargreaves argues, educational crises

... take the form of a succession of critical points where a significant gap is perceived between educational policy and practice on the one hand and society's needs on the other; where existing solutions are seen to be exhausted or to have failed, or where new needs are felt to have emerged. It is under such conditions that support grows for new styles of state management, different patterns of initiatives; for a new overall strategy which promises to produce a closer match between schooling and society's needs. (Hargreaves, 1989: 103)

Drawing on the work of Habermas (1976), Hargreaves suggests that educational crises in the areas of administration, curriculum and assessment can be linked to wider social crises of rationality, legitimation and motivation. As defined by Habermas, a crisis of rationality reflects a widespread disillusionment with the administrative arrangements by which the state distributes goods and services. A crisis of legitimation reflects a breakdown of consensus in relation to the prevailing pattern of beliefs and values. A crisis of motivation occurs when there is a loss of confidence in the state's capacity to reward effort and encourage dispositions towards enterprise and achievement amongst its citizens.

This analysis provides a useful framework within which to locate assessment policy and relate it to other major areas of educational reform. Table 1 presents a diagrammatic summary adapted from Hargreaves (1989: 103).

Table 1
Crises in Education

Type of Crisis	Type of Crisis	Examples of Change	Locus of Response	of Policy
Rationality	Administration and reorganisation	Administrative (access and opportunity)	Educational provision	Self-managing schools;
Legitimation	Curriculum content (knowledge and values)	Curriculum aims and levels; Modularisation of curricula	Educational aims and levels;	National
Motivation	Assessment and accountability	Assessment outcomes (selection and allocation)	Educational technologies; New credentials; National monitoring of standards	New assessment

That the New Zealand state has been responding to a perceived crisis of motivation became clear early in 1991 with the publication of what is known as "The Porter Project". This project began in 1990 when Professor Michael Porter of the Harvard Business School, in collaboration with a team of economists and business consultants, carried out a detailed analysis of the New Zealand economy. The aim was to diagnose the causes of New Zealand's economic decline and prescribe remedies for recovery.

The Porter Project was funded mainly by the Trade Development Board and was supported by key government ministries and agencies, including Treasury and the Reserve Bank. Its report, entitled Upgrading New Zealand's Competitive Advantage, concludes that New Zealand's economic recovery would require "sustained and systematic change in our education system, attitudes towards competition, and prevailing management

philosophies, to name but a few" (Crocombe et al, 1991: 156). New Zealanders, according to Porter, are poorly motivated, inappropriately skilled and insufficiently competitive. The total absence of evidence to support such assertions does not deter the authors of this report from their graphic portrayal of a crisis of motivation, the responsibility for which is unequivocally laid at the door of the education system. Thus, in Porter's words,

There is a glaring mismatch between the skills needed to upgrade the New Zealand economy and those provided by our education system. (Crocombe et al, 1991: 161)

Although it avoids the extreme political rhetoric of the New Right, there can be little doubt that the report of the Porter Project is in accord with the New Right's agenda. It is significant, therefore, that during 1991 this report was to become "an ideological buttress for launching the Government's so-called 'achievement initiative'" (Peters, 1992: 3). When the Minister launched the draft National Curriculum of New Zealand in May 1991 at the Post-Primary Teachers Association Curriculum Conference, he made frequent reference to the Porter Project. Two months later, in his foreword to the 1991 Budget document, the Minister would state that:

Studies, like the Porter Project, questioned the relevance of our current curriculum with its excessive focus on social issues and poor preparation for the

competitive world. It confirmed other recent studies that show inadequate skilling in technology compared with other qualifications. (Minister of Education, 1991: 1)

There is no suggestion here that "confirmation" of such assertions ought to be based on evidence, nor are the mysterious "other recent studies" identified. Nevertheless, the Budget document would announce that:

Through re-ordering priorities in educational funding, the government will achieve its goal of maximising New Zealand's economic performance with a

more highly skilled, better educated, and more adaptable workforce.
(ibid: 3)

The government was clearly responding to the New Right's perceived crisis of motivation.

It was predictable, therefore, that student assessment would feature prominently in the draft National Curriculum of New Zealand and that the document would be couched in New Right rhetoric, with frequent use of such phrases as "competitive economy", "essential skills" and "educational standards".

The Achievement Initiative: Conflicting Agendas

From the outset, it has been clear that the Achievement Initiative policy is to be concerned primarily with the development of clear learning objectives and related assessment procedures in the "basic subjects" of English, mathematics, science and technology. This agenda has been strongly promoted by the Ministry of Education, with such statements as the following:

The Achievement Initiative policy aims to better meet the learning needs of all children by clarifying the desired levels of achievement to be attained in the various subjects. The development of levels statements will be followed by the development of assessment exemplars to assist classroom teachers in making appropriate judgements about the achievement levels of their students. The emphasis of the assessment will, therefore, be on good diagnostic monitoring which is formative and ongoing, and makes use of a wide range of proven procedures that are already well established in New Zealand schools.

(Ministry of Education, 1991c: 2)

There is confusion, however, when this statement is compared with statements in the draft National Curriculum of New Zealand, such as:

It is envisaged that assessment procedures will be developed to enable national monitoring of students' progress against clearly specified levels of

achievement
in the basic subjects. Information derived from national monitoring will
assist
with decisions on the targeting of resources.
(Ministry of Education, 1991b: 25)

This indicates that another agenda is also driving the Achievement
Initiative policy. It is an
agenda that has been pursued overtly since the mid 1980s by the New Zealand
Treasury and
by New Right pressure groups such as the Business Round Table.

In our earlier paper, we suggested that Treasury's advocacy for national
monitoring of the
education system is based in economics and reflects a general shift from a
demand-led to a
supply-side approach to the management of government funds. In education,
as in all other
areas of public expenditure, this means a shift of focus from inputs to
outputs. At the policy
level, there is a parallel shift away from an emphasis on improving the
amount and quality
of provision towards an emphasis on controlling the level of outputs (Codd,
McAlpine and
Poskitt, 1991: 8).

The Treasury's role in this area of government management is defined by the
Public Finance
Act (1989). This Act requires that Treasury assessments of public
expenditure be based on
the quantity, quality and cost of outputs. Another requirement of the
Public Finance Act is
the assessment of outcomes, which are described by a Treasury official as
"Government-
determined views on what is to be achieved by the outputs" (Greig, 1990:
55). For schools,
"outcomes refer to skills or other aspects of learning" (ibid).

Thus, within this managerial model of accountability, student assessment
becomes the
mechanism for measuring educational outcomes. But assessment also plays a
key role in a
market system of education. As Willis (1992) argues in a recent paper:

Within the context of New Right thought the key to improving education is
by
creating an educational market in which parental choice and competition
trigger

the drive to improved standards. In a market system of education, assessment plays a specific and central role because it can be used as a 'market signal' - a means by which the educational performance of teachers and schools, as well as students, can be judged. (Willis, 1992: 205)

It is this "market model" of accountability that New Zealand Treasury officials promoted strongly in their 1990 briefing papers to the incoming government. Here, they argue that the recent administrative reforms rely too heavily on "bureaucratic controls" and do not go far enough in promoting choice and competition. Central agencies, according to the Treasury, should be "seeking to improve the capacities of parents (and boards of trustees) to monitor performance, by making available comparative information, and establishing minimum requirements for assessment systems and methods" (Treasury, 1990: 134).

Thus, the ideology behind the New Right approach to student assessment is one which emphasises management accountability and market forces. It is an approach that conflicts deeply with a professional model of accountability in which assessment is seen to be a major responsibility of the classroom teacher, and where its form and frequency are matters of practical judgement rather than managerial imposition. It is very clear, as Willis (1992: 210) points out, that the New Right's approach to assessment is one in which "the needs of the learner may take second place to the requirements of the market or efficient management".

Although the Achievement Initiative may have originated in the state's response to a crisis of motivation conjured up by proponents of New Right ideology, the policies developed from it have been shaped by other agendas. Probably the most significant of these developments has been the move away from a comprehensive "high stakes" programme of national testing towards forms of assessment that fall within the province of daily classroom practice.

The alternatives for any system of national monitoring are clearly drawn by Broadfoot in the following way:

National assessment may be conceived of as a resource which teachers are encouraged, but not obliged to use, or it may be mandatory. Secondly national assessment may embrace only a small sample of students and schools or it may

require comprehensive coverage. In those countries, where testing is both compulsory and comprehensive, it typically becomes 'high stakes' exerting a powerful and arguably undesirable wash-back effect on the education system as

a whole.

(Broadfoot, 1992: 7)

That the Ministry of Education takes a similar position is shown in the following statement

made in response to Broadfoot's paper by the Manager of the Educational Assessment

Secretariat:

The paper, I believe, provides strong support for the careful and considered approach that New Zealand has been following in developing assessment policies for national monitoring and for the achievement initiative. This approach is based on the view that we can learn from overseas experience, that there are

pitfalls that we can avoid, that we should not repeat expensive mistakes, and that

we need to develop systems which take into account our unique socio-cultural context.

(Irving, 1992a: 2)

To a large extent, the move away from mandatory testing has been the result of concerted

efforts on the part of professional educationists within the Ministry of Education formulating

well researched and cogently argued policy proposals and presenting these to a Minister

whose respect and confidence they have gained. As Capper points out, the current Minister

of Education

... has observed the problems England has had with across the board testing, has talked to his officials, and has consequently amended policy on schools testing towards the entirely sensible model of providing teachers with diagnostic instruments which they can use voluntarily as a developmental tool for their own students.
(Capper, 1992: 21)

Thus, it appears that the Achievement Initiative policy is being shaped by conflicting political agendas and that under the influence of key Ministry officials its emphasis is shifting from centrally imposed national testing towards enhancing the professional practices of classroom teachers.

The National Qualifications Framework: A Counter-Movement

At the same time as the Ministry of Education was launching the Achievement Initiative and the National Curriculum, another agency of the state was advancing a policy that would have major implications for curriculum and assessment reform at the senior secondary and tertiary levels. In March 1991, the New Zealand Qualifications Authority (NZQA) released a discussion document about restructuring national qualifications entitled *Designing the Framework*. It proposed a unified, logically planned and systematic approach to the design and issuing of all educational and vocational awards, qualifications and credentials. The new form of assessment was to be "standards based" and the new credential was to be a National Certificate.

The New Zealand Qualifications Authority was established in 1990 as a quasi-governmental organisation with its own independent board. Somewhat paradoxically, it represents a strongly centralising element within the state's otherwise more decentralised system of educational administration (Selwood, 1991). Significantly also, it is

staffed by many professional educators who had held positions in the former Department of Education before it was restructured.

The National Qualifications Framework (NQF) embodies an approach to assessment that is intended to be flexible, responsive to local needs and facilitative of learning and achievement.

It is also to be a centrally administered system, with a pre-ordained structure of eight qualification levels and comprising measurable units of learning (or credits) that can be accumulated or transferred within and across the full range of credentials and qualifications.

While there are obvious tensions within such a rationale, the general aims of the NQF are educationally sound.

The discussion document announced that:

The framework will be a co-ordinated set of units of learning, available to students at senior secondary school, in polytechnics, colleges of education, universities or private training establishments, and in wananga, marae or

community agencies. Such learning must be equally available in the workplace, and through self-directed study or experiential learning. (NZQA, 1991: 36)

Experiential learning is defined as "learning acquired in the normal course of life and work rather than through organised education and training" (ibid). Standards are to be maintained through a national catalogue of units of learning. This implies a modularisation of all curricula at the senior secondary level, which goes much further than the system developed by the Scottish Vocational Education Council (SCOTVEC). Whereas the Scottish model has a similar national catalogue of modules (mostly of a vocational nature) it also involves considerable devolution of responsibility to colleges and schools for curriculum development and student assessment (McLaughlin, 1988). Academic curricula on the other hand remain linked to public examinations.

While the NQF represents a major curriculum innovation, it also raises many unanswered questions concerning the arbitrary compartmentalisation of learning, the fragmentation and simplification of complex domains of knowledge, and increased standardisation of curricula. These are similar to issues that have been identified in the Scottish context (Jonathan, 1987).

At the centre of the NQF is the new National Certificate which is to sit (most uneasily) alongside existing public examinations at the fifth form (School Certificate) and seventh form (Bursary) levels. It is claimed that:

The National Certificate will offer an alternative route to employment and further study, and will enjoy equal status with Bursaries. The Government does not intend that there will be first and second class qualifications based on the discredited distinction between academic and vocational. Both are equal in their worth. (NZQA, 1991: 32)

While this may express the hopes of the policy-makers, social realities may determine otherwise.

In the context of a review of 1991 curriculum reforms, Capper (1992: 23) comments that the NQF, taken as a whole, "has the potential of inducing a major revolution in which access to learning and certification becomes democratised as never before". The NQF, in Capper's view, represents a fundamental conceptual shift in the delivery of education that "has been made possible by technical advances in assessment methodology" (1992: 25). This is no doubt a reference to the substantial shift towards criterion-referenced assessment contained within the NQF, although considerable confusion surrounds this aspect of the policy (Codd, McAlpine and Poskitt, 1991).

There is a major tension within the NQF between standards-based assessment

(which is largely criterion-referenced) and norm-referenced assessment with which it is starkly contrasted (NZQA, 1991: 62-3). It is assumed that these forms of assessment are mutually exclusive, but any comprehensive system of educational evaluation would need to include both. The issue, which we address more fully in the next section, is not "which is better?" but rather, "when and how are they to be used?"

The proposed National Certificate also highlights tension between local curriculum initiatives and nationally standardised units. If "clear learning outcomes" are to be "assessed against national standards of performance" (NZQA, 1991: 30), and if "the majority of the units of learning will be common to the National Certificate and Bursaries", there will be very strong pressures towards a common curriculum.

Units of learning are designed to broaden the curriculum and increase flexibility, but the "smorgasbord" approach would have no way of ensuring overall curriculum structure. National standards could not be maintained when there is no basis for comparing one particular combination of units with another combination in terms of overall (i) cognitive complexity or (ii) logical coherence.

What we have in the NQF and the National Certificate are major educational reforms that are clearly being driven by professional educationists while paying lip service to the demands of the New Right for increased competition and improved national standards. Moreover, as with the Achievement Initiative and the National Curriculum of New Zealand, policies to establish

a new framework of educational credentials are being contested at the same time as they are being developed.

For teachers, students and the wider community, the current situation is confusing to say the least. Considering the complex, and unstable, socio-political environment described earlier, it is perhaps predictable that the direction of assessment reform in New

Zealand remains uncertain and that strong tensions persist within the current policies. We now turn to an examination of these tensions and their implications for assessment practices.

Two Forms of Assessment: Unresolved Tensions

The policy initiatives we have described have resulted in the promulgation of two different forms of assessment. The first form considers individual student growth in learning is largely formative, diagnostic and standards-based and is related to criterion referenced assessment. The second considers the national monitoring of standards - students, schools and systems - and is more related to norm referenced assessment.

The differences between these two approaches to assessment are best seen in terms of broad brush strokes rather than in sharp contrasts. Furthermore the architects of these two approaches to assessment in the current debate see them as complementary rather than in conflict. However, considerable tensions exist between the two approaches to assessment and they result in quite different educational and socio-political consequences.

In our earlier paper (Codd, McAlpine and Poskitt, 1991) we outlined some of the tensions, and in particular, concerns over proposals for the national monitoring of standards. It was claimed that "Extensive national monitoring and comprehensive testing at three different points from Standard 4, Form 2 and Form 5 cannot be justified in terms of student learning and national cost. If the aim is to monitor national standards", we argued, "then a system of 'light sampling' as advocated by the Working Party on Assessment for Better Learning (1989) should be adopted" (Codd, McAlpine and Poskitt, 1991: 19). This has now indeed been accepted, and a recent paper from the Ministry of Education (Irving, 1992b: 3) states that "Following careful consideration it has been agreed that national monitoring will be undertaken by light multiple matrix sampling at four-yearly intervals".

Tensions between the two forms of assessment have also been reduced to some extent since

1991. How has this occurred? In part it has been the result of an increasing emphasis by the Ministry of Education and the New Zealand Qualifications Authority on criterion referenced assessment, i.e. standards-based assessment (Irving, 1992b; Woods, 1992; Peddie, in press). At a recent National Workshop on Professional Development in Educational Assessment organised by the Ministry of Education it was emphasised that "a basic philosophic shift is occurring in New Zealand from norm referenced assessment procedures towards standards-based and criterion referenced assessment" (1992: 15).

Tensions reduce if the balance between the two forms of assessment is made less equal. The less heavy-handed approach to national monitoring and the increasing emphasis on standards-based assessment, i.e. competency based and achievement-based assessment, have helped to achieve this and have generally been approved of by educationists (Ministry of Education, 1992). However the uneasy relationship between the two forms of assessment continues to exist. How the relationship will work out in practice over the next few years remains to be seen.

Before considering some of the practical implications, it is useful to examine some of the main attributes of criterion referenced and norm referenced assessment.

Criterion referenced assessment aims at measuring performance in terms of clearly defined domains of learning. It describes what a student knows, or can do, and does not make any overt comparisons with other students. Standards-based assessment with its two forms of competency based assessment and achievement based assessment are examples of criterion referenced assessment. Historically, criterion referenced testing (CRT) has had links with mastery learning, behavioural objectives and specific learning outcomes, minimum competency testing, and equality of educational opportunity.

Norm referenced testing (NRT) measures the performance of an individual in relation to some

known group. The aim is to establish a spread of scores and to discriminate amongst items and students. Interpretations are often made against state or national norms. Historically, norm referenced assessment has had links with standardised tests of ability, aptitude and intelligence. To some extent they have been used as devices of educational and social control acting as gatekeepers in terms of progress through the system, e.g. ability grouping and national exams.

What are some of the educational and socio-political implications of these two forms of assessment? A careful examination reveals some significant contrasts.

Criterion referenced assessment is associated with:

- ~ individual student progress,
- ~ learning as process as well as product,
- ~ intrinsic motivation,
- ~ self evaluation
- ~ diagnostic and formative assessment
- ~ local school control and initiatives, and
- ~ teacher professionalism and empowerment.

It has leanings towards the political agenda of professional educators.

Norm referenced assessment is associated with:

- ~ accountability,
- ~ monitoring the system,
- ~ summative evaluation,
- ~ national exams, and
- ~ centralised control which involves national and even international monitoring of standards.

It has leanings towards the political agenda of the New Right.

As stated earlier, the current policy makers see of these two forms of assessment as complementary, without apparent tensions or conflict. Thus we have statements such as assessment "should begin with the learning needs of the students, while acknowledging the public right to clear information on overall educational standards" (Irving, 1992b: 1). Again,

assessing individual student progress and monitoring national educational standards "are two forms of assessment carried out for different purposes and therefore requiring different assessment instruments and procedures" (Irving, 1992b: 2). But how happily do these two approaches coexist?

It is clear that with criterion-referenced assessment the major emphasis is on the learning needs of the student and that "assessment should focus on the individual student rather than on comparisons between students" (Irving, 1992b: 1). This form of classroom based assessment should be, as Wylie and Smith (1992) emphasise, practical, not too time consuming to carry out and record, and should be directly relevant to the existing curriculum and classroom work of individual children.

Criterion referenced approaches require that teachers will have control over the curriculum and assessment procedures. Accordingly, the Ministry of Education gives an assurance that:

"The emphasis for this type of diagnostic, formative assessment of students' performance is placed where it belongs - with the teacher as part of the normal teaching and learning process" (Irving, 1992b: 1).

Self evaluation is also an important aspect of criterion referenced assessment. However, Willis (1992) warns that "unless students are able to play a significant role in discussion about the purpose of what they are learning, a sense of ownership seems unlikely to develop (Hargreaves, 1989, in Willis, 1992: 124). Valid self evaluation implies that the learner should be thoroughly informed about the learning objectives from the outset and hopefully have had some input into them. It is difficult to see how this could be possible with learning objectives predefined and assigned to fixed levels within the National Curriculum.

Further implications arise for motivation and learning. While contrasts should not be drawn too sharply some differences between the two forms of assessment can be made. In terms

of motivation, differences between intrinsic and extrinsic are aligned respectively with criterion referenced and norm referenced assessment. Some claims are made (Ministerial Working Party on Assessment for Better Learning, 1990; Codd, McAlpine and Hansen, 1990) that achievement-based assessment assists the intrinsic motivation of students. Learning goals are also seen as more attainable for a wider range of students and learning and assessment are seen to be fairer for different cultural groups. It is also claimed that the quality of learning in terms of problem solving skills and long term retention is enhanced when students are intrinsically motivated (Nicholls et al, 1985, in Willis, 1992).

However, if learning and teaching associated with criterion referenced assessment overemphasises lower level knowledge and memory at the expense of higher level thinking and originality there are threats to intrinsic motivation. Furthermore, one of the dangers of intrinsic motivation is that the value of what is learned may not be questioned. Willis (1992: 216) comments that "Without an analysis of what one should be motivated towards, intrinsic motivation becomes an end in its own right".

Pre-specified Learning Outcomes: Boon or Bane?

The proposed National Curriculum of New Zealand (1991) provides a structure of objectives in eight clearly defined levels for each of the broad aims, or strands of a subject. Initially in mathematics, science, English and technology, the levels will include the development of assessment exemplars which will provide guidance to classroom teachers of the type of assessment that can be used to make appropriate judgements about the students' progress. Clearly defined objectives are intended to help teachers provide clear profiles of individual student achievement and assist in reporting to parents.

One of the purported strengths of the National Curriculum is its emphasis on defining knowledge, understanding and skills as clear learning outcomes. Another purported strength is the close link between learning outcomes and assessment procedures. The Draft

Documents on Mathematics and Science in the National Curriculum (1992) are current

examples of how these two principles can be translated into practice. Yet despite general approval of these two principles there have been criticisms both in New Zealand and overseas on some aspects of learning and assessment associated with pre-specified learning outcomes and standards-based assessment. It is informative to examine some of these criticisms.

As Woods (1992; 1) comments, "The so-called 'new right' economic thinking has shifted the emphasis on to the outcomes [product] rather than input or process". Some critics (Stables, 1992) see the shift in emphasis to outcomes or product to be at the expense of process in learning. In speaking of the National Curriculum in the U.K., Stables comments that teachers of English in one of the trials felt that too much emphasis was being placed on outcomes or product. He claims that teachers concerned with diagnostic and formative approaches should be emphasising processes of learning and thinking.

Critics have also seen dangers in a curriculum that views knowledge as a series of discrete bits of information that everyone must learn, and assessment as a method of measuring pre-specified learning outcomes. Fragmentation of the curriculum into a host of specific learning outcomes is likely to lead to a narrowing of content and assessment requirements. There are also dangers of developing a whole generation of learners who are simply good at answering questions on prescribed segments of knowledge. In contrast, cognitive researchers have shown that "learning is reflective, constructive and self-regulated" (Herman, 1992: 75). Good assessment therefore should be built around such components. Recent studies of learning and motivation also highlight the importance of affective and metacognitive skills.

One of the problems with assessment is that it is easier to measure the lower level knowledge and recall than higher level thinking. Yet it is the higher levels of thinking which hook the

student into more sustained learning. There are dangers that lower level learning outcomes will dominate assessment and that such learning will become superficial, unconnected, fragmented and associated only with the shadows of knowledge rather than with knowledge itself (Schwab, 1989).

Willis (in press) points out that "assessment should focus on the ability of students to identify relationships rather than encourage them to break information into discrete facts". Broadfoot

et al (1991) also comment on the dangers of narrowing the curriculum as the result of an overemphasis on specific learning outcomes. Romberg and Zarinnia (1989: 175) claim that since current criterion referenced tests "do not measure disciplined inquiry, interpretation of knowledge, production of discourse on novel problems, critical thinking, cultural knowledge and the like, teachers ignore these attainments and later lose the capacity to produce or even imagine them" (in Broadfoot et al, 1991). Willis (in press) also comments that high quality learning and assessment "should assume that real learning is active and relevant to real life issues. It is important to develop assessment that reflects this perspective", continues Willis, "if we are to use assessment to improve learning rather than just measure it" - that is, a technician approach.

With reference to measuring specific learning outcomes through standards-based assessment Elley (1992a) acknowledges its suitability for skills and basic knowledge. He claims, however, that it does not do a good job of assessing some of the most important objectives in the school curriculum associated with higher order thinking, critical, creative and integrative thinking. "Are we to abandon such praiseworthy objectives to achieve tighter assessment?" Elley asks, and continues "If so, standards-based assessment would be a tragedy" (1992a: 1).

Elley claims that standards-based assessment falls down because: the standards themselves

are often general and vague; most skills require a more meaningful context in which to be placed; and the assumption of discrete sequential units or a ladder of standards is simply not true. "Until we can solve these problems", says Elley, "standards-based assessment is only an impracticable ideal which will offer too much power to examiners and too little justice to students" (1992a: 3). Elley concludes that all of these factors make standards based tests too blunt in high stakes situations.

Further controversy surrounds the division of learning outcomes into eight levels. Elley, (1992b) argues that the eight levels are somewhat arbitrary. Woods (1992: 1) also realises that "There are problems with assigning outcomes to levels. The levels are somewhat arbitrary, however, since no universal learning theory has yet been developed (or is likely to be developed) which can give us guidance in appropriately categorising learning outcomes".

In the U.K. teachers in the English trials of the National Curriculum felt that the "Levels as a whole often failed to give a coherent overall picture of a child's progress" (Stables, 1992: 114). English teachers also felt that much of the domain of English could not be put in neat packages and assigned clearly to one level. They claimed that there are elements of English teaching which are too important to trade off against a purely reductionist skills-based approach to assessment. They also felt that there could be constraints in allowing uneven development and faster progress through the levels.

Elley (1992b), in commenting on the Draft Science in the National Curriculum (1992), makes similar comments. He states that: students progress at different rates in different aspects of each learning area; if students are expected to progress by only one level about every two years they, and their parents, will see signs of progress much too rarely; and, as many objectives are important at many levels, they can scarcely be said to form a progression. Overall, Elley says that eight levels are arbitrary and that the structure

is a "straightjacket designed to create an illusion of tidiness in a political scheme, rather than a reflection of students' developmental stages" (1992b: 1).

New Testing Technologies and Demands of the New Curriculum

It appears that political pressures may be distorting the development of educationally sound assessment practices. Shavelson, Baxter and Pine (1992) also doubt whether new testing technologies can be developed within the short time frames demanded by politicians of the New Right. However he reports some interesting developments at the University of California and the Californian Institute of Technology in alternative methods of assessing science - assessments consistent with emerging constructivist assumptions about learning and teaching. These assessments are focusing on meaningful performance tasks and examine thinking processes and multiple solutions rather than single right answers. Such assessment procedures establish close dynamic links between teaching, learning and assessment. As Shavelson et al (1992: 22) comment, "A good assessment makes a good teaching activity and a good teaching activity makes a good assessment". Shavelson et al are currently investigating the reliability and validity of benchmarks connected with teacher observation and assessment with real life investigations and simulations in science, and while they say they

are optimistic, they warn that the measurement reality is more sobering than the political rhetoric. They warn that:

... before rushing out declaring advantages of new testing technologies associated with performance based assessments and profiling the nation should follow a more reasoned pace. We need more time to coordinate research in classrooms with these new approaches. If new accountability systems are developed hastily ... they may just as likely drive education in unwanted directions as in the desired direction. More debate and tryouts, informed by basic and evaluative research are needed. The political rhetoric is simply too far ahead of the technical reality. Our research has not addressed the social

impact
of new accountability systems. Untested assumptions are made about
salutary
effects of performance assessments on students and teachers.
(Shavelson et al, 1992: 26).

Shavelson et al conclude that bully-pulpit use of high stakes testing is
likely to be counter
productive to the quality of education as a whole. They state that we
should rather be putting
more resources into teacher inservice training in assessment and
encouraging a more
enlightened community of teachers. Finally they claim we should also be
paying particular
attention to our assessment of students from diverse social, economic and
cultural
backgrounds.

In New Zealand the Ministry of Education is also emphasising that
educational assessment
be fair to all students. "Assessment must recognise and value the
differences in race, culture,
language, gender, background and experience ... Educators must also ensure
that the
appropriate emphasis is given to the knowledge, skills and attitudes of
Maori and Pacific
Island students and of both boys and girls" (Irving, 1992b: 1). However,
this may be easier
said than done and Herman (1992) claims that one of the most demanding
challenges in high
stakes national testing of standards is mediating the demands for equity
(race, gender and
language background) on the one hand, with demands for standardisation and
comparability
on the other.

If assessment reform is to meet the teaching and learning needs of the new
curriculum, more
attention should be given to current research. Policy makers should be
more aware of the

important contemporary developments in assessment. Herman (1992)
highlights the
importance of real life problems or their simulation, and cooperative
learning contexts rather
than competitive ones. Cooperative groups facilitate learning by modelling
effective thinking
strategies, scaffolding complicated performances and valuing elements of

critical thought.

Assessments are also being carried out with team or group performances.

Recent research on assessment (Herman, 1992) emphasises such factors as the value or worthiness of the content of the assessments. (including the meaningfulness and relevance for the individual student) and also the consequences, the fairness, the transfer and generalisability, the cognitive complexity, and the cost effectiveness of the new performance-based assessments. Wiggins (1992) emphasises similar principles, arguing that assessment tasks should be "rich, realistic and enticing", supporting multiple approaches to problem solving. Assessment tasks should also be, wherever possible, authentic and meaningful - reflecting knowledge that is worth mastering. Performance standards should reflect genuine benchmarks using a variety of assessment approaches. Assessment should get away from over-measuring simple learning outcomes and move toward higher order learning and thinking. Such higher order thinking, according to Wiggins (1992: 29) is: non algorithmic, i.e. the path of the problem solving should not be predetermined; complex, with the total path not visible from any single vantage point; capable of yielding multiple solutions, each with costs and benefits; and self regulative. Wiggins (1992: 29) emphasises assessment which searches for larger, "more interrelated but complex chunks of content" compared with discrete bits of knowledge.

Assessment and the National Monitoring of Standards

While the demands of the new curriculum are pushing assessment policies in one direction, the demands of the New Right for national accountability continue to push them in quite a different direction. The Minister of Education (Smith, 1992: 14) insists, "We simply must ensure that we know how well our education system is serving our students and the stakeholders - the taxpayers and parents - who fund it".

At one level there is to be monitoring of student progress at three key

transition points -
school entry, standard 4 to form 1, and form 2 to form 3. Using item banking from nationally standardised items, schools will be able to assess the relative performance of their students against national standards. The aim is to "help determine the needs of groups of students as they enter new schools and assist in the better targeting of resources" (Irving, 1992b: 3). Such testing however is open to abuse both as a criterion for extra resources and in terms of invalid comparisons being made between the performances of different schools. As Willis (1992: 206) notes, "the use of assessment as a market signal is both problematic and controversial". Nevertheless, the Minister of Education (Smith, 1992: 10) states that "To ensure that schools who face greater challenges are provided with extra resources the National Government will develop tracking procedures at entry into primary school, intermediate school and secondary school".

What are the criteria going to be to gain these extra resources? High scores or low scores? If low scores gain extra funding what is there to stop some teachers from encouraging students to fake low to gain extra resources for the school? Later in the same speech (Smith, 1992: 12) hints it will be extreme scorers that gain the resources. "The Government can direct resources to schools facing specific challenge in education for either disadvantaged or particularly gifted students."

Another view suggests that assessment for the national monitoring of standards focuses on the education system. In this view, national monitoring

... uses standardised procedures, takes place at specified points in the system and at a set time during the year, and does not need to involve all students. Usually it is undertaken periodically such as every four years, in order to build up a picture of student achievement nationally over time and thus provide clear information on how well national standards are being maintained and where improvement might be needed. (Irving, 1992b: 2)

It is interesting to note that when it comes to monitoring educational standards the issue is not only national but also international. The Minister of Education claims that "We require information to establish just how well our students are performing compared to their

counterparts overseas" (Smith, 1992: 4). He further states (Smith, 1992: 14) that information from the national monitoring of standards "will also be complemented by our participation in the international survey of educational achievement carried out every five years by the International Association of Educational Achievement, or I.E.A."

At this level, national monitoring will be undertaken by light multiple matrix sampling at four-yearly intervals. This will involve about 5% of New Zealand students (approximately 2,500 students). With light multiple matrix sampling individual students answer a proportion only of all items but a comprehensive picture is built up of all major learning objectives over the total national sample. Such an approach enables systematic monitoring of national standards without placing unnecessary burdens on children, teacher or the system as a whole. It enables a much wider range of assessment tasks to be included in the monitoring, and it is cost effective (Irving, 1992). National monitoring will require a three year lead in time and will be put into place at standard 2 (8 year olds) and form 2 (12 year olds).

While this form of national monitoring operates at the system level, however, it does not meet the political demand for market accountability at the school level. In this area, there are recent ominous signs that the New Zealand Treasury is continuing to press for measures that will enable schools to be compared in terms of their "effectiveness". Hence, on 4 September 1992, the government introduced the Public Finance Amendment Bill No.3 which has specific provisions relating to the accountability of schools. It appears that through this legislation schools could be required to furnish assessment data as part of "effectiveness reviews" to be

conducted by the Education Review Office (Austin, 1992). The New Right may be readying itself to recapture the high ground of assessment policy.

Conclusion: Signs for Optimism

The reform of student assessment in New Zealand is occurring against a background of unprecedented change, both within the education system and throughout the wider society.

Much of this change has been influenced by the prevailing political ideology of the New Right. The emphasis, therefore, is on managerial accountability within a free market environment. The main values to be promoted are those of possessive individualism,

competition, and consumer choice. Professionalism is viewed as a manifestation of self-interest which, without rigorous accountability, leads to "provider capture".

Given such a socio-political context, it is predictable that there would be strong demands for a comprehensive system of mandatory student assessment that would enable schools to be compared and national standards to be monitored. It is perhaps too soon to conclude that such policies will not be adopted, but it is clear that educators to date have been relatively successful in promoting curriculum and assessment reforms that reflect professional concerns. Although, as we have shown, many tensions remain, there are also signs of optimism. In concluding this discussion of recent developments, the following positive signs are worth noting:

1. Feedback on the Draft Mathematics in the National Curriculum has been generally positive (Ministry of Education, 1992). Eighty-five percent of respondents gave an "approve of all" or "approve of most" rating. High approval also was given to the philosophic approach to maths and the emphasis on problem solving and learning maths in context. This ties in with new testing technologies overseas (Shavelson et al, 1992; Wiggins, 1992; Herman, 1992).

2. The Ministry of Education is developing a Handbook on Assessment for the Achievement Initiative which should be available in draft form before the end of 1992.

The handbook will provide comment on:

- ~assessment principles;
 - ~schoolwide assessment policy;
 - ~the uses of diagnostic and formative assessment - including peer assessment, self-assessment, checklists and cumulative portfolios;
 - ~making judgements about student performance against learning objectives specified in the new curriculum statements;
 - ~the use and development of assessment examples/activities across the curriculum;
 - ~assessment and fairness;
 - ~recording student achievement;
-
- ~ways of reporting student achievement;
 - ~role of standardised tests;
 - ~teacher development in assessment; and
 - ~key assessment terms.

This should be a timely and helpful guide for teachers and other professionals.

3. The New Zealand Qualifications Authority has also published a number of useful explanatory documents and booklets on the National Qualifications Framework, e.g.

The Framework and Schools (1992) outlines key elements in the framework including:

- ~units of learning;
- ~standards-based assessment;
- ~quality systems;
- ~recognition of prior learning; and
- ~records of learning.

A comprehensive publication on standards-based assessment (Peddie, 1992) is about

to be published. This is a timely publication and describes basic measurement concepts

and terms, different forms of standards-based assessment viz a viz competency-based

and achievement-based assessment. Assessment samples are given and critical issues

discussed. It is important that N.Z.Q.A. and the Ministry of Education complement

each other with such publications and not confuse readers.

4. The Ministry of Education is also designing subject-specific assessment guides to accompany the new curriculum statements for the Achievement Initiative subjects.

Examples of sound assessment strategies based on proven classroom activities and practice will be presented as case studies.

5. Teacher development programmes in the Achievement Initiative subjects of Mathematics and Science are being planned for 1993.

6. The reporting of student performance as student profiling is also a positive sign. A rich array of information related to knowledge, skills and attitudes can be reported in this way. In line with this philosophy a trial School Leaver Documentation Project (Nightingale, 1990) has recently been completed.

7. A recent National Workshop on Professional Development in Educational Assessment (1992: 24-5) organised by the Educational Assessment Secretariat of the Ministry of Education made the following encouraging recommendations:

- ~ that the primary purpose of educational assessment be recognised as the support and enhancement of student learning;
- ~ that the Ministry of Education accord a high priority to teacher development in the use of educational assessment to support and optimise student learning;
- ~ that teacher development in assessment be school-based;
- ~ that teacher development should promote informed reflective practice and a problem solving approach;
- ~ that the crucial role of principals as the professional leaders of schools requires their early involvement with regard to developments in assessment;
- ~ that a focus of professional development should be to increase awareness of ways that assessment can impede learning and to provide ways to encourage action to avoid these harmful effects;
- ~ that the community be involved in information sharing and decisions with regard to teacher development in assessment;
- ~ that the Ministry of Education publish guidelines to ensure that assessment data

are appropriately used, that access to them is safeguarded and that the rights and interests of individuals and groups be protected;
~that the Ministry of Education establish an integrated policy framework which ensures future changes will: identify issues for target groups; address bias/discriminatory practice; reflect the inclusive curriculum;
~that professional development on assessment take account of the needs of Maori, Pacific Island and other ethnic groups, girls and women, and students from low socio-economic backgrounds.

These positive developments echo many themes from around the world and suggest common concerns in the areas of curriculum and assessment reform. It is to be hoped that they provide an agenda for the future.

REFERENCES

- Austin, M. (1992) A Principal's Role - Today and Tomorrow. Speech to the Bay of Plenty Principals' Association Conference (8 October 1992).
- Broadfoot, P. (1992) Changing Priorities in Assessment: An International Reform Agenda. Paper to an international conference organised by the New Zealand Qualifications Authority (January 1992).
- Broadfoot, P., Abbott, D., Croll, P., Osborn, M., Pollard, A. and Towler, L. (1991) Implementing National Assessment: Issues for Primary Teachers. Cambridge Journal of Education, 21 (2), 153-168.
- Capper, P. (1992) Curriculum 1991, in H. Manson (ed.) New Zealand Annual Review of Education, Department of Education, Victoria University of Wellington, pp. 15-27.
- Codd, J.A. (1990) 'Educational Policy and the Crisis of the New Zealand State'. In S. Middleton, J. Codd and A. Jones (eds.) New Zealand Education Policy Today: Critical Perspectives, Wellington: Allen and Unwin, pp.191-205.

Codd, J., Gordon, L. and Harker, R. (1990) Education and the Role of the State: Devolution and Control Post-Picot. In H. Lauder and C. Wylie (eds.) Towards Successful Schooling, London: The Falmer Press, pp.15-32.

Codd, J., Harker, R. and Nash, R. (1990) Education, Politics and the Economic Crisis. In J. Codd, R. Harker and R. Nash (eds.) Political Issues in New Zealand Education, Second Edition, Palmerston North: Dunmore Press, pp.7-21.

Codd, J., McAlpine, D. and Hansen, S. (1990) Achievement, Assessment and Consensus: An Evaluation of the Sixth Form Certificate Practical Art Moderation Trial. Department of Education, Massey University.

Codd, J., McAlpine, D. and Poskitt, J. (1991) The Achievement Initiative: Myths and Mysteries of Assessment. Paper to Annual Conference of New Zealand Association for Research in Education, Dunedin.

Crocombe, G.T., Enright, M.J. and Porter, M.E. (1991) Upgrading New Zealand's Competitive Advantage. Auckland: Oxford University Press.

Draft Science in the National Curriculum (1992) Ministry of Education, Wellington, New Zealand.

Draft Mathematics in the National Curriculum (1992) Ministry of Education, Wellington, New Zealand.

Education Reform Act (1988) London: HMSO.

Elley, W. (1992a) Is Standards-Based Assessment Feasible?: Some Blunt Words About Our Blunt Instruments. Paper to a National Workshop on Professional Development in Educational Assessment. Ministry of Education, Wellington.

Elley, W. (1992b) Comments on Draft Science in the National Curriculum: A Response to the Ministry of Education. University of Canterbury, New Zealand. Mimeo.

Grace, G. (1990) Labour and education: the crisis and settlements of education policy. In M. Holland and J. Boston (eds.) The Fourth Labour Government, Second Edition, Auckland: Oxford University Press, pp.165-191.

Greig, D. (1990) Perspectives on National Monitoring. In C. Croft (ed.) Proceedings of the Conference on National Monitoring in Education. Wellington: NZCER, 54-62.

Habermas, J. (1976) Legitimation Crisis, London: Heinemann.

Hargreaves, A. (1989) Curriculum and Assessment Reform, Milton Keynes, Open University Press.

Herman, J.L. (1992) What Research Tells Us About Good Assessment. Educational Leadership, May 1992, pp.74-78.

Irving, J. (1992a) Response to Patricia Broadfoot's paper, NZQA Conference, Wellington (January 1992).

Irving, J. (1992b) Educational Assessment and Monitoring in New Zealand, Ministry of Education, Wellington, New Zealand (August 1992).

Jonathan, R. (1987) The Case For and Against Modularisation, Scottish Education Review, 19 (2), pp.86-97.

McLaughlin, R. (1988) The Scottish Vocational Educational Council, Scottish Educational Review, 20 (1), pp.52-8.

Minister of Education (1988) Tomorrow's Schools. Wellington: Government Printer.

Minister of Education (1991) Education Policy. Investing in People - Our Greatest Asset. Wellington: Government Print.

Ministerial Working Party on Assessment for Better Learning (1990) Tomorrows Standards. Government Printer, Wellington, New Zealand.

Ministry of Education (1991a) The Achievement Initiative. New Zealand Education Gazette 70 (7): 1-2.

Ministry of Education (1991b) The National Curriculum of New Zealand. Wellington: Learning Media, Ministry of Education.

Ministry of Education (1991c) The Achievement Initiative: Progress. New Zealand Education Gazette 70 (16): 1-2.

Ministry of Education (1992) Mathematics in the National Curriculum. New Zealand Education Gazette 71 (15): 1-2.

National Workshop on Professional Development in Educational Assessment (1992) Educational Assessment Secretariat, Ministry of Education, Wellington, New Zealand (17-20 May 1992).

New Zealand Qualifications Authority (1991) Designing the Framework. A Discussion Document About Restructuring National Qualifications. Wellington: NZQA.

New Zealand Qualifications Authority (1992) The Framework and Schools: A Qualifications Framework for New Zealand. Wellington: NZQA.

Nightingale, D. (1990) School Leaver Documentation Development Project: Evaluation report on the trend in nine New Zealand secondary schools, 1989-1990. Ministry of Education, Wellington, New Zealand.

Peddie, R. (1992) Beyond the Norm? An Introduction to Standards-Based Assessment. New Zealand Qualifications Authority. Wellington, New Zealand.

Peters, M. (1992) A Critique of the Porter Report. Delta, 46, pp.3-12.

Public Finance Act (1989) Wellington: Government Printer.

Romberg, T.A. and Zarinnia, E.A. (1989) The Influence of Mandated Testing on Mathematics Instruction: Grade 8 Teachers' Perceptions, University of Wisconsin,

Madison, U.S.A.

Selwood, S. (1991) Progress or Paradox? NZQA: The Genesis of a Radical Reconstruction of Qualifications Policy in New Zealand. Unpublished Master of Educational Administration Thesis, Massey University.

Schwab, J. (1989) Testing and the Curriculum. Journal of Curriculum Studies 21 (1), pp.1-10.

Shavelson, R.J., Baxter, G.P. and Pine, J. (1992) Performance Assessments: Political Rhetoric and Measurement Reality. Educational Researcher, pp.22-27, May 1992.

Smith, L. (1992) Assessment and Evaluation in the Primary School. Paper presented at the University of Canterbury, Christchurch, May 1992.

Stables, A. (1992) Speaking and Listening at Key Stage 3: Some Problems of Teacher Assessment. Educational Research, 34 (2), pp.107-115.

Treasury, New Zealand (1990) Briefing Papers to the Incoming Government. Wellington: Government Print.

Wiggins, G. (1992) Creating Tests Worth Taking. Educational Leadership, pp.26-33, May 1992.

Willis, D. (1992) Educational Assessment and Accountability: A New Zealand Case Study. Journal of Educational Policy, 7 (2), pp.205-221.

Willis, D. (in press) Learning and Assessment: Exposing the Inconsistencies of Theory and Practice.

Woods, M. (1992) New Assessment Skills For the National Curriculum and the Qualifications Framework. New Zealand Qualifications Authority, Wellington, New Zealand.

Wylie, Cathy and Smith, Lesley (1992) Assessment and Reporting Practices
in the First
Three Years of School: Interim Report to the Ministry of Education, January
1992,
New Zealand Council for Educational Research, Wellington, New Zealand.