

Curriculum Reform: Forwards or Backwards?

WARWICK B. ELLEY

In April 1993, the Minister of Education launched the *New Zealand Curriculum Framework*, which sets out the essential skills and achievement objectives to be developed in every school. During the year, also, the detailed national statements of aims and standards for Mathematics and Science were published, along with suggested instructional activities and assessment tasks. In addition the draft statement for English was released in October, for comment, and work was well advanced on the Technology and Social Science curricula. The Government's hope is that these various documents would "specify clear objectives" and identify "for boards of trustees, teachers, students, parents and the wider community a progression of desirable standards of learning – against which students' progress can be assessed" (Ministry of Education, 1993a). This is the hallmark of the reformed curricula – that the progression in each subject be clearly specified. The standards are spelled out in the form of eight successive levels in each learning area and students' progress is to be assessed in relation to these levels.

These statements are the first curriculum documents to be developed under the Ministry's new contract system, and it remains to be seen whether they will acquire sufficient levels of teacher ownership and attract sufficient resourcing and teacher development to be successfully implemented, as intended. This reviewer has serious doubts on each of these scores. The tight

time frames allowed for development have meant that very few teachers have been involved, few trials have been undertaken in schools, and while there has been newspaper debate over the role of knowledge vs. skills in the Science curriculum, there has been too little discussion on the general model adopted for the new curricula.

There is certainly much to commend in the curriculum statements themselves. The teacher groups who identified the key objectives and suggested appropriate teaching activities worked hard within the politically-driven deadlines to produce the desired "clearly-specified standards", showing the progression of skills and knowledge to be attained by students. They had numerous precedents to follow, in the shape of the 10-level Attainment Targets model in the United Kingdom, and various local handbooks and curriculum reviews prepared for recent governments. However, the dramatic protests and U-turns which have plagued the UK national curriculum should have sent signals to the National Government that the way ahead in New Zealand will not be straightforward. In spite of a number of differences between the UK and New Zealand curriculum models, there are several key problems common to each.

The Curriculum Structure

In calling for an eight-level progressive structure in each curriculum area, the Government is presuming a similarity across subjects, which has no basis in research or teacher experience. A neat uniform structure across all subjects may please the tidy mind, but teachers know, even if politicians do not, that the link between content and aims is not the same in mathematics or technology as it is in English or Social Science. Furthermore, there is no rationale for dividing any of the curriculum areas into eight (or ten) evenly spaced levels. There are no natural breaks in development at the selected cut-off points and no pilot studies have been carried out to determine whether they represent challenges of similar magnitude, within or across subjects. As Dr Peter Dines (1991), an architect of the British system conceded,

when visiting New Zealand in 1991, there is no research on children's learning to support the levels structure. In both the UK and New Zealand, the levels have been set where they are, chiefly on the basis of the subjective opinions of the teachers who served on the committees. Some critics have claimed that eight is an arbitrary number and it is certainly too few, as the average student working up the continuum from five to 16 years will linger on a given level for close to two years; slow students for three or more. Most teachers can surely provide more frequent incentives and a greater sense of progression than this.

More serious is the question of whether the sequencing of knowledge and skills constitutes a clear progression at all. In many vocational skills, and in some areas of mathematics, it is true that children do tend to follow a pre-ordained sequence (Young-Loveridge, 1993). One should have mastered basic number combination before fractions; one can normally count to 100 before 1000. These regularities were recognised in the 1928 syllabus, and others before it. But in many parts of the curriculum, students' knowledge growth is individual and idiosyncratic. Their knowledge consists of an infinity of particulars, not of logically organised packages – mastered in all-or-nothing fashion. It expands in irregular spurts, sidetracks, inconsistencies and misconceptions. So students may master level 5 objectives before level 4 in some specific areas, while lagging at level 2 in others of the same strand. The ideal linear progression which may be found in physical education, or typing, is alien to most knowledge-based subjects. History, for instance, is multi-dimensional in character. Student progress is found in increasing levels of knowledge, processes and skills, in different history topics. So, teachers who attempt to bring their students' knowledge growth into line with the prescribed sequence in history – or in science, or English literature, will often feel pressure to bypass the teachable moment in the interests of achieving the desired outcomes. There is little evidence to suggest that instruction is more effective if it is outcome-driven rather

than interest-driven. Evidence from the UK and the United States suggests otherwise (Torrance, 1993; Shepherd, 1991).

Some of these difficulties can be illustrated in the English Draft Curriculum. This document was released for comment in October, after an all-too-brief development phase, restricted as it was, by teachers' industrial action through late 1992 and early 1993.

Firstly, it should be recognised that the contractors were unable to spell out eight distinguishable levels of achievement in more than five of the 33 achievement objectives which they identified in the English curriculum. Personal reading has three levels, handwriting four, expressive writing five and interpersonal listening seven. The arbitrariness and the difficulty of identifying specific forms of advancement in students' growth, is obvious here. Key thresholds in student development are reached at different ages, depending on their learning styles, and the nature of the subject.

Furthermore, the awkwardness of trying to impose a level structure on a subject like English is illustrated in the wording of the levels. In *transactional listening* for instance the first four levels are defined as follows:

1. Listen to, respond to and recall text.
2. Listen to texts and recall and *respond to main ideas*.
3. Listen to texts and recall and respond to the main ideas *in sequence*.
4. Listen to texts, recall and respond to the main ideas and *identify the purposes of the text and the presenter*.

To determine what is meant by progression, in transactional listening, we can study the emphasised words, which represent the difference between each level and the one preceding it. Leaving aside, for the moment, the problem of varying difficulties of "text", we can see that the curriculum developers assume first that there is a central aim running through all levels. Who could

claim otherwise? Language development is like this. Mathematics by contrast, can set different aims at different levels. Secondly, the developers assume that responding to main ideas is harder than recalling text, that responding to the main ideas *in sequence* is harder than responding in any other sequence, and that identifying the purposes of the text and presenter are harder tasks again. Of course, there is little foundation for these assumptions in research. Asked for main ideas about a message, most students do present them in the sequence they heard them. And if students are interested in, and familiar with the topic addressed by a text, they are likely to achieve all these objectives readily; if they have no relevant well-formed schema for the topic, they may well have difficulty with all of them. As research in reading and listening have already shown, the variance due to text familiarity far outweighs variance due to skills (e.g., Farr and Carey, 1986). Again, if the chosen text is so structured that the main ideas are listed in sequence, or highlighted by the author, then the level 3 standard will be readily attained; if the author spells out his or her aims, but not the main ideas, then level 4 will be more easily attained than level 3.

The point is that difficulty, in reading or listening, is more a function of the characteristics of the text – its structure, its complexity, its vocabulary load, and the match between the interests of the student and the content of the text, than it is a function of the particular skills defined in these levels. A skill-based level structure may operate successfully in athletics or woodwork, but not in receptive language modes. The inherent progression in language is not captured by these level statements.

Indeed, if we were to examine the real difference between a typical level 3 child (about eight to nine years) and a level 7 child (about 15 years) in reading or listening it would be found not in such skills as the ability to identify main ideas, or to detect purposes, or to think critically, so much as the ability to understand progressively more difficult text. Young children often can identify main ideas in simple texts; older students cannot do so in prose which is abstract, dense, technical or

characterised by a wide range of unusual vocabulary. Typical 15-year-olds know the meanings of 10,000 words or more; typical eight-year olds know fewer than half that number. We could, then, establish a progression by describing text difficulty, perhaps in terms of vocabulary load. However, if the curriculum developers had focused on such an obvious dimension in their attempts to describe the progressions in reading, we would have been sending quite the wrong signals to teachers, that vocabulary knowledge is a desired outcome; teach to it. Whereupon teachers would abandon their time-honoured productive strategies of story reading aloud, shared reading and sustained silent reading – which do result in such an outcome (Elley, 1989; Nagy et al., 1985) and systematically teach new vocabulary instead – a painfully slow way of achieving similar ends.

The difficulty of pinning down clear standards is obvious throughout the English statement. In the *poetic writing* strand, for instance, pupils in each of the first four levels are expected “to write descriptively on a variety of topics”. The progression, then, is to be found in the fact that level 1 children are “experimenting with language and form”; in level 2 they will use “a number of forms”, and “make language choices”; in level 3, the wording is identical, except that the language choices are “imaginative”, while in level 4 the language is “selected” (not chosen) “for imaginative effect”. These subtle semantic distinctions are contrived and arbitrary, and surely unhelpful in teaching or assessing children’s progress from levels 1 to 4, (i.e., about eight years of schooling). Some level 4 children will still be “experimenting with language and form”, while some level 1 children will sometimes “select language for imaginative effect”. And why should imagination appear only at level 3? Are these statements any improvement over those of the 1928 syllabus?

Some of the level statements do provide guidance to teachers about the English objectives to be aimed at – but they are largely objectives which are common to all levels. The growth children exhibit as they become more mature is their ability to use language in a greater range of contexts, with a greater range of

text types. The curriculum developers recognise this, but the eight-level structure cannot accommodate it.

It is unclear, at this stage, whether the new curriculum model will result in any improvement in overall standards. In the teaching of English, New Zealand has an enviable reputation overseas, for the quality of its instructional programs. Successive IEA studies (Thorndike, 1973; Elley, 1992) have confirmed that it produces very high achievement levels. Nevertheless, a new curriculum model is being introduced, which is ostensibly outcome-driven, rather than interest-based – without trial or adequate rationale, and following a British model which is recognised locally as unworkable. Strange policy-making indeed!

Assessing Students Against the Levels

The new curriculum, structured as it is in an eight-level framework, is also intended to provide for easier assessment of pupils. In the UK, this function was given high priority, within their ten-level structure, with the result that many teachers in 1990-1992 have spent literally hundreds of hours testing their children on the recommended standard assessment tasks (SATs). Once the teachers saw the way in which school authorities and newspaper reporters started using the test results to compare schools, and to produce “league tables”, the tests quickly took on a “high stakes” image, with all the evils usually implied by such status. Almost every edition of the *Times Educational Supplement* through 1992 and early 1993 reported abuses, anomalies and protests. Teachers coached children for the tests, administered them repeatedly, encouraged certain children to absent themselves, and used the lack of standardisation procedures to their advantage. Far from serving any helpful diagnostic purpose, the SATs took on an accountability role which effectively minimised their value for improving children’s learning, and distorted the nature of the curriculum (Torrance, 1993; Lofty, 1993).

By 1991 the Scottish education authorities had backed off the new assessment system, after teacher protests. Then the teachers in England and Wales boycotted the tests in 1992-3, with wholehearted support from the community, culminating in an official investigation chaired by Sir Ron Dearing. In his Interim Report, released in August 1993, Dearing concluded that the required testing was far too unwieldy, and that some alternative to the level structure was called for. Why, then, is New Zealand embarking on a similar journey in our curriculum reform?

The Ministry of Education claims that our curriculum model is really different from that of the United Kingdom (Perris, 1993) that ours is more flexible, and promotes assessment which is intended for diagnosis rather than accountability. This may be the Ministry’s intention at this stage, but the National Curriculum in Britain was launched with similar pious hopes in 1989. In fact, the similarities between the two national curriculum models are more apparent than the differences. Both the UK and the New Zealand models have a similar multi-level structure, with about six major strands in each curriculum area. The target objectives and standards to guide teaching are similar in their level of specificity and there is every indication that teachers are to be requested to assess children in relation to these levels in both countries. Already, the New Zealand Minister’s visionary document, *Education for the 21st Century*, (Ministry of Education, 1993b) released in September 1993, has set targets relating to the percentages of students at given class levels who should have attained particular levels of achievement by 1998, and again by 2001. Thus, 65% of Form 4 students should have reached level 5 or better in English, Mathematics and Science by 2001. To determine whether these targets are met, schools will clearly be asked to assess their children formally in relation to the level statements. Just as in the UK, we can predict that what started as an idealistic, teacher-friendly model will rapidly deteriorate into an accountability-driven unfriendly system (Gipps, 1992).

Viewed another way, good diagnostic assessment is best done informally, according to need, in unstandardised fashion. By contrast, accountability assessment must be uniform and tightly controlled. Running records of children's reading, for instance, are very helpful diagnostic tools for teachers, but they are most helpful when they are least standardised, and they are so time-consuming that they are best confined to selected children. The same assessments cannot serve both purposes well. Likewise, writing conferences are very helpful for formative assessment of children's writing, but they too should be tailored to children's local need at the time, not standardised for acceding to authoritative decree. Diagnosis and accountability are never comfortable bedfellows (Brown, 1992).

Assuming then, that the new curriculum structure of eight levels is to be used not just for guiding teachers, but also to provide summative information about the levels attained by all children (i.e., for reporting to parents, boards of trustees, ERO or Ministry) – the question must be asked as to whether the level statements are sufficiently clearly-specified to enable teachers to do this in fast and accurate fashion. This form of assessment is new for most teachers. In the jargon of professional assessors it is criterion-referenced, or standards-based, rather than norm-referenced. Teachers will be asked to assess their children's progress in relation to where, in the series of abstract level statements, they fit, rather than how they are achieving in relation to others in the school. This would be fine if we were dealing with clearly-specifiable skills. We could accurately assess pupil progress in swimming, or typing, or learning their basic number combinations. But how can teachers reliably assess important academic skills which are not easy to specify in steps? How can they accurately assess whether pupils "show an awareness of style" or "can recall a wide range of facts"? These are the kinds of level statements proposed in English and geography. How can they agree on whether their pupils "can detect main ideas in sequence", when the ability to do this depends on the topic of the passage to be read, the structure of the passage, the

vocabulary load, the time limits, the format of the question, the identity of the teacher, the extent of coaching, and many other context factors? The standard assessment tests (SATs) in England provided many examples of such unreliable assessment. In one school, teachers could report that either 90% of children had attained level 2 standard, or that 38% had, depending on the book they chose to test from.

The point is that testing is not an exact science. Teachers can only sample pupil behaviour – they cannot capture it all. Pupil behaviour differs from task to task, and from day to day. It is affected by the particular tasks the teacher selects for assessing, by the wording of the questions, by the attitude of the tester, by the time allowed for completion, by the recency of their study of such tasks, and many other factors. Such distorting influences can be accounted for, and minimised, in formal standardised tests and external examinations, especially if supplemented by moderated internal assessments. If the tasks were the same for all, if national norms were available to show how others of similar age typically perform on this particular task, under these particular conditions, then teachers could make reasonably accurate judgements of the percentage of their students reaching particular levels of the New Zealand curriculum. But this is a far cry from the helpful diagnostic assessments proposed for assessment under the new model. If teachers are to retain the promised flexibility in planning, in teaching, and in assessment, then the high-stakes accountability functions of assessment must be clearly separated from the helpful diagnostic ones.

Of course, one may well question the desire of the Government to record students' achievement levels at all in this way. There is already a well-researched proposal to monitor student achievement levels across the school system as a whole, every four years, in all curriculum areas (Crooks & Flockton, 1993). The plan calls for assessment of a light sample of 3% of standard two and form two children. This proposal, which will have no undesirable "backwash" effects on schools, has been accepted by the Minister, and pilot testing is well advanced. This

should tell the Government what it needs to know about changes in educational standards, over time. If, however, the purpose of the more formal universal assessment of pupils is to identify schools which require special funding, there are already other more defensible approaches which can be taken (Dialogue Consultants, 1991).

It would be a pity if what benefits there are in the curriculum reforms were to be undermined by needless assessment requirements which cut across desirable aims. This reviewer would recommend that the levels notion be abandoned, that curriculum guidelines focus on aims to be sought and topics to be addressed, rather than standards to be attained. At the very least, the notion of calculating the number of pupils who attain given levels in a school should be avoided. In the present climate, where schools clamour for more enrolments and seek to build up a positive image, the potential for allowing vague level standards to drive instruction will only be counter productive.

References

- Brown, S. "Raising standards: factors influencing the effectiveness of innovations" in *Critical Reflections on Curriculum Policy*, Edinburgh: Scottish Council for Educational Research, 1992.
- Crooks, T. & Flockton, L. *The Design and Implementation of National Monitoring of Educational Outcomes in New Zealand Primary Schools*, Report No.1. Dunedin: University of Otago, 1993.
- Dearing, Sir Ronald. *Interim Report on National Curriculum*, London: Ministry of Education, UK, 1993.
- Dialogue Consultants. *Development of Socio-Economic Indicators of Disadvantage in Schools*, Browns Bay, Auckland, 1991.
- Dines, P. "National Curriculum and Assessment in UK", Address to National Conference of Educational Administrators, Christchurch, January, 1991.
- Elley, W. B. *How in the World Do Students Read?* The Hague: IEA, 1992.
- Elley, W. B. "Vocabulary acquisition from listening to stories", in *Reading Research Quarterly*, 24, 2, 1989, pp. 174-187.

- Farr, R. & Carey, R. F. *Reading: What can be Measured?* Newark Delaware: International Reading Association, 1986.
- Gilmore, A. M. *An Evaluation of the Sixth Form Certificate English Moderation Trial*, Christchurch: University of Canterbury, 1991.
- Gipps, C. "Equal opportunities and the SATs for seven year olds", *Curriculum Journal*, 3, 2, 1992, pp. 171-183.
- Lofty, J. S. "Can Britain's National Curriculum Show America the Way?" *Educational Leadership*, 50, 2, 1993, pp. 52-55.
- Ministry of Education. *The New Zealand Curriculum Framework*, Wellington: Learning Media, Ministry of Education, 1993a.
- Ministry of Education. *Education for the 21st Century: a Discussion Document*, Wellington: Learning Media, Ministry of Education, 1993b.
- Nagy, W. E., Herman, P. A. & Anderson, R. C. "Learning Words from Context", *Reading Research Quarterly*, 20, 2, 1985, pp. 233-253.
- Perris, L. "Ministry response on New Zealand curriculum model", *Dominion Sunday Times*, 29 September, 1993.
- Shephard, L. "Will national tests improve student learning?" *Phi Delta Kappan*, 71, 1991, pp. 232-238.
- Thorndike, R. L. *Reading Comprehension Education in Fifteen Countries*, Stockholm: Almqvist & Wiksell, 1973.
- Torrance, H. "Combining measurement-driven instruction with authentic assessment: some observations of National Assessment in England and Wales", *Educational Evaluation & Policy Analysis*, 15, 1, 1993, pp. 81-89.
- Young-Loveridge, J. *The Effects of Early Mathematics Intervention: Report to Ministry of Education*, Hamilton: University of Waikato, 1993.

The author

Warwick Elley is Professor of Education at Canterbury University. He has been the leader of an international project on literacy for the IEA. He was formerly at the University of the South Pacific and has become a leading commentator on assessment in New Zealand.