THE ASSESSMENT RESOURCE BANKS AND OTHER APPROACHES TO SCHOOL-BASED ASSESSMENT IN NEW ZEALAND

Paper presented at the Association of Commonwealth Examination and Accreditation Bodies Conference, Malta, March 18–22, 2002

Cedric Croft

Research Director, Learning, Curriculum and Assessment

New Zealand Council for Educational Research

Abstract

Assessment Resource Banks (ARBs) for English, mathematics and science are now being further expanded and developed at the New Zealand Council for Educational Research. The ARBs in mathematics and science became available to all New Zealand schools via the Internet in early 1997, with English added in 1998. They include assessment material linked closely to curriculum statements and suitable for students levels 2–5 (8-year-olds – 15-year-olds) with material at the upper end of the range used within secondary schools for school-based purposes.

The levels-based structure of the curriculum and the need for teachers to assess to levels presents a challenge for all school-based assessment, particularly for assessing students' writing. One recent innovation in the English ARBs has been to develop level-based scoring guides, and link these to exemplars of writing that illustrate major hallmarks of writing at particular levels.

In addition to the ARBs, there are three other national initiatives plus a range of achievement tests which are utilised by schools for a variety of school-based and formative purposes. The paper will overview the nature of the ARBs and other standardized tests available within New Zealand for school-based and formative uses.

1. The Assessment Resource Banks in English, Mathematics, and Science*

The Assessment Resource Banks (ARBs) have been developed at the New Zealand Council for Educational Research (NZCER) under contract to the Ministry of Education to match New Zealand curriculum statements in English, mathematics, and science, for levels 2 to 5 (eight-year-olds to fifteen-year-olds). Mathematics and science resources first became available via the Internet in small numbers in March and May 1997. English was added in September 1998. As at March 2002, there were in excess of 3000 resources available, made up of 1250 in mathematics, 1300 in science, and 550 in English.

The ARBs are not developed as alternatives to a school's own assessment procedures, but as a source of complementary material. They are provided to help teachers assess progress within strands or achievement objectives, or judge the relative performance of their students against the "typical" performance of national samples of students at given year levels. The ARBs are not designed to be used by students or parents, except that a new range of animated resources require students to respond to a series of on-screen displays. These are a forerunner to the ARBs having some future interactive capacity.

A strength of the ARBs is that the published assessment materials are prepared co-operatively by assessment specialists and teachers, trialled with a national sample of students, then selected by teachers to represent their own teaching objectives. This development process ensures acceptable levels of curriculum validity and reliability. The process of selection should ensure classroom validity, as teachers select assessment resources that best match their curriculum objectives and teaching programmes.

Access which is at **http://www.nzcer.org.nz** begins at the NZCER homepage. To use the banks it is necessary to hold a username and password, available from an on-line registration form. Staff from registered schools, registered teachers, and staff from teacher support services, are among those eligible for a password. Access is available by arrangement for assessment staff internationally.

All resources are presented in a format that may be printed and photocopied. It is also possible to cut and paste resources electronically, and save ARB files to a word processing package so that the assessment tasks may be adapted as needed.

_____

* Adapted from Croft (2002).

2. Background and Development

The ARB project began in February 1993, when NZCER was contracted by the Ministry of Education to investigate the feasibility of establishing a bank of materials to assist in the assessment of students at the 'transition points' of Year 6/7 and Year 8/9.

The second stage for the ARBs, beginning in December 1994, was to implement the ARBs on a trial basis in 27 schools. The third stage, from April 1996, was to improve a search engine developed during stage two, extend the number of mathematics and science resources in the banks, and make these available to all schools via the Internet. The fourth stage, from July 1997 to June 1999, was to increase the number and range of mathematics and science assessment resources, continue to improve the flexibility of the search strategy, and add English assessment resources. For the fifth stage, from July 1999 to June 2001 the ARBs continued to grow in size and encompass a more diverse range of assessment material with strong growth in English a feature. During this period the ARBs were confirmed as a school-based resource where prior to this time they were seen as having a role in national testing.

The development and growth of the ARBs during the period 1993–2000 have been covered in Croft, Reid and Livingstone (1995), Croft, Gilbert, Boyd, Burgon, Dunn, Burgess and Reid (1996), Croft (1996), Brown and Strafford (1997), and Croft (1999).

In the current phase from July 2001 to June 2003, a redesigned website is evident, more material with a strong formative basis is available, there are marked increases in web traffic and school-based uses, and criteria have been developed to help determine the optimum size and general shape of the ARBs in future (Croft, 2001).

3. Development and Expansion of the ARBs 1997–2002

Since opening with 125 accessible resources, the ARBs have grown to 3100 resources as at March 2002, including 297 resources added from the Third International Mathematics and Science Study (TIMSS). The development and expansion of the ARBs is outlined in the following tables.

For Table 1, the numbers of resources developed and published at NZCER are combined for 12-month periods, 1 July to 30 June.

Table 1

Number of NZCER Resources Published to ARBs in English, Mathematics and Science by 12 months Period 1997–2001

12 month periods	English	Mathematics	Science	Total
1 July 1997 – 30 June 1998	–	147	268	415
1 July 1998 – 30 June 1999	140	233	210	583
1 July 1999 – 30 June 2000	163	236	249	648
June 2000 – July 2001	183	216	208	607

Taking the three full years from 1 July 1998 to 30 June 2001, the annual average number of NZCER resources published to the ARBs has been 613. This number breaks down to 162 for English, 228 for mathematics and 222 for science. It is apparent from Table 1 that the numbers of resources published in mathematics and science is less for the last year reported than for the two preceding years. This resulted from a policy decision to increase the number of resources designed to engage students in extended responses, with an aim of emphasising the bank's formative uses.

These figures indicate approximate future annual average growth of NZCER resources published to the ARBs, assuming present resourcing and reasonably constant specifications for the styles of assessment resource to be developed and the curriculum levels to which they apply.

Table 2 records the number of TIMSS resources incorporated into the ARBs over a twelve-month period.

Table 2

Number of TIMSS Resources Added to the ARB in Mathematics and Science, by Six-monthly Periods, 1999

Six month period to	Mathematics	Science
June 1999	77	62
December 1999	74	66
June 2000	–	18
	151	146
	===	===

4. The Structure of the ARBs

Whether we are talking of English, mathematics or science, the general principle is that the ARB learning area, e.g., (mathematics), follows the structure of the respective National Curriculum Statement. Each learning area is organised into a series of strands, achievement objectives or functions, levels, and process skills. A focus on English will give a general picture of this structure.

Each resource in English is classified in curriculum terms, by strand, function, level and process skill. Additionally, there are keywords and resource types to provide additional dimensions for each resource's classification.

The classification fields which come directly from the curriculum are as follows:

Strand

In English, the learning strands are written language, visual language and oral language.

Functions

Each learning strand in English has a number of functions. The functions by strand are:

Written language

Reading functions	–	Personal reading
		Close reading
Writing functions	–	Expressive writing
		Poetic writing
		Transactional Writing

Visual language

Viewing function	–	Viewing
Presenting function	–	Presenting

Oral language

Listening functions	–	Interpersonal listening
		Listening to texts
Speaking functions	–	Interpersonal listening
		Using text

Curriculum Level

Each learning area has eight levels. Resources in English are for levels 2–5. These correspond to about years 4-10, or in age terms, 8 year olds to 15 year olds. As used generally in Curriculum Statements, levels are more descriptive than definitive. They have elements of mastery learning or criterion referencing implicit, but they are generally a looser conceptualization than either. Generally, there are few systematic attempts by schools to ensure that a student has 'mastered' all content or processes implicit or explicit within a level prior to moving to the next. The workable approach developed in schools is to describe a student as "working within" a particular level. There is no requirement to assess a student's performance within a level as for example, 'superior', 'average' or 'failing'. But there are requirements to report the level at which a student is achieving.

Process Skill

Most ARB resources are further classified by a predominant process skills. In English the process skills are exploring language, processing information and thinking critically. Each process skill is broadly defined within the English curriculum statement with key components of the skill further elaborated by strands and levels. For example:

WRITTEN LANGUAGE ACHIEVEMENT OBJECTIVES			
Reading and Writing Processes			
	Exploring Language	**Thinking Critically**	**Processing Information**
LEVELS 5 and 6	*In achieving the objectives of understanding and using written language, students should:* • using appropriate term-inology, describe, dis-cuss, analyse, and apply the distinctive convent-ions, structures, and language features of a range of texts and explain how they suit the topic and purpose	*In achieving the objectives of understanding and using written language, students should:* • interpret, analyse, and produce written texts, identifying and dis-cussing their literary qualities, and explore and identify attitudes and beliefs in terms of personal experience and knowledge of other texts	*In achieving the objectives of understanding and using written language, students should:* • using appropriate tech-nologies, retrieve, select, and interpret informa-tion from a variety of sources, and present accurate and coherent information for a range of purposes, analysing the processes used

Ministry of Education, 1994, p. 46.

Keywords

Keywords are another form of resource classification. Keywords are not a feature of the curriculum.

Each resource has keywords or phrases designed to further describe the content and predominant skills tapped by the resource. Wherever possible the keywords are taken directly from New Zealand curriculum statements, but because of variations in terminology for similar concepts, some alternative terms are required.

There is an on-line dictionary of keywords used to construct this type of search. As resources are added to the ARBs, dictionaries are expanded to include new keywords. This ensures that there is at least one resource in the bank for each entry in the dictionary. The keyword search is a very powerful aspect of the search engine. It is popular with users, as it allows a search of the ARBs to be undertaken by topic.

An extract from the Keyword dictionary for Visual language follows:

VISUAL LANGUAGE KEYWORDS			
Acronym	Drawings	Message	Signs
Advertisement	Effectiveness	Metaphor	Similarities
Arrows	Ellipsis	Mime	Speech bubbles
Audience	Emphasis	Mood	Static image
Background	Exclamation mark	Movement	Stereotypes
Body language	Expression	Movement lines	Storyboard
Book cover	Facial expressions	Myth	Sub-heading
Brochure	Feelings	Newspaper	Sub-title
Bullet points	Films	Pamphlet	Symbols

A search to select resources for school-based assessment may be undertaken by a single classification field or any combination of the fields discussed so far.

Resource type

All resources are classified also by the style of response expected from students. The following four types are used in mathematics and science resources:

- **Selected response (SR)**. Students *select* a response from a range of options incorporated in the resource. Two or three multiple-choice or matching items may be grouped to form one resource. Examples include multiple-choice items, matching items, true/false and other alternate-choice items.

- **Brief constructed response (BCR)**. The student *constructs* the response. Short answers, such as a word or two, a number or two, a phrase, or brief sentence are the essence of a BCR. Correct brief responses will encapsulate a single main idea.

Completing entries in tables, graphs, or diagrams constitute a BCR. Examples include short-answer questions, completion items for tables, graphs, diagrams, plans, etc., problem-solving tasks requiring brief structured responses.

- **Longer constructed response (LCR)**. These have the same general characteristics as a BCR but require a more extended response from students. The LCR resource is generally more open-ended than the BCR, and inferences may be needed to determine relationships within the task. *Producing* tables, graphs, or diagrams constitute an LCR. Other examples include short essay-type questions (structured or unstructured); a written plan for an experiment, investigation, or practical task, conceptualising and/or producing tables, graphs, diagrams.

- **Practical (PRA)**. These are based around a performance component involving responses including, but additional to, paper-and-pencil. An investigation may be undertaken, data may be analysed, conclusion drawn, or a product may be completed. Examples include simple investigations or experiments in science or mathematics; classifying tangible materials in science; undertaking measurement tasks in mathematics; constructing shapes or figures in mathematics.

In English there are six resource types: Selected response (SR), Short written response (SWR), and Longer written response (LWR) resources have similar characteristics to SR, BCR, and LCR resources in mathematics and science.

- **Oral response (OR)**. The predominant response is oral, although a minor written component may be included.

- **Student rating or assessment (SRA)**. The essence of these resources is that a rating or assessment is undertaken by students. This category makes provision for student self-assessment or peer assessment by way of rating scale, observation scale, or checklist.

- **Teacher rating or observation (TRO)**. Resources of this type are included to assist teachers' assessments of expressive skills, mostly in the written and oral strands. Multi-level marking guides come within this category, although some multi-level material are included for LWR resources as well.

Resource type also provides a classification field for searching for resources. Examples of a TRO follow.

Levels-based Assessment — Transactional Writing

	Range of Tasks Write in authentic contexts	Deep Features of Writing [organisation, sequence, sentences, vocabulary]	Surface Features of Writing [punctuation, grammar, syntax , spelling]
Level 1	• *Write instructions.** • *Recount events.*	• Writes several related sentences on the topic. • High frequency vocabulary predominates	• Beginning use of full stops, capitals. • Beginning use of conventional syntax [word order] • More than 20% spelling errors (excluding proper nouns); some conventional spelling patterns evident but mostly 'semiphonetic' attempts.+
Level 2	• *Write instructions.* • *Write explanations.* • *Recount events.* • *State facts and opinions.*	• Includes several ideas some with supporting detail. • Some sequencing is evident. • Beginning to vary sentence beginnings and length. • Beginning to extend sentences with conjunctions. • Vocabulary broadening beyond high frequency.	• Mostly correct uses of full stops, capitals, commas for listing, and question marks. • Beginning to use quotation marks. • Conventional syntax generally evident. • Between 10% and 20% spelling errors (excluding proper nouns) and moderated by breadth of vocabulary with majority recognisable. Increasing conventional spelling patterns evident, with mostly 'phonetic attempts'.+
Level 3	• *Write instructions.* • *Write explanations.* • *Write factual accounts.* • *Express personal viewpoints.*	• Beginning to support main ideas with some details. • *Sequences ideas logically.* • Beginning to organise some ideas into paragraph. • Varies sentence beginnings and length. • Beginning to structure sentences in a variety of ways and may use complex sentences, consisting of more than one subordinate clause.. • Beginning to use vocabulary appropriate to task/genre.	• Mostly correct use of full stops, capitals, commas, question marks, exclamation marks and quotation marks. • Control of verb forms i.e. singular/plural agreement ,subject/verb agreement and tense. • *Conventions such as spelling appropriate to genre.* • Between 5% - 10% spelling errors (excluding proper nouns) and modified by breadth of vocabulary. Shows clear phonetic mapping; conventional patterns increasing in number and variety.+
Level 4	• *Write instruction.* • *Write explanations.* • *Write factual accounts.* • *Express and explain a point of view.*	• Consistently includes details to support main ideas. • Organise ideas into coherent paragraphs. • *Organises and links ideas logically.* • *Make language choices appropriate to the audience.* • Varies sentence beginnings and sentence length to suit purpose. • Structure sentences in a variety of ways with increasing use of complex sentences consisting of more than one subordinate clause. • Vocabulary generally appropriate to task/genre.	• Accurate use of full stops and capitals, commas, question marks, exclamation marks, speech marks, apostrophes, parentheses, dashes, colons, semi-colons, ellipses. • *Using appropriate spelling.* • 3% - 5% errors (excluding proper nouns) and moderated by breadth of vocabulary.+

Level 5	• *Write coherent*
	    - logical instructions
	    - explanations
	    - factual accounts.
	• *Express and argue a point of view.*
		• *Links main and supporting ideas.*
		• Strong sequential structures evident within and between paragraphs.
		• *Structures material in appropriate styles.*
		• Evidence of vocabulary carefully chosen for task.
			• *Using conventions of writing accurately and confidently (punctuation, grammar).*
			• Wide use of subordinated structures in sentences with errors rare and variety in length.
			• Conventional spelling predominates.

Level 6	• *Write clear*
	    - coherent instructions
	    - explanations
	    - factual reports.
	• *Express and justify a point of view.*
		• *Uses appropriate styles for different audiences.*
		• *Justifies point of view persuasively.*
		• *Structures material confidently.*
			• *Using conventions of writing accurately and with discrimination.*

* Statements from Curriculum, pp. 34-35, 92-100 in *italics*.
+Based on data from NZCER National Survey of Primary Writing, Croft & Mapa. 1998.
NZCER [July 2000] Assessment Resource Banks. Revised, March 2001.

Levels-based Assessment — Poetic Writing

	Range of Tasks Write on a variety of topics	Deep Features of Writing [language, organisation, sentences, vocabulary]	Surface Features of Writing [punctuation, grammar, syntax , spelling]
Level 1	• *Beginning to shape ideas.**	• Writes several related sentences on topic. • High frequency vocabulary predominates	• Beginning use of full stops, capitals. • Beginning use of conventional syntax [word order] • More than 20% spelling errors (excluding proper nouns); some conventional spelling patterns evident but mostly 'semiphonetic' attempts.+
Level 2	• *Shaping ideas in a number of genres such as* • *letters* • *poems* • *narrative*	• *Making choices in language and form.* • Story line with sequential structure evident, some descriptive detail. • Beginning to vary sentence beginnings and sentence length. • Beginning to extend sentences with conjunctions. • Vocabulary broadening beyond high frequency.	• Mostly correct uses of full stops, capitals, commas for listing, and question marks. • Beginning to use quotation marks. • Conventional syntax generally evident. • Between 10% and 20% spelling errors (excluding proper nouns) and moderated by breadth of vocabulary with majority recognisable. Increasing conventional spelling patterns evident, with mostly 'phonetic attempts'.+
Level 3	• *Shaping, editing, and reworking texts in a range of genres* • *letters* • *poems* • *narrative*	• Beginning to incorporate some descriptive detail of setting and character to support story line. • Beginning to organise ideas into paragraphs. • *Sentence structure appropriate to genre.* • Varies sentence beginnings and length. • Beginning to structure sentences in a variety of ways and may use complex sentences, consisting of more than one subordinate clause. • *Vocabulary appropriate to genre.*	• Mostly correct use of full stops, capitals, commas, question marks, exclamation marks and quotation marks. • Control of verb forms i.e. singular/plural agreement ,subject/verb agreement and tense. • *Conventions such as spelling appropriate to genre.* • Between 5% - 10% spelling errors (excluding proper nouns) and modified by breadth of vocabulary. Shows clear phonetic mapping; conventional patterns increasing in number and variety.+
Level 4	• *Shaping, editing, and reworking texts in a range of genres* • *letters* • *poems* • *narrative*	• *Expressing ideas and experiences. imaginatively;* occasional use of figurative language and or innovative use of vocabulary. • Organises ideas into coherent paragraphs. • Narratives include descriptive detail of character and setting. • *Using appropriate sentence structure.* Varies sentence beginnings and sentence length to suit purpose. • Structure sentences in a variety of ways with increasing use of complex sentences, consisting of more than one subordinate clause. • *Using appropriate vocabulary.*	• Accurate use of full stops and capitals, commas, question marks, exclamation marks, speech marks, apostrophes, parentheses, dashes, colons, semi-colons, ellipses. • *Using appropriate spelling.* • 3% - 5% errors (excluding proper nouns) and moderated by breadth of vocabulary.+
Level 5	• *Shaping, editing, and reworking texts in an extended range of genres* • *letters* • *poems* • *narrative*	• *Selecting appropriate language features.* Uses figurative language and innovative use of vocabulary with control and intent. • Strong sequential structure evident within and between paragraphs. • Maintains appropriate vocabulary throughout.	• *Using conventions of writing accurately and confidently (punctuation, grammar).* • Wide use of subordinated structures in sentences with errors rare and variety in length. • Conventional spelling predominates.
Level 6	• *Shaping, editing, and reworking texts to express ideas imaginatively in a range of genres*	• *Choosing appropriate language features.*	• *Using conventions of writing accurately and with discrimination.*

A series of trials and refinements were undertaken until the levels-based scoring guides were published in December 2000. A general conclusion from the trials was that the scoring guides were instrumental in helping achieve 80 percent agreement by level, for groups of teachers assessing samples of writing.

5. The Makeup of a Resource

The ARBs are now the major nationally developed assessment resources in New Zealand linked to national curriculum statements. They provide nationally derived performance data for each item, question, or task within a resource.

Each resource is presented in two parts: the questions or tasks for the student, and the scoring guide for the teacher. This latter part includes performance data, and in many instances, examples of diagnostic information. The diagnostic information is developed from an intensive analysis of errors found in responses from the national samples of students. At the time of writing, mathematics, science and a handful of English resources include diagnostic information.

Diagnostic information includes common errors, examples of common misunderstandings, and, where possible, likely misconceptions and incorrect calculations (Neill, 1997). The term "diagnostic" is incorporated in the keyword dictionary, so a search may be made for all resources in the banks which have this information. If a more specific outcome is needed, the search may be directed to particular strand(s), level(s), achievement objective(s), resource type(s) and so on.

The scoring guide for each response includes correct answers to questions and appropriate responses to tasks, a scale of marks, and information about the difficulty level of each question or task within a resource. These data are obtained from trials on groups of about 200 students from seven or eight representative schools. The five descriptive statements and corresponding difficulty levels are:

Difficulty estimate	Percentage correct
Very easy	80% and above
Easy	60% to 79%
Moderate	40% to 59%
Difficult	20% to 39%
Very difficult	19% and below

The scoring guide also shows the year level(s) for which the data were collected and the date of the trial.

11

6. National and School-based Uses of the ARBs

From the outset, national and school-based uses of the ARBs have been dual elements of the project. In the initial stages national uses predominated. Indeed, the aim of using the ARBs on a national basis at the 'transition points Year 6/7 and Year 8/9, provided the main impetus for initial funding.

In discussing future national uses of the ARBs, Croft, Gilbert, Boyd, Burgon, Dunn, Burgess and Reid (1996), p. 81, had noted:

> "The possibility that the banks could provide information on science and mathematics achievement at Years 7 and 9, points which mark the transition primary/intermediate/secondary, and the completion of primary/intermediate schooling, appealed as contributing to the government's Achievement Initiative and its policy on accountability within the school system. Hence the ARBs were to become the major vehicle for implementing transition-point assessment. The general intention of using ARBs to provide some form of national information on Year 7 and Year 9 cohorts had considerable government support."

As the ARBs began to expand and offered more extensive coverage of curriculum statements, alternative possible national uses of the banks had been put forward. These were outlined in Croft (1999), pp. 57–58.

A 1998 Green Paper, *Assessment for Success in Primary Schools* included a proposal for externally-referenced tests, also referred to as national tests. In its Response to the Green Paper, NZCER had recommended against national testing, concluding, p. 12:

> We are unaware of research findings or reputable literature which supports the validity of a single test for large-scale testing. We are aware of reputable literature which notes the invalidity of a single test for large-scale testing, e.g., Brown, McCallum, Taggart, and Gipps (1997), Jones (1997), Davey and Neill (1991).

and, p. 16:

> The case against the GP version of national testing is multi-faceted and educationally compelling. We conclude that national testing as outlined in the GP will not fulfil its purposes. Our view is that it should not proceed in the form it is outlined.
>
> We recommend that our suggestion for Intact tests within the ARBs be developed to the point where they may be field tested.

The Intact tests referred to above were to be assembled from the existing resources within the ARBs. Intact tests were seen as secondary to the main purposes of the ARBs, which were to make available an expanding collection of **individual assessment resources** that schools could match to their own teaching programmes.

Although the school-based uses of the ARBs were dominated by their potential national uses during the period 1996–1999, there was no doubt in the minds of the developers that the school-based aspects would pay better dividends long term. It had been noted in Croft, Gilbert, Boyd, Burgon, Dunn, Burgess and Reid (1996), pp. 85–86:

> We have no doubt that the proposed school-based uses of ARBs represent an innovative approach to improving the quality of school assessment practices. When combined with electronic delivery, a broadening of the item-bank concept, and development of a classification system representing the New Zealand Curriculum Framework, school-based uses of ARBs constitute a unique development. We find no references in assessment literature to any other system which combines these features in a similar manner.

As the ARBs have been consolidated as a school-based resource in policy terms, and by virtue of the style of resources now being developed, their focus has turned more towards supporting teachers' use of formative assessment strategies. At present 505 resources in mathematics and 227 in science are reporting diagnostic information in the form of common errors made by students, or common misconceptions that may be inferred from their responses, Neill (2001). It has been noted by Mendelovits et al. (2000), that the former summative purposes of the ARBs have probably influenced the nature of some material published prior to 2000.

A recent survey at NZCER which is yet to be published, has found that the predominant uses of the ARBs as recorded from a random sample of teachers in a representative sample of schools, are for teaching and learning purposes, monitoring students' progress and providing information on achievement to parents and caregivers. In comparison to Gilmore and Hattie (2000), the emphasis towards formative uses of the ARBs seems to have increased markedly.

Other Approaches to School-based Assessment

With the main purpose of the ARBs clarified towards school-based uses, they are now established as the premier source of nationally developed assessment material for use by schools within their own assessment programmes.

There are at present three other national tools available (or becoming available this year) for schools to use. Their role within present policy may be summarised as follows:

- The **NEMP** is primarily for system-wide monitoring or broad accountability purposes at Year 4 and Year 8. Once the tasks have been released they become available to schools. These tasks do not restrict their measurement to achievement objectives within curriculum levels, as they are more broadly conceived. Consideration is being given to releasing NEMP tasks via the ARBs, in much the same way as TIMSS items have been released.

- The **AsTTle** will provide literacy and numeracy tests for Years 5, 6 and 7 in English and Maori. In concept they are little different from the 'intact tests' proposed in 1998 for the ARBs. AsTTle will provide tests with items that 'map the curriculum'. Schools are unable to determine the composition of the test at the item level. AsTTle provides some choice, but at the whole test level only. They are to become available in April 2002 on CD ROM, with provision for schools to analyse their results against national norms. Their present structure will not allow national reporting of results.

- **Exemplars** are to provide a range of students' responses to a variety of tasks at levels 1–5 of national curriculum statements.

  > "An exemplar is an authentic example of student work annotated to illustrate **learning, achievement,** and **quality** in relation to the levels described in the relevant national curriculum statement. Each exemplar highlights significant features of that work and important aspects of students' learning.
  >
  > **WHAT IS THE PURPOSE OF AN EXEMPLAR?**
  > An exemplar:
  > - signals important features of student work to watch for, collect information about, and act on to support growth in learning;
  > - provides students, teachers, and parents with a basis for discussing important qualities, aspects, or indicators of learning;
  > - provides reference points that will support teachers' professional judgments about the quality of their students' work."
  >
  > (Ministry of Education, 2002).

In addition to the four current initiatives outlined to this point, New Zealand schools have a range of standardised achievement tests available with the NZCER *Progressive Achievement Tests* being the most prominent and frequently used.

The *Progressive Achievement Tests* include:

Listening Comprehension (Revised)
(Reid and Johnston)

Mathematics (Revised)

(Reid and Hughes)

Reading Comprehension and Vocabulary (Revised)

(Reid and Elley)

PAT: Study Skills (Reid, Croft, Jackson), has recently been withdrawn and replaced by the new:

Essential Skills Assessments: Information Skills (Croft, Dunn and Brown, 2001). This consists of six modules with fourteen tasks as follows:

–   Finding Information in a Library
–   Finding Information in Books
–   Finding Information in Reference Sources
–   Finding Information in Graphs and Tables
–   Finding Information in Prose Text
–   Evaluating Information in Text

They are designed for formative assessment but may be used for comparative purposes, as each student's score can be covered to a stanine. They cover two years at each level, so that the tests may be used to identify strengths and weaknesses and to monitor progress over time. Each test contains a range of item and question types and may be used at any time between March and November. They are group administered.

The Essential Skills Assessments were planned as the beginning of a new achievement series for NZCER, as the overall structure of the *Progressive Achievement Tests* no longer articulates closely with current curriculum statements. The *Essential Skills Assessments: Work and Study Skills* is the next in this series to be published.

All of these tests typically cover the age range of 7 years to 15 years, report results in stanines and/or percentile ranks by class and age, and in the case of *PAT: Reading* and *PAT: Listening,* level scores as well.

In addition to the achievement tests mentioned schools have available to them:

Burt Word Reading Test — New Zealand Revision

(Gilmore, Croft, & Reid)

Age: 6 – 13

Proof-Reading Tests of Spelling (PRETOS)

(Croft, Gilmore, Reid, & Jackson)

Age: 8 – 14

STAR Supplementary Test of Achievement in Reading

(Elley)

Years 4–6 and 7–9

A full description of all the tests noted here may be viewed at the NZCER website as noted at p. 2.

Conclusions

In terms of public policy and school practice, nationally-developed and standardized assessments have been a strong recent focus within New Zealand.

Standardized tests as typified by the *Progressive Achievement Tests* have been available and used widely in New Zealand schools for the past 25 years or so (Croft and Reid, 1991). Even before the publication of the first *Progressive Achievement Test* in 1969, standardized tests in reading, spelling and mathematics were commonplace.

The four national development projects as discussed earlier have come about for a variety of reasons:

- NEMP was instituted in 1995 in response to debates and concerns expressed in some quarters about standards and outcomes in primary schools. The sampling approach adopted by NEMP and the wide range of curriculum areas included is seen as vastly superior to national cohort testing.
- The ARBs became available to schools in 1997. Originally they were to have been for national testing of two cohorts at 'transition points' in their schooling, but they have now been consolidated as a resource to be utilised by schools, for their own assessment purposes. The ARBs now represent New Zealand's major resource for school-based assessment.
- AsTTle are the current government's response to the previous government's policy of introducing mandatory cohort testing at three year levels in the primary school. AsTTle are to be used voluntarily. They will not permit national cohort data to be obtained on a common set of items. In essence they are a hybrid of the ARBs and a standardized test, in that schools are given some choice at the whole test level, and there are data to be provided on students' performances in relation to a variety of school variables. Schools are to be given the capacity to interrogate data via a CD-ROM.
- The most recent exemplar project is to provide for schools 'typical' responses from students to a variety of tasks, with the intention of illustrating standards within the curriculum and helping schools assess their students' progress within these standards.

What is being reported here is in addition to the present restructuring and overhaul of the secondary school qualifications. At present, assessment and qualifications are a strong focus within the New Zealand system.

References

Brown, G., & Strafford, E. (1997, November). *The feasibility of expanding the assessment resource banks to incorporate English.* Unpublished project report. Wellington: New Zealand Council for Educational Research.

Croft, C. (1996). Assessment resource banks for school-based and national assessment. In Ian Livingstone (Ed.), *New Zealand Annual Review of Education, 6,* 127–139.

Croft, C. (1999). *School wide assessment: Using the assessment resource banks.* Wellington: New Zealand Council for Educational Research.

Croft, C. (2001). *Assessment resource banks in English, mathematics and science: Criteria for determining the optimum size and nature of the ARBs: Report to the Ministry of Education.* Wellington: New Zealand Council for Educational Research.

Croft, C. (2002, in press). *The assessment resource banks: From national testing to a school-based resource.*

Croft, C., Gilbert, A., Boyd, S., Burgon, J., Dunn, K., Burgess, L., & Reid, N. (1996, March). *Assessment resource banks in mathematics and science B implementation trial.* Wellington: New Zealand Council for Educational Research.

Croft, C., & Reid, N. (1991). *How often and for what purposes are NZCER tests used in primary and secondary schools?* Wellington: New Zealand Council for Educational Research.

Croft, C., Reid, N., & Livingstone, I. (1995). Assessment resource banks in mathematics and science. *set: Research Information for Teachers, Best of set Assessment, 2,* item 14. Wellington: New Zealand Council for Educational Research.

Gilmore, A., & Hattie J. (2000). *Evaluation of the assessment resource banks in schools: Final report.* Unit for Studies in Educational Evaluation, Department of Education, University of Canterbury and School of Education, University of Auckland.

Mendelovits, J., Farkota, R., & Lindsey, J. *Evaluation of the New Zealand assessment resource banks' methodology for Ministry of Education, New Zealand.* Hawthorn, Australia: Australian Council for Educational Research.

Ministry of Education. (1994). *English in the New Zealand curriculum.* Wellington: Ministry of Education.

Ministry of Education. (1998). *Assessment for success in primary schools* (Green paper). Wellington: Ministry of Education, Learning and Evaluation Policy.

Ministry of Education. (2002). *The national exemplars trial.* Wellington: Ministry of Education.

Neill, A. (2001). An introduction to the assessment resource banks (ARBs) and their diagnostic potential. *The New Zealand Mathematics Magazine, 38* (1), 29–38.

New Zealand Council for Educational Research. (1998). *Assessment for success in primary schools: Green paper. A response from the New Zealand Council for Educational Research.* Wellington: (author).

Reid, N., Armstrong, L., Atmore, D., Boyd, S., Croft, C., & Livingstone, I. (1994). *Assessment resource banks feasibility study: Summary report.* Wellington: New Zealand Council for Educational Research.

Author

Cedric Croft is Chief Researcher and Research Director, Learning, Curriculum and Assessment at the New Zealand Council for Educational Research. He has written widely on the Assessment Resource Banks and school assessment more generally. Recently he was a member of a Ministerial Literacy Experts Group. Recently he has studied large-scale testing programmes in the United States and United Kingdom as a Visiting Researcher at the Educational Testing Service, the National Centre for Research on Evaluation, Standards, and Student Testing, and the Institute of Education.