The role of an online repository of assessment tasks and resources: The place of the Assessment Resource Banks

6th Conference of the International Test Commission: The Impact of Testing on People and Society: Enhancing the Value of Test Use, Liverpool, July 14–16 2008.

Authors: Chris Joyce and Charles Darr, *New Zealand Council for Educational Research*

Since 1997 New Zealand schools have been able to draw on an online bank of assessment tasks for mathematics, English, and science to assist in the development of classroom and school-wide assessment. There are now more than 3000 assessment tasks available. In more recent times the focus of the banks has shifted to formative assessment. As well as the assessment tasks themselves, teachers are provided with a range of support material to enhance pedagogical content knowledge and provide a context for the assessment focus.

Introduction

In recent years formative assessment has received a lot of attention both in New Zealand and internationally. A new curriculum released last year (Ministry of Education, 2007) reaffirms improvement of learning and teaching as assessment's primary purpose.

In New Zealand there is no national testing until students reach Year 11, and schools are expected to develop their own assessment systems. There is a range of assessment tools available, some provided by the Ministry of Education (MOE), others developed independently. The Assessment Resource Banks (ARBs) are one of the assessment tools developed under contract for the Ministry. Their primary purpose is to support formative assessment in the classroom in mathematics, science, and English. This article describes the assumptions about formative assessment that underpin the resources and how these influence the design of tasks and teacher support material, the objective of which is to enhance learning. The discussion also includes how the resources support teachers in their use of assessment for formative purposes.

A brief history of the ARBs

The ARBs are collections of assessment resources developed by the New Zealand Council for Educational Research (NZCER) under contract to the MOE. They are an online assessment resource for three curriculum areas, mathematics, science, and English, catering for students from Years 3 to 10. They are freely available to all New Zealand schools and teachers. The resources are password protected, although other support material is freely available to any potential users.

The initial contract began in February 1993 with a feasibility study (Reid et al., 1994). The ARBs were originally a response to a government policy of the day known as "transition point

assessment", and the focus was on Years 7 and 9, two major transition points within the New Zealand education system. However, they were never officially used for this purpose.

The ARBs have been available to schools since March 1997, beginning with the mathematics bank. The science bank was added in May 1997 and, finally, the English bank was added in September 1998.

The rationale for the ARBs was to provide banks of items from which teachers could choose for themselves what would be assessed and when. The banks include a classification system which enables teachers to search for appropriate resources to match curriculum and teaching objectives.

The original resources were designed for summative assessment. The ARBs did incorporate a broader range of resources than traditional multiple-choice items, as they included constructed-response and practical tasks. Over time, as new resources were developed and added to the banks, more diagnostic information was added to assist teachers in making decisions about students' learning needs.

This gradual shift to a more formative purpose was formalised in 2003 when the MOE contract stipulated that the development of new resources was to assist teachers with their formative assessment practices. The focus of this paper is on the ARB resources developed from 2003 to the present time.

Developing a resource—the research process

In effect, each assessment resource is a small research project.

Each task, as it is designed, is informed by theoretical understandings about assessment, an in-depth knowledge of New Zealand curriculum content and intentions, and "informed intuitions about best practice" (Davis & Krajcik, 2005, p. 4). The task is reviewed and is likely to be modified several times before it is trialled with students.

There are two main trial methods.

Most commonly, a group of tasks is put together to form a pencil and paper "test" set. The tasks may have a common theme, but not always, although up until the present time they have always involved just one curriculum area—mathematics, science, or English. The test sets are piloted with a small group of students, and are then sent to a representative sample of New Zealand schools to trial with students at a nominated year level. The aim is to have about 200 students to trial the material.

A coding schedule is designed to capture not only how many students get the preferred answer, but also partially correct or less sophisticated responses, as well as common misconceptions. Open-ended items undergo a content analysis that has a phenomenographic slant. This may capture such things as the different strategies students use, or attitudes they bring to the task. This

is all useful information to provide to teachers to assist them to use the students' responses formatively.

The coded responses are then analysed using either a traditional item analysis that looks at point-biserials and Cronbach coefficients. This was originally done using Conquest, but latterly the statistical package R has been used. Relationships between responses are sometimes explored using cross-tabulations or concatenations. An example of this is in mathematics, where the relationship between strategies used and the answers the student gives is examined.

Some tasks are not suitable for this type of trialling and are instead trialled by a researcher working in classrooms with students (and sometimes teachers). The number of students and schools involved varies, but is fewer than the 200 students completing the pen and paper national trials. One advantage of this type of trialling is that the researcher can talk to the students as they are working on the task, or during follow-up discussions, which provides more information about their thinking processes. These tasks are analysed using largely a phenomenographic method (Marton, 1994), looking at the different approaches students use to respond to the task being trialled.

In some cases, tasks are trialled both in classrooms and in the national trials.

The data collected from student responses, supplemented by NZCER and other relevant research, inform teacher support notes. This teacher material is developed for each resource. It provides the answers or, often, a range of responses that signal a level of competence. A broad difficulty level is assigned to each question based on the number of students getting the preferred response. No difficulty level is assigned to the classroom-trialled questions (because the trial numbers are too low), or when the task is too open-ended to capture this sort of data. The rest of the teacher pages provide further information to support analysis of student tasks, feedback, and decisions about what to do next. The teacher support material is discussed in more detail in the next section.

The challenges of designing tasks for formative assessment

The purpose of formative assessment is to promote student learning. There is a myriad of definitions for formative assessment, but the following one reflects the intention of the ARB resources:

> An assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their pupils … to modify the teaching and learning activities in which they are engaged. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching work to meet learning needs. (Black, Harrison, Lee, Marshall, & Wiliam, 2002, *i*)

According to this definition, any task performed by a learner, ranging from a formal standardised test to an informal discussion between teacher and learner, or between learners, therefore has the potential to be formative *if the information gained is used to inform what happens next in the*

classroom. The assessment tasks may be planned, or the teacher may "seize the teaching moment" as it occurs during interactions with students.

This view of assessment *during learning* underpins the ARB resources. The ARB assessment tasks have been designed to be carried out as part of normal classroom activities. They help teachers to make sense of what students are saying, doing, and thinking, and to make decisions about what to do next. They can also be used to help students to reflect on their learning *as* they are learning.

Formative assessment, by its very definition, presents some challenges to resource developers. Three that are particularly pertinent to the development of the ARBs are:

- the role of the teacher in formative assessment
- the purpose of the assessment as inquiry rather than measurement (Hargreaves, 2007)
- addressing the specific needs of a learner.

This has required the resource development team to consider not only the design of the assessment tasks, but also how best to provide support for teachers, including the technical design of the website.

Supporting the teacher

The tasks by themselves cannot be formative. It is what the teacher and/or the learner do with the information gained from doing the task that makes it formative (Sadler, 1998). For example, a student may complete an English comprehension task. If their performance of this task is analysed and indicates that the student has difficulty making inferences, then an assessment judgement has been made. The teacher or the student may then use that judgement to decide to concentrate on using clues in the text to infer the characters' feelings, at which stage the task has been put to formative use. How teachers and students use the tasks, therefore, is the really important factor in whether they achieve the aim of helping students to learn (Newton, 2007).

Effective formative assessment, therefore, is dependent on three teacher-related actions involving different skills and knowledge. Cowie and Bell (1999) describe these actions as eliciting, interpreting, and acting. Eliciting involves engaging in some way with students so they can show what they know or can do. Interpreting involves analysing their responses, recognising areas of strength or difficulty, and understanding why they are having difficulty. Acting includes having strategies for helping students to learn.

Even when teachers know about using assessment formatively, in practice they often still struggle with it. There are a number of reasons why this might be so, but one seems to be that teachers need an *in-depth* pedagogical content knowledge in the context being assessed (Black et al., 2002) and most teachers have some areas where this knowledge is weaker. If teachers don't have this

strong pedagogical content knowledge they will have difficulty both in analysing student responses and knowing how to promote new learning.

An ARB task provides the opportunity to elicit information about what the student knows and can do. However, if the teacher doesn't know what to do with this information, then it won't be used effectively for formative purposes.

To provide this support, each ARB task is accompanied by educative curriculum materials (Davis & Krajcik, 2005), compiled to grow teachers' understanding of assessment, curriculum content, and how best to teach the content. These teacher pages provide information and support that will assist teachers to analyse their students' responses, and make decisions about the next learning steps; i.e., it helps them to use the data generated by the students completing the task to improve their learning. This information includes:

- a clear description of the assessment focus
- links to the New Zealand curriculum
- student responses appropriate to the task, assessment focus, and the level of students
- diagnostic information including common misconceptions (gathered both from research literature and the student trials)
- important and relevant background content knowledge
- next learning steps tied to particular responses
- other resources that support further learning.

As well as information specific to a resource, additional support material that is useful for professional development is also available. Where appropriate, there are links from the individual resources to this material. Some of the support material relates to formative assessment, while another section is to do with curriculum and is informed by classroom research carried out by the researchers involved with the resource development. In other cases the support material is a teacher-friendly summary of national and international research literature.

This additional support material may look very different in the three different curriculum areas. For example:

- In mathematics, a large component of the support material is a series of concept maps. These are a series of framework statements that have been developed in significant areas of mathematics. They provide information about the key mathematical ideas involved, link to relevant ARB resources, and make suggestions for some ideas on the teaching and assessing of that area of mathematics. These concept maps are added to as further research and trialling in the respective areas are conducted.
- In English, *Thinking about how language works* provides teachers with additional information about language that will help them to analyse student responses to ARB items. It

is also useful when analysing responses from students with English as a first or subsequent language, and is a helpful aid when writing assessment questions.

- In science, students' understandings about some important "big ideas" in science, such as interdependence, the water cycle, and fair testing, have been investigated and reported on, and relevant ARB resources identified.

Assessment as inquiry

A second challenge is to write tasks that uncover what the student is thinking. Understanding how a student gets to a particular answer is often more helpful for formative purposes than whether they know the "right" answer, because it helps a teacher, or the students themselves, to pinpoint areas of strength and weakness, and where further work is needed.

The construction of the tasks varies to model a range of assessment strategies (which teachers can use when they design their own tasks). Important aspects of many tasks include students providing evidence for their response, justifying and defending their ideas, or showing the strategies they use to arrive at an answer. Many identify opportunities for student–student or student–teacher discussion, because conversations about learning have been identified as an important element of formative assessment (Black & William, 1998). As well, collaborative work may be suggested. In some resources, specific self- and peer-assessment activities are included in the resource, while in others, suggestions for these are referred to in the teacher support material.

Meeting the needs of specific students

A third challenge is that assessment for formative purposes should be specific to the needs of the students. Although there are a large number of items, they cannot always comprise exactly what a teacher wants to assess, nor will they always be relevant to all students. A technical solution to this is to provide a Microsoft Word version of the resource that teachers can save and adapt. For example, a teacher may want to delete or add a question, or they may want slightly different versions of the same task to cater for the differing needs of their students. They are able to change a context, or even personalise a task. While these changes are likely to invalidate some of the supporting information, having a trialled model to work from has the advantage of both saving time and assisting teachers to develop good-quality tasks. The educative nature of the teacher materials also plays a role in enabling teachers to successfully adapt the model for a new context.

The Microsoft Word version was introduced prior to 2003, but has been an important component in supporting assessment for formative purposes. However, new technologies present new opportunities, and these advances are being monitored for their ability to provide this feature, the ability to adapt the task, in a more efficient way.

Where do the ARBs fit in New Zealand's assessment toolkit?

It is important to have a balance between day-to-day assessments that focus on in-depth understandings and keeping learning on track, and less frequent, more formal assessments that provide information about general progress (Wiliam, 2006). The ARBs are best suited to the first category. While New Zealand's commonly used standardised tools such as Assessment Tools for Teaching and Learning (asTTle) and Progressive Achievement Tests (PAT) also have a big role to play in formative assessment, ARB tasks have the potential to investigate a narrow focus at greater depth. This fine-grained assessment allows teachers and students to drill down to find out more about a specific aspect of a student's knowledge, skill, or understanding.

The designers of both asTTle and PAT have developed a "what next" feature as part of their support material. This directs teachers to other relevant resources, including particular ARB tasks, that explore a specific concept or skill that may have been identified in their tests as possibly needing attention or extension. Teachers can then choose to incorporate the tasks into their regular classroom teaching. This cross-referencing between different assessment tools provides a model of good assessment practice whereby it is easier for teachers to select an assessment that is suited to a particular purpose, or use a range of assessment methods to investigate a particular understanding.

Conclusion

To summarise, the ARBs are online banks of assessment resources, the main purpose of which is to provide assessment resources for formative purposes (although teachers may of course choose to use them for other purposes). They are designed for use as part of normal classroom learning and teaching activities, of which assessment and decision making as a result of collecting evidence is an integral part. They are part of a toolkit of assessment resources available to New Zealand teachers that supports schools to develop their own integrated assessment system.

References

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box.* London: nferNelson.

Black, P., & Wiliam, D. (1998). *Inside the black box.* London: nferNelson.

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education: Principles, Policy & Practice 6*(1), 101–116.

Davis, E., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3–14.

Hargreaves, E. (2007). The validity of collaborative assessment for learning. *Assessment in Education: Principles, policy & practice, 14*(2), 185–199.

Marton, F. (1994). Phenomenography,. In Torsten Husén & T. Neville Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.) (pp. 4424–4429). Oxford: Pergamon. Retrieved June 24 2008 from http://www.ped.gu.se/biorn/phgraph/civil/main/1res.appr.html

Ministry of Education. (2007). *The New Zealand curriculum.* Wellington: Learning Media.

Newton, P. (2007). The purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice, 14*(2), 149–170.

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.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice, 5*(1), 77–84.

Wiliam, D. (2006). *Does assessment hinder learning?* ETS Europe Breakfast Seminar, July 11 2006. Retrieved June 13 2008 from http://www.uk.toeic.eu/no_cache/toeic-sites/toeic-default/news-toeic/?news=136&view=detail