Using Rich Assessment Tasks in Mathematics to Engage Students and Inform Teaching

Doug Clarke
Australian Catholic University
d.clarke@patrick.acu.edu.au

Barbara Clarke
Monash University
barbara.clarke@education.monash.edu.au

Assessment: Why?

"Nobody ever got taller by being measured" (Professor Wilfred Cockroft, United Kingdom)

"You don't fatten a pig by weighing it" (American equivalent)

Over a number of years, we have had the opportunity to ask many teachers the question: "Why do we assess students?" The answers fall broadly into the following (overlapping) categories:

- To find out what my students know and can do.
- To help me to know what to teach next.
- To measure the effectiveness of my teaching.
- To provide feedback to students on their learning.
- To inform parents, other teachers, employers and interested others of academic progress.
- Because my principal/school/community expects it.

These reasons seem reasonable and appropriate. The first four answers imply some form of subsequent action, but such action is not always evident. We recall a comment from a teacher that the end-of-topic test is "a kind of official announcement that the topic is over and we will be moving on to the next one". Although the comment was somewhat "tongue-in-cheek", he explained that the curriculum was "crowded", and there wasn't really time to address any individual or group difficulties with the previous topic that emerged from the assessment. These would have to be addressed the following year, when the topic was revisited.

We have often used the two related quotes at the beginning of this section to illustrate that assessment divorced from subsequent action is unlikely to lead to improved learning. Sometimes, politicians and other educational policy makers believe that it is *the act* of assessment that will lead to improved learning, when in fact it is the *action* that follows, using the information gained from the assessment that is potentially most powerful. Having said that, there are many examples of the way in which changes in assessment at the state level have led to changes in both teaching and assessment practice across schools. For example, the emphasis in senior secondary mathematics in

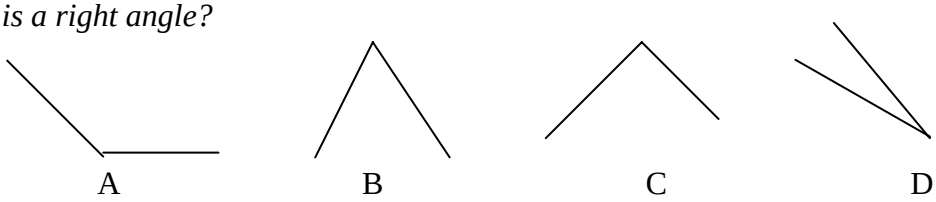
Victoria (Australia) on projects and investigations had a considerable (and, in our opinion, positive) effect on both classroom teaching and assessment methods in the lower secondary years.

Assessment: How?--The Tools of Assessment

In the 1980s, an increasing consensus emerged among classroom teachers that traditional forms of assessment were inadequate in meeting all the revised goals which teachers held for assessment. The argument was that if we value genuine understanding, problem solving and group skills, and the ability to use what has been learned in "real" situations, then we need to broaden the repertoire of assessment techniques from the classic pen and paper test, combining "informal" assessment with a greater range of formal methods of assessment. As David Clarke (1997) has argued, it is through our assessment that we communicate most clearly to students those activities and learning outcomes we value.

Both the educational and wider communities have considerable faith in the pen-and-paper test as a form of assessment, but we all know students for whom the pressure of the test leads to performance that is not representative of their knowledge and understanding. In addition, students whose background in the language of instruction is not strong may also find it difficult to demonstrate what they know and can do using this assessment format. A concrete example of such conflict between written responses and understanding is the case of a Grade 6 child in a little Catholic primary school, reported by Anne Newman (1983) in a research study. During a maths test, Jacinta was presented with a multiple-choice maths item involving four angles (see the picture below).

Which angle is a right angle?



The question asked, "which angle is a right angle", requiring her to choose from four possibilities. After ten minutes, Jacinta raised her hand and asked, "miss, when it says 'which *angel* is a right *angel*?' does that mean the *wings* go this way (pointing to one angle) or this way (pointing to another)? For this girl, the task was all about angels and wings, and is a powerful reminder that with a pen and paper test, you are not always assessing what you think you are.

Recent research has generated data which suggests that students who give correct answers to pen and paper mathematics items sometimes have little or no understanding of the concepts and relationships which the tests were designed to measure. The work of Ken Clements and Nerida

Ellerton with students in three Australian states (see Clements, 1995) contrasted the information that teachers and researchers can gain from written tests and interviews. They interviewed 80 Australian Year 8 students (around 14 –years old), on mathematics tasks for which they had produced written responses earlier. Clements and Ellerton produced data which indicated that around one-quarter of students' responses to multiple-choice and short answer maths items could be classified as either (a) correct answers given by students who did not have a sound understanding of the relevant knowledge, skills, concepts and principles; or (b) incorrect answers given by students who had partial or full understanding.

The late 80s saw an increased emphasis on the use of anecdotal records, checklists, portfolios, student self-assessment and so on. These assessment alternatives continue today, with teachers refining them in light of experience. In the 90s, issues of accountability at classroom, school and system level brought a greater emphasis on nominating desired outcomes and ways of collecting, documenting and reporting student growth over time.

Irrespective of how these changes are viewed, much has been learned about assessment in the process. The challenge remains therefore to make assessment meaningful but at the same time manageable.

Assessment: What?

Not everything that counts can be counted. Not everything that can be counted counts.

(quotation claimed to be posted on Albert Einstein's wall)

Whether true or not, this quote reminds us that just because we can easily assess memory of procedures and facts, this does not necessarily mean that these should be the major focus of our assessment attention. Unfortunately, for the busy teacher, it remains true that the easier a given form of assessment is to use, the less useful the information it is likely to provide. We do face the challenge of making mathematics manageable as well as meaningful.

The topic test issue that was raised earlier also brings into focus whether we are teaching for short, medium or long-term understanding. Most secondary teachers will recall the horror that greets an announcement that "owing to the sports' day/excursion/school assembly or other reason, today's test will be delayed until tomorrow". A common response is "you mean I have to learn it all over again tonight?" In these students' minds, there is no expectation that any understanding of the assessed content will be retained for more than 24 hours, let alone well into the future. The "teach it/test it/teach it/test it" process may actually be creating the illusion of learning.

Making Assessment an Open Process

Student: I wondered why I got a "C" for the last assignment. I thought it was pretty good".

Teacher: You got a "C" because you did what I asked you to do. If you want a better grade, you need to do much more".

This true story from a Year 11 student illustrates the importance of clear communication of assessment expectations. If the teacher expects a student to know the features of quality work and be able to produce it, it is important to share examples of such quality, and come to some agreement of what constitutes quality, preferably with a discussion of the assessment criteria. If these criteria can be developed together with the students, all the better. One strategy is by sharing, anonymously, examples of student work of both high and somewhat lower quality, and discussing these in relation to the criteria.

But there is more that a teacher can do to allow students to be genuine participants in the assessment process. Conversations about expectations should occur early in the teaching and learning process, in order that students understand what they need to know and be able to do, how this will be assessed, and the consequences of assessment.

Student self-assessment can provide a greater degree of ownership of the assessment process to students, but can also provide a benefit to both teachers and students as students begin to internalise performance standards. Tools in current use include questionnaires, rating scales, journal entries, checklists, and constructive written feedback. As well as assessing their own growth in knowledge and skills, students also have much to offer about their preferred learning styles and ways in which the teaching and learning process can be enhanced.

Some teachers have discovered the power of student-constructed tests, where, as part of a review of content prior to formal assessment, students are invited to create assessment tasks that they believe would assess fairly the key ideas in the topic under study. The students then work on each other's tasks. This has proved to be excellent revision, and is particularly well received when the teacher makes a commitment to use at least some of the student-created tasks in the final assessment.

In trying to make mathematics assesssment meaningful and manageable, attention has turned to teachers making greater use of the information they gain during instruction, as compared to the holding of "assesssment events". Professor Alan Schoenfeld from the University of California, in teaching mathematical problem solving to high school students, advised the students at the start of the term that he reserved the right, at any time, as he moved around the classroom to ask individuals one of three questions in relation to their work on particular problems: (1) "what are you doing?"; (2) "why are you doing it?"; and (3) "how does it help you in solving the problem?" (see Schoenfeld, 1992). After several weeks, he noticed that students would notice him coming and he could see their minds ticking over as they said to themselves: "Oh-oh, here he comes. . . . What am I doing? Why am I doing it? How does it help me to solve the problem?" By the end of the

term, he found he no longer needed to ask the questions as the students had internalised the questions, and increasingly become "metacognitive"—assessing their own thinking. In this way, the process was both informative for the teacher and the students.

As the examples above hopefully show, attempts to make assessment more of an open process has benefits for both teachers and students.

Rich Assessment Tasks: Some Features and Some Examples

So what kinds of assessment tasks and processes are most powerful? All teachers develop favourite assessment tasks—those that provide high quality, reliable information about what students know and can do. It is interesting to engage in a discussion of what makes these tasks so powerful. What are the critical features? In discussions with teachers, we have gradually formed a list of characteristics of "rich assessment tasks". Some of these characteristics are offered now.

Rich assessment tasks:

- connect naturally with what has been taught
- address a range of outcomes in the one task
- are time efficient and manageable
- allow all students to make "a start"
- engage the learner
- can be successfully undertaken using a range of methods or approaches
- provide a measure of choice or "openness"
- encourage students to disclose their own understanding of what they have learned
- allow students to show connections they are able to make between the concepts they have learned
- are themselves worthwhile activities for students' learning
- provide a range of student responses, including a chance for students to show all that they know about the relevant content
- provide an opportunity for students to transfer knowledge from a known context to a less familiar one
- help teachers to decide what specific help students may require in the relevant content areas
- authentically represent the ways in which the knowledge and skills will be used in the future.

We will now share briefly two examples of tasks that we believe to be rich, outlining some of the characteristics we believe they possess. Further examples can be found in Beesey, Clarke, Clarke, Stephens, & Sullivan (1998) and Clarke (1998).

1. Academy Awards, Age and Gender

Each year, we hear of the winners of the Academy Awards (the "Oscars") in the United States. The back-to-back stem and leaf plot below shows the ages of the Best Actors (male and female) for each year up to 1997.

Actors (male)		Actors (female)
	2	1244444
	·	56666667778889999
443322110	3	000011233334444444
9988888887775555	·	5556778889
443333222111110000	4	01111122
999888776655	·	5589
432211	5	0
6665	·	
2100	6	0112
	·	
	7	4
6	·	
	8	0

3 | 2 means 32 years

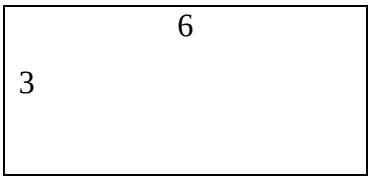
1. Use these data to find the median age of male winners and median age of female winners. Please write these below:

2. Write approximately 100 words about some things you've noticed from the data, and some possible reasons for what you've observed. (Please use the terms "median", "range", and "outlier" in your discussion if possible.)

Although other characteristics may be evident, experience with this task shows that it has the potential to

- address a range of outcomes in the one task (measures of central tendency, data analysis, understanding and use of appropriate mathematical terms, mathematical communication);
- engage the learner (many students follow the Academy Awards each year, and so this tends to provide a motivating context); and
- help teachers to decide what specific help students may require in the relevant content areas (the ways in which students analyse these data provides clear information to the teacher of their understanding of relevant content).

2. Area = Perimeter



Consider the rectangle with dimensions 6 units by 3 units. We can calculate easily that the area is 18 square units and the perimeter is 18 units. So, if we ignore the units, *the magnitude of the area and perimeter are the same* for this shape.

Investigate the following questions, reporting what you find:

- Are there other rectangles that have this property? Please explain.
- Are there any circles for which area = perimeter? (ignoring units) Please explain.
- Are there any squares for which area = perimeter? (ignoring units) Please explain.
- (Extension task) Select another kind of shape (e.g., triangle, hexagon, etc.), and explore the situations in which the area is the same as the perimeter, ignoring units.

The use of this task with students from age 14 to 18 leads us to believe that it has the following characteristics:

- it allows all students to make "a start" (we have found that even the students who struggle with mathematics are able to start considering shapes, and test whether they have these properties, using a "guess and check" strategy)
- it allows students to show connections they are able to make between the concepts they have learned (students spend a lot of time on the topics of area and perimeter in class, and this combines the two concepts in a way that reinforces both the relationships and differences for various shapes); and
- it provides a range of student responses, including a chance for students to show all that they know about the relevant content (this is a task that is particularly motivating for very capable students, as it leads, particularly in the case of the rectangle, to generalisations that can be represented with asymptotic graphs—it is interesting that a task that starts in the domains of geometry and measurement moves into algebraic expressions, manipulation and graphing at the higher grade levels).

Of course, there are few tasks that have all of the listed characteristics of rich tasks, but these points may serve as a kind of checklist against which the appropriateness of a given task can be measured. Readers may also wish to add to (or subtract from) the list. Earlier, we discussed

some of the disadvantages of pen and paper tasks. We have found that the use of these kinds of tasks, with discussion with students about their written products, as they are working on them, afterwards or both, can provide opportunities for students to show what they know and can do, as well as giving important information for teachers that can help them with planning for further instruction.

Making More Holistic Judgements of Students' Work in Mathematics

One interesting development in assessment in recent years, has been the use of scoring "rubrics" with a greater emphasis on making more *holistic* judgements on student work, with less emphasis on counting up "rights and wrongs". Rubrics are brief outlines that describe the content and quality of work needed to achieve a specific grade in an open-ended task, and enable the person assessing the piece of work to determine the evidence of students' understanding and communication of that understanding. Such an approach can enable attention to "big ideas", rather than simply facts and procedures. An example of such a rubric is given below.

Score	Summary/ Description
Goes beyond (4+)	Fully accomplishes the task, but uses methods, makes interpretations, or shows insights or generalisations significantly beyond those expected at this level.
4	Task accomplished. Central mathematical ideas clearly understood and demonstrated and understood.
3	Substantial progress towards completing the task; indicative of understanding of relevant knowledge, concepts and skills; but some key ideas may be missing.
2	Attempt at the task makes some progress; partial but limited grasp of the central mathematical ideas; reveals gaps in knowledge, conceptual understanding and/or relevant skills.
1	Little progress or understanding evident.

The use of such a rubric contrasts with the notion of "counting up" right and wrong parts of the question in assigning a score to a piece of student work.

In Conclusion

Educational and assessment policies come and go, but the major purpose of assessment will continue to be *to inform teaching and learning*. Such an orientation will impact on the kinds of tasks and processes we use for assessment, the timing and use of such tasks and processes, the ways in which students view assessment, and the subsequent quality of our teaching. The most effective teachers are likely to be those who approach assessment as an opportunity for students to show what they know and can do, who link assessment and teaching (making these links clear to students), and regard assessment as a prelude to action.

References

Beesey, C., Clarke, B. A., Clarke, D. M., Stephens, M., & Sullivan, P. (1998). *Effective assessment for mathematics*. Carlton, Victoria: Board of Studies/ Addison Wesley Longman.

Clarke, D. J. (1997). *Constructive assessment in mathematics.* Berkeley, CA: Key Curriculum Press.

Clarke, D. M. (1998). *Securing the future: Subject based assessment materials for the school certificate—Mathematics*. Ryde, New South Wales: Department of Education.

Clements, M. A. (1995). Assessing the effectiveness of pencil-and-paper tests for school mathematics. In B. Atweh & S. Flavel (Eds.), *Galtha: MERGA 18: Proceedings of the 18th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 184-188). Darwin, NT: University of the Northern Territory.

Newman, A. (1983). *The Newman language of mathematics kit*. Sydney, Australia: Harcourt, Brace & Jovanovich.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-370). New York: Macmillan.