The Mathematics Educator
1998, Vol. 3, No. 1, 26-37

Assessing Problem Solving: Give And Take

Kaye Stacey
Barry McCrae

Introduction

In Australia, as in Singapore, teaching mathematics as a creative subject focussed on problem solving is recognised as an important goal. For nearly a decade now, this has been the position in all official curriculum statements. When this goal first came to prominence in the early eighties, many innovative curriculum materials were produced and many educational sessions for teachers were held giving new ideas and discussing classroom issues. As a consequence, many teachers experimented with new classroom activities. Although the change in teaching mathematics was widespread, the actual proportion of teachers actively making changes was small and therefore relatively few children were seen to benefit. After several years, it was obvious that there was little incentive for most teachers to change their teaching methods and goals, especially as the assessment system remained unchanged. The strategy of achieving change in curriculum and teaching by change in assessment was therefore attractive.

This article outlines one such instance of assessment driven change in secondary schools in one part of Australia. There is no doubt that changing the assessment had a large, quick and lasting effect on teaching throughout the schools (Clarke & Stephens, 1996). However, the thrust of this article is to examine how the assessment itself has been altered by the total school system and how community acceptance of assessment of broader curriculum goals has gradually been increased. Two features stand out. Firstly the introduction of a test, to supplement project work, has allayed community unease about cheating. Writing the test poses many challenges, which have impacted on the style of the project. Secondly, pressure from teachers, parents and students has also influenced the style of the project. On balance, we judge that 'creative problem solving' is still being assessed, although in a curtailed way.

The Situation and the Constraints

In Victoria, the state of Australia with the second largest population, about 45 000 students aged around 18, completed Year 12 each year. The students in Year 12 work towards the Victorian Certificate of Education (VCE), from which a score is derived that strongly influences which university course they can enter. Teachers, students and their families regard the VCE and the scores as very important. There are several mathematics subjects, but here we will only discuss *Specialist Mathematics*, the hardest mathematics subject and generally taken by students who are likely to need mathematics in their tertiary studies. Since 1989, one of the components of the *Specialist Mathematics* assessment has been a problem solving task, contributing 33% of the marks for the subject. This task is set by an external examination panel but marked by teachers and there have been various schemes employed to maximise fairness in the marking across the state. Students work on the task at home and at school and the report that they write is the basis of their assessment.

Attributes of Creative Problem Solving

The problem solving task was introduced because it was felt to be important that all components of mathematics that are regarded as valuable should be visible in the assessment schemes. Problem solving in the VCE is defined as 'the creative application of mathematical skills and knowledge to solve problems in unfamiliar situations, including real-life situations' (Board of Studies, 1996, p. 7). We have analysed the extent to which the set tasks assess problem solving under this definition, by examining a number of dimensions of problem solving. The key dimensions on which we judge this are creativity (two aspects), unfamiliarity (three aspects), the need to generalise, the nature of the context and the time over which a student can be immersed in solving the problem. More particularly, they are:

1. *Creativity of direction* – does the student have some input into the direction of the investigation, deciding to some extent what questions should be answered?
2. *Creativity of method* – is some creativity and insight required to discover how to solve the problem?
3. *Unfamiliarity of solution* – to what extent does the solution require the student to combine familiar ideas in new ways?
4. *Unfamiliarity of techniques* – to what extent does the solution require the student to use mathematical techniques which they may have not previously learned?
5. *Unfamiliarity of context* – how well known to the student is the context (real world or mathematical world) in which the problem is set?
6. *Opportunity to generalise* – to what extent does the problem require the student to go beyond a single case and to identify general results?

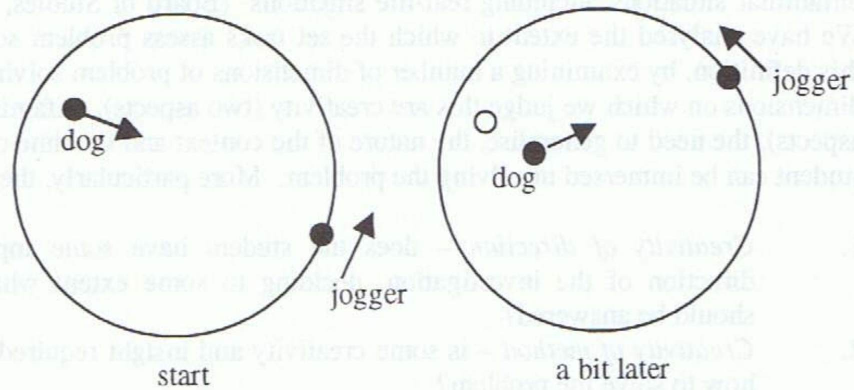
7. *Real world context* – is the problem a genuine real world problem, a contrived real world problem or is it set within the mathematical world?
8. *Extended time for thinking* – does the solution demand (or can it be given) time for 'perspiration' and 'inspiration'?

The VCE Problem Solving Task 1989 – 1997

In this section and the next section, we briefly outline the stages that the assessment of problem solving has gone through since its introduction in 1989 – for a fuller description see Stephens and McCrae (1995). An abbreviated example of one of the first VCE problem solving tasks is given in Figure 1.

Dog and Jogger

A jogger runs around a circular track at a constant speed. The jogger's dog has wandered off and the jogger whistles it to come back. The dog runs at a constant speed towards the jogger. The dog's direction changes as the jogger continues running around the track.



- If the dog starts inside the circular track and runs more slowly than the jogger, what path will the dog follow?
- What path will the dog follow if it runs more quickly than the jogger? What if the dog and jogger run at the same speed?
- What is the dog starts outside the track?
- How do your results compare with the behaviour you would expect of a real dog in the situations described above?

Figure 1. One of the Problem Solving Tasks set in 1989

Students had four weeks to work on this problem (or a number of alternative problems) at home and at school, with some assistance from the teacher within defined bounds, and presented their work as a written report of about 1000 words. There are a number of features to observe. The task is very open, and students themselves can decide significant features of what they are going to do. The task can be tackled in various ways, some of which might involve mathematics that students will not have learned at school (such as calculus with vectors). On the other hand, some progress could be made on the task using very little high-level mathematics, for example by drawing diagrams. There is an opportunity for students to formally compare the mathematical solution with the real situation. Some of these characteristics are recorded in Table 1.

Table 1 : Ratings of VCE Tasks on the Definition of Problem Solving

Feature	*Dog and Jogger (1989)*	*The Great Escape (1997)*	*Trends from 1989-1996*
Creativity of direction	Medium	Low	Reducing
Creativity of method	High	Medium	Reducing
Unfamiliarity of solution	High	Medium	Reducing
Unfamiliarity of techniques	High or Medium	Low	Techniques must be in syllabus now
Unfamiliarity of context	Medium	Medium	No change
Opportunity to generalise	High	Medium	Generalisation must be present
Real world context	Medium	Medium	No change – seen as a key feature
Extended time for thinking	4 weeks	2 weeks	Reduce, but still substantial

The Introduction of the Problem Solving Test

After several years of operation, public confidence in the problem solving task was eroded because of allegations that many students were submitting solutions that were not their own – having been copied from other students, done by parents or privately paid tutors, or purchased from other sources. As a consequence, the task was redesigned. From 1994, the time to work on the problem was shortened to 2 weeks. It seemed likely that there would be less opportunity to obtain solutions from others if there was less time. The shorter time period is also less disruptive to school programs, lessening the time when students are preoccupied with the assessment. The task was made more prescriptive, changing from an open set of directions to a sequence of questions. This was to assist teachers in assessing the work reliably, to help students get started, and to be clearer about expectations. The problems set were chosen so that mathematics from the syllabus was directly relevant and other mathematical knowledge was probably not of much help. Students were given a choice of three problems, one of which was set 'within the mathematical world'. This pure mathematics problem was only considered 'moderately suitable' by teachers (McCrae, 1995) who apparently see a real world setting as a key feature of problem solving.

In addition to these changes to the problem solving task, all of which have been retained up to now, a test designed to assess understanding of the solution was added. This is a 1 hour test, done at school and marked by the teachers, to confirm that students understand the work in their reports. The composition of the final grade is 60% from the problem report and 40% from the test. If a student's mark on the report is significantly higher than his or her mark on the test, then the student is interviewed by teachers at the school to find out why. Student who are found not to have authored their reports are disciplined.

One of the problems set in 1997, *The Great Escape*, and the associated test are given in Figures 2 and 3. *The Great Escape* is much more structured than the earlier tasks, making it easier for teachers to mark, easier for students to structure their work and easier to ensure that all students have had a common experience on which the test can draw. There is a specific part of the problem (q6) that asks students to make a generalisation. When a generalisation was not asked for in 1994 (the first year of the redesigned task), the very closed nature of the problem again gave more opportunity for dishonesty and teachers and examiners felt that the problem format had seriously reduced its validity as a measure of problem solving ability (McCrae & Stacey, 1997). The characteristics of *The Great Escape* are summarised in Table 1.

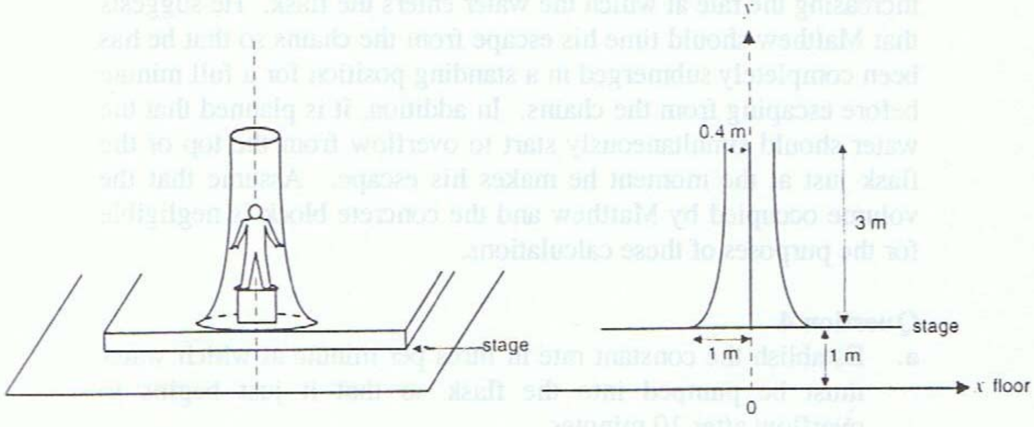
The Great Escape

The mathematical techniques which might be required for this task include:

functions – domain and range, and their graphs
volumes of revolution
calculus
logarithms and exponentials
approximate methods of solutions of equations.

Matthew is an escape artist who makes a living performing daring escapes from dangerous situations. He believes he has developed an escape trick that will prove he is the equal of the great escape artist Houdini. In his escape his feet are chained to the top of a concrete block attached to the bottom of an enormous flask. He makes his escape from the chains as water is pumped into the flask. The flask has a circular transverse cross section. Its longitudinal cross section takes the shape of a curve with equation

$y = \frac{a}{x^2} + b$, using the floor as the *x*-axis and the vertical centre of the flask as the *y*-axis (refer to the diagram).



Question 1
Find the values of a and b.

Question 2

a. Find, in litres, the total volume of water that would be required to completely fill the flask, giving your answer to the nearest litre.
b. Express the volume of water (in litres) in the flask as a function of the height of the water above the floor.

When Matthew is placed in the flask, he plans to make his dramatic escape just as the level of water reaches the top of his head, when he is standing upright. Matthew is exactly 180 cm tall and it takes him 10 minutes to escape from the chains on the top of the concrete block. Assume that the volume occupied by Matthew and the concrete block is negligible for the purposes of these calculations.

Question 3
Find the height of a concrete block required which will allow Matthew to make his escape after exactly 10 minutes, when water is pumped into the flask at a constant rate of 300 litres per minute.

Matthew's manager plans to make the escape more exciting by increasing the rate at which the water enters the flask. He suggests that Matthew should time his escape from the chains so that he has been completely submerged in a standing position for a full minute before escaping from the chains. In addition, it is planned that the water should simultaneously start to overflow from the top of the flask just at the moment he makes his escape. Assume that the volume occupied by Matthew and the concrete block is negligible for the purposes of these calculations.

Question 4

a. Establish the constant rate in litres per minute at which water must be pumped into the flask so that it just begins to overflow after 10 minutes.
b. Hence, find the height of the concrete block required for this situation.
c. Express the height of the liquid in the flask, in terms of the time in minutes, after filling has commenced.

d. **i.** Find an expression that Matthew could use to find the height of a block for a given submersion time. [The submersion time is the time between the water reaching the top of Matthew's head (when he is in a standing position) and it reaching the top of the flask.]

ii. Sketch the graph of height versus submersion time.

Question 5

Matthew now realises that he has not considered his own volume or the volume of the block. Assume that the block is a cylinder of radius 13 cm, and that Matthew is a cylinder of radius 13 cm and height 180 cm.

a. If the block is of height 60 cm, find the number of litres of water to be pumped into the flask so that Matthew's head is just covered, when he is standing upright.

b. **i.** Suppose the block is of height H cm. Express, in terms of H, the number of litres of water to be pumped into the flask to just cover Matthew's head.

ii. If water is pumped in at 300 litres per minute, find the height (in cm) of the block so that Matthew's head is just covered in exactly 10 minutes.

Question 6

Matthew discovers that the pump can be adjusted to produce a constant flow rate of between 180 and 330 litres per minute and that he can change both the radius and the height of the cylindrical block.

a. Suppose that the block is of height H cm and radius r cm. Express, in terms of H and r, the number of litres of water to be pumped into the flask to just cover Matthew's head.

b. Establish 5 sets of values for the flow rate H and r which will allow Matthew's head to be just covered in exactly 10 minutes, commenting on any possible restrictions on the values of these variables.

Figure 2. One of the Problem Solving Tasks set in 1997

Having the test has increased public credibility in the assessment, but it has imposed its own demands on the nature of the task and raised other issues. If the test is to assess students' understanding of how they did the problem (and this is

Can Problem Solving Survive in High-stakes Assessment?

In the last column of Table 1, we have tried to identify the overall trends in the assessment from 1989 to 1996 on the key dimensions of problem solving. As can be seen from the table, many of the characteristics of the early problem solving tasks are now significantly reduced or not present. Clearly the problem and test is less an assessment of creative problem solving than before. However, the new arrangement has support from teachers as practical and fair (see McCrae & Stacey, 1997). Cheating in the take-home assessment tasks of the VCE is still an issue of public concern as we are writing this, yet another article has appeared in the Melbourne paper on this issue (Brett, 1997), but the mathematics problem solving task is no longer regarded as a particularly vulnerable area. The compromises reached have ensured that the task is not too onerous to conduct in the school, that students can be reasonably sure when they have done enough work, that there is sufficient public confidence that cheating will be identified and dealt with and that teachers can grade the work reliably across the state. These are important advances. Without these issues being attended to, the task would not still exist.

So is the task still a valid assessment of problem solving? Students still find it challenging. It provides them with one of the few experiences of working intensely on a substantial mathematics problem that they experience at school. Although we have no direct evidence, we believe that schools are still preparing their younger students as well for creative problem solving as they were in the first years after the VCE was introduced. This, we feel, is an important criterion for the success of assessment driven change. Our judgement, on balance, is that the current arrangements are worth defending because of the way in which they continue to encourage creative problem solving in schools.

References

Board of Studies. (1996). *Mathematics study design (1997-2000).* Carlton, Victoria: Author.

Brett, J. (1997, August 14). Morality cheated in VCE tests. *The Age*, p. A7.

Clarke, D. & Stephens, M. (1996). The ripple effect: The impact of the systematic introduction of performance assessment in mathematics. In M. Birenbaum & F.J.R.C. Dochy (Eds.), *Alternativeness in Assessment of Achievements, Learning Processes and Prior Knowledge* (pp. 63-92). Boston: Kluwer.

McCrae, B. (1995). An evaluation of system-wide assessment of problem solving at year 12 by report and related test. In B. Atweh & S. Flavel (Eds.), *Proceedings of the 18th Annual Conference of the Mathematical Education Research Group of Australasia* (pp. 409-415). Darwin, Northern Territory, Australia: Mathematical Education Research Group of Australasia.

McCrae, B. & Stacey, K. (1997). Testing problem solving in a high-stakes environment. In E. Pehkonen (Ed.), *Use of open-ended problems in mathematics clasroom* (pp. 34-48). Helsinki: Department of Teacher Education, University of Helsinki.

Stephens, M. & McCrae, B. (1995). Assessing problem solving in a school system: principles to practice. *Australian Senior Mathematics Journal, 9(1)*, 11-28.