Reforms in assessment practices in mathematics: Innovations or rigmaroles?

Wong Khoon Yoong

The Mathematics Educator, 1(1), 67-81

Association of Mathematics Educators, Singapore

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

Archived with permission.
Reforms In Assessment Practices In Mathematics: Innovations Or Rigmaroles?

Wong Khoon Yoong

Abstract

This paper begins with a series of questions about a typical assessment task. These questions highlight issues that cry out for reforms in assessment practices. The characteristics of current reforms are then compared with the traditional approach, citing recent Australian mathematics education projects as examples. Teachers are encouraged to reflect on these changes and to make the personal journey to effect changes in assessment whenever these could promote better learning of mathematics.

Introduction

In 1990, I ended a talk on problem solving to Singapore teachers with this quote by Tanner (1989): “Curiosity kills cats. The danger in teacher has always been that out [of] a desire to protect the cat, we have killed the curiosity instead” (p.269).

Assessment practices exert an overwhelming influence on what teachers teach and how students learn. However, we must not allow assessment practices to kill our students’ curiosity about mathematics. We should constantly ask ourselves whether the ways we assess our students and report on their performance will nurture their curiosity, since this is the key trait that will enable them to learn mathematics in school and beyond.

Recent reforms in mathematics education in many countries have experimented with assessment strategies that aim to promote a wide variety of curriculum objectives, ranging from skill mastery to attitude to problem solving. Some reforms are more successful than others, and we should critically examine their impacts before we decide to make changes to our practices. In this paper, I wish to arouse your curiosity about these changes by focusing on the following aspects:

- Questions about a typical assessment task
- Characteristics of recent reforms in assessment
Questions about a typical assessment task

When we talk about assessment in mathematics, we often think of asking students to solve mathematics problems. Below is a typical question and its solution.

If 40% of John’s salary is $600, how much is 60% of his salary?
Solution:  
40% = $600  
60% = $600 × 60/40 = $900

We can ask a number of questions about this item in order to highlight the need to make assessment more valid, reliable, and beneficial to the learning process.

(a) What can the student do here?

The solution suggests that the student can (a) translate a word problem into symbol, (b) apply proportional thinking and (c) compute percentages accurately. These are undoubtedly desirable outcomes and they belong to the categories of Skills, Processes, and Problem solving of the Singapore mathematics curriculum. However, being able to use a routine rule does not necessarily imply understanding the concepts underlying the rule. Likewise this item cannot assess the outcome of Attitude and Metacognition mentioned in the Singapore curriculum. The corollary is that a single test item cannot assess all the objectives we wish to achieve. That is why most reforms have included a variety of assessment tasks.

An algorithmic method is used to solve the given question, but is it the only way? In recent years, mathematics educators have documented the informal ways (called street mathematics) that people (children and adults) use to solve arithmetic problems in their daily life. For example, Nunes et al. (1993, pp.18-19) reported the following interview with a 12 year old Portuguese child (M):
Customer: How much is one coconut?
M: Thirty-five.
Customer: I'd like ten. How much is that?
M: [Pause] Three will be one hundred and five; with three more, that will be two hundred and ten. [Pause] I need four more. That is... [Pause] three hundred and fifteen... I think it is three hundred and fifty.

This kind of research shows that school mathematics is not often used in everyday life, and it raises the question: should our assessment be limited to the methods taught in the classroom only, or should it allow for alternative techniques 'discovered' by the students?

(b) **Is appropriate context provided in the test item?**

The above test item is devoid of a meaningful context. There is no realistic situation in which the calculation is used. Quite the contrary, a more likely situation is this: if John's salary is $600 and he has to set aside 40% for rent, how much is the rent? When students are used to working out problems that do not have sensible contexts, they may fail to interpret the results appropriately. The computation has lost its meaning. This is shown by the following well-known NAEP result, quoted by Nunes et al. (1993, p.54):

"An army bus holds 36 soldiers. If 1,128 soldiers are being bussed to their training site, how many buses are needed?" Of the 45,000 students who were tested, 70% did the long division correctly; however, 29% gave '31 remainder 12' and 18% gave '31' as the answer.

Schoenfeld (1988) observed: "If the 'bottom line' is error-free and mechanical performance, students come to believe that is what mathematics is all about" (p.161). We should not reinforce this undesirable belief by asking students to practise too many mechanical problems. The outcome of stressing only mechanical skills is succinctly summarised by the newspaper headline: "Singapore's top students are exam-smart but can't think for themselves" in the New Straits Times (21 March 1989).

Reformers are aware of this weakness and have extended the goal of assessment "to draw on the habits of thinking developed by students in their studies, not on specific problems they have learned to solve in their earlier work" (Measuring what counts, 1993, p.55). Adding meaningful contexts to assessment tasks will enrich the learning experience. For example, Wong &
Schibeci (1994) show how the quadratic function is embedded in an activity related to the Body-Mass Index used in health promotion.

(c) What errors has the student made in the solution?

Some teachers may consider the step: \( 40\% = \$600 \) an improper mathematical statement and deduct some marks. Is this focus on formal expression an important assessment goal? How does this fit into the current emphasis on developing communication skills in mathematics?

Mathematics can be communicated orally or in written work. Oral assessment is now increasingly used to test students' understanding of mathematics. A typical example is to ask students to explain something to a fellow student who has missed a lesson. On the other hand, written communication is promoted through the use of journal writing (Borasi & Rose, 1989) and reports of investigations and course work. Such written work is not limited to text only, as students learn to express mathematical ideas using symbols, charts, graphs, and diagrams. A concern with this assessment mode is to distinguish between language ability and mathematical understanding.

There is already extensive research into students' misconceptions of various topics in mathematics. These research findings have been used to devise learning activities that use students' errors to improve learning. For example, given the common error: \((x+y)^2 = x^2+y^2\), it is not sufficient to correct it. It is more helpful to guide students to explore the consequences of following through this error; to depict it geometrically; to look for mathematical systems in which this might be true, and so on. Szetela (1993) gave a different but interesting example: a solution to an extended problem contains an error, and primary students are asked to play the role of the teacher to explain the error to a fictitious student. A similar approach is also discussed in Herrington, Wong & Kershaw (1994).

(d) What we do not know about the student from this piece of work?

As mentioned earlier, a single item cannot assess all the intended objectives. In particular, we cannot answer the following worthwhile questions using a single test item:

- Can the student work individually and in groups as a team member?
- Is the student imaginative, creative or just imitative?
• Is he/she lazy or hardworking?
• How does the student learn? What is his/her learning style?
• What can the student use this knowledge for?
• Weak students fail tests, but know much more than we can test. How do we find out?

To answer these questions requires the teachers to use several different assessment strategies: extended tasks, portfolios, interviews, checklists, observations, learning logs, etc. All these have been enthusiastically endorsed by many reformers.

(c) What feedback do we give to our students regarding their work?

Traditionally, teachers will put a tick if the solution is correct and a cross where an error is found. The teacher may go over the common errors in class and explain how to correct them. Rarely do teachers give detailed feedback in written work, especially when they have to mark the work of many students, as is the case in Singapore.

For public examinations, teachers and students are concerned whether the marking is fair and consistent. Since the test papers are never returned to the students, they cannot countercheck the marking, contenting themselves with the grade reported. Since this is inevitable but is unduly restrictive, many reforms have developed a combination of public examinations and school-based assessment. The school-based component typically includes several assessment tasks mentioned earlier. In some cases, students are given a say in the assessment process (for example, develop a contract with students on what to assess, and use student generated questions in assessment) or the grade given (as in peer assessment).

(f) What do students do with our feedback?

We know very little about this since not much research has been done in this area. Observations will confirm the fact that most students will just look at the marks and file the work away. Some will try to understand the corrections given in class, and only the really motivated will do more examples on their own initiative. Clearly greater effort has to be made to help students use teacher feedback to improve their learning.

One successful attempt from the Netherlands is to use a two-stage testing as follows (cited in Measuring what counts, 1993, p.73):
1. Students take a test made up of short questions and essays.
2. Teacher marks test and notes major errors.
3. Students work on corrections and do more study.
4. After several weeks, students re-take a similar test.
5. The final grade is a combination of the first and second scores.

(g) What can we learn about our own teaching from assessment?

Teachers can learn much from reflecting on their practices, and, wherever possible, sharing their reflections with colleagues. Some issues for reflection are:

- What do student grades tell me about the quality of my teaching?
- Did my assessment focus unduly on mechanical skills at the expense of understanding, thinking, attitude, and genuine problem solving?
- To what extent does assessment drive my teaching style and student learning?
- How does my assessment of student relate to other types of assessment, such as common school-based assessment, public examination, student self-assessment?

Attempting to answer these questions will alert teachers to the possible need to change their current practices.

(h) What do the stakeholders think about assessment?

Different stakeholders are likely to have different perceptions of school assessment. Since mathematics is often seen as a very important subject and part of the compulsory ‘3R’, major changes in the mathematics curriculum or methods of assessment have attracted wide publicity in the media. Typical concerns of different stakeholders are given on the next page.
Teacher: Is May Lin making good progress in mathematics?
Student: I can't solve this problem. Will this come out in the examination?
Parent: Will my child get an A grade in O-Level mathematics?
Principal: Does Mr Tan give monthly tests and mark them promptly?
Inspector: Do the teachers in this school follow the ministry guidelines for examination?
Politicians: Can learning mathematics help our children to think creatively?

How do different people understand and accept changes in assessment practices? While the answer will be different depending on the countries, it is probably true that “Parents and community members need to understand that doing well by old standards is no longer good enough to secure children's future” (Measuring what counts, 1993, p.101). Assessment must satisfy the emerging needs of the individuals and society.

To summarise, we have examined a few issues that are prompted by considering a typical test item and the attempts to address these issues. The purpose of raising these issues is to hope that teachers will become more aware of them and will not take their assessment practices for granted. The next section will examine briefly the recurring themes found in many recent projects on assessment.

Characteristics of current reforms in assessment

The assessment projects being implemented in several countries share some common themes that are distinct from the traditional approach. As Olssen (1993) wrote:

much of the assessment practice in mathematics is formal, remains in the control of the teacher, requires written responses and occurs without access to much in the way of resources ... The teachers with more innovative teaching and learning situations displayed assessment practices that were more actively engaging and inclusive of students in the assessment process, and allowed more access to resources in formal assessment events. (p.100)
The contrast between the traditional practice and new projects is presented in the following table.

<table>
<thead>
<tr>
<th>Traditional Practice</th>
<th>Reforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>The goals of assessment are stated in general terms in the official syllabus. Most students are not aware of these goals.</td>
<td>The objectives are stated explicitly and in great details, such as the Standards (NCTM, 1989), Outcome Statements, and Profiles (Olssen et al., 1994). These assessment criteria are widely publicised and are sometimes given to the students.</td>
</tr>
<tr>
<td>Students learn routine rules and are tested on simple tasks that have one right answer only.</td>
<td>Learning is more meaningful and is embedded in an appropriate context. Students are expected to sue the knowledge in real situations. Open-ended or challenging assessment tasks are more complex, allowing students to attack them at different levels and in diverse ways that reflect their ability.</td>
</tr>
<tr>
<td>Tests are 'closed-book' and have to be completed within a fixed time.</td>
<td>Include take home projects. Examinations may include test items proposed by the students.</td>
</tr>
<tr>
<td>Resources used are paper and pencil. Mathematical tables, protractor, and compasses are used for specific topics.</td>
<td>Extended tasks require the use of calculators at all levels, graphing software, spreadsheet, Cabri, concrete objects, raw data and so on.</td>
</tr>
<tr>
<td>Stress individual performance, hence encourage comparison and competition among students.</td>
<td>Allow for group work. This promotes cooperation, but cheating and unfair assessment remain contentious practical problems.</td>
</tr>
<tr>
<td>Students solve problems given by the teachers.</td>
<td>Encourage students to pose their own problems and solve them (Brown &amp; Walter, 1993). Another technique is to give an answer (e.g., the profit is $100) and ask students to make up the questions (Herrington, Wong &amp; Kershaw, 1994).</td>
</tr>
<tr>
<td>Focus on what the students can do.</td>
<td>Include also how the students think and their beliefs about mathematics and mathematics learning.</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Minimise students' errors since accuracy is of paramount importance.</td>
<td>Use errors as springboards to improve understanding; see section (c) above.</td>
</tr>
<tr>
<td>Limited communication in mathematics. Workings consist of formal methods learned by imitating standard examples.</td>
<td>Test communication skills in different modes: written, oral, concrete objects. The scope of communication encompasses more than standard mathematical workings. It is important to distinguish between language and mathematical ability.</td>
</tr>
<tr>
<td>Achievement is reported by a single score or grade.</td>
<td>Use various sources of information (observations, checklists, interviews, etc.) to arrive at a profile of student achievement. Outcomes are reported quantitatively and qualitatively.</td>
</tr>
<tr>
<td>Teachers control the assessment process and reporting: students are the objects of assessment.</td>
<td>Some projects involve students in self-assessment and peer assessment.</td>
</tr>
<tr>
<td>Many systems are constrained by public examinations with historical links to outside examination bodies. These systems are fairly stable.</td>
<td>Assessment in a state of flux with new techniques, such as problem posing and group work, being tested and modified.</td>
</tr>
<tr>
<td>Mathematics is treated as an abstract subject that is the same all over the world. Hence, assessment does not take into account cultural differences.</td>
<td>Assessment tasks take into account the cultural background of students; see Nelson, Joseph &amp; Williams (1993).</td>
</tr>
</tbody>
</table>

The traditional methods are not necessarily wrong but they are too narrow for measuring a wide range of curriculum goals. The alternative strategies, using extended tasks and multiple approaches, have greater potential to improve the quality of teaching and learning.
Reforms of mathematics assessment in Australia

As a consequence of traditional responsibilities and political rivalries, there is no national assessment in Australia. Australian educators are also suspicious about the way any national testing results might be misused for political gains. They also question the validity and reliability of such a national examination system.

Despite these unfavourable conditions, the Australian federal government has introduced a national framework for school curriculum. The main developments are described below.

(a) 1989 - School Goals

The former Australian Education Council (now called the Ministerial Council for Youth, Education, Employment & Training) proposed 10 common goals for schooling in Australia. It also published national standards and profiles for eight areas of learning: English; mathematics; science; technology; language other than English; health and physical education; studies of society and environment; and the arts. While these goals and profiles may dictate teaching sequence, they do not constitute a national curriculum, or a national test, or a certification to enter tertiary institutes.

(b) 1991 - The National Statement on Mathematics

The Mathematics Statement, published in 1990, aims to establish a common language for curriculum development and innovations in teaching, learning and assessment. It divides the school mathematics curriculum into eight strands: Attitudes and appreciations; Mathematical inquiry; Choosing and using mathematics; Space; Number; Measurement; Chance and data; Algebra. Within each strand, the mathematics content is divided into four broad bands covering Years 1 to 12. Sample activities are suggested to achieve the broad objectives. A constructivist approach is emphasised.

As an example, "experience with applying mathematics should be provided which enable students to represent practical problems using objects, diagrams, symbols or mental images". Possible activities include: "use visual imagery to represent a problem situation ... investigate the use of mathematics in fiction or daily newspaper" (AEC, 1990, p.68).
To help teachers, parents and the community understand the philosophy and practice underlying the Statement, training workshops were organised throughout the country. The materials for these workshops have been published as the *Maths Work Series* by the Australian Association of Mathematics Teachers.

### 1993 - Profiles

These Profiles are based on the Statement. It covers six strands: Algebra; Chance and data; Measurement; Number; Space; Working mathematically. The Profiles divides attainment in each strand into eight levels covering Years 1 to 10. They are meant to be used for reporting student achievement in a nationally comparable way. Work samples have been published to assist teachers in assessing the levels attained (Olssen et al., 1994). There is no compulsion for the various States to adopt the Profiles. However, several States have modified the Profiles to suit local requirements.

### Responses to the reform

There were mixed reactions to the reform. In general, most mathematicians objected to the perceived lack of balance between mastery of skills and personal 'experiences' with mathematics. They also complained that they had not been properly consulted (Bryce, 1992). On the other hand, mathematics educators, teachers, and ministry officials are actively developing materials and strategies to reflect the changes in philosophy and expectations. Given such a large scale reform, modifications are to be expected in the coming years.

### Some State projects

Most States in Australia have implemented reforms in the mathematics curriculum based on local requirements and in response to the national initiative.

In Western Australia, the Monitoring Standards in Education project, which began in 1990, has monitored student progress in Years 3, 7 and 10 in mathematics for about 1500 students at each level. The results were linked to the official Outcome Statements, which delineates six levels of attainment in five strands. The average levels of attainment in Algebra and Number are shown in the table (adapted from Titmanis et al., 1993) on the next page.
It is clear from the above table that the average student who graduates from WA high school has not reached the highest Level 6. Recently, the WA Ministry of Education has decided that, from 1996, high school students can graduate from Year 12 without having to study mathematics in Years 11 or 12 (Western Australian, 10 May, 1995, p.1). These changes have led to heated debate about the deterioration of mathematics standards among WA students.

Another State reform is the mathematics curriculum under the Victorian Certificate of Education (VCE) for Years 11 and 12 students. This reform began in 1991 and is now in its second phase of implementation (Michell & Stephens, 1995). The assessment consists of a mix of school-based assessment and end of the year written examinations. The main components are (CAT = Common Assessment Task):

- CAT 1: Year 11: Investigative project, not counted in final grade
- CAT 2: Year 12: Problem solving, school-assessed (60%), verification test (40%)
- CAT 3: Facts and Skills - Multiple choice, externally set
- CAT 4: Analysis Tasks - externally set

The strength of the VCE mathematics courses is "to provide challenging and relevant experiences for a wide range of students" (ibid, p.12). There are still practical problems like pressure on the teachers and
validating the authenticity of students' project work that have not been resolved satisfactorily yet.

Conclusion

The above brief survey shows that an enormous amount of resources has already been spent on implementing innovative projects to improve mathematics assessment. Are these reforms effective in achieving the intended objectives? Research on these reforms is rather scanty, and the answer is far from certain. But one particular message seems to emerge clear and loud from all these efforts: there is no going back to a single final examination as the only method to assess student learning.

In recent years, the learning of mathematics in Singapore has been enriched with interesting activities such as Mathematics Camps, as reported in the ministry's Mathematics Newsletter. However, public examinations remain the main assessment method. This puts a tremendous pressure on students to excel in these examinations and, to a large extent, it has maintained credible standards in mathematics among Singapore students. As Resnick and Nolan (1995) have found: "Countries known for their outstanding students have several practices in common: clear, consistent, demanding public education standards head the list" (p.6). Given such a solid base, it is time for Singapore teachers to fine tune their assessment practices in order to help students learn even better.

Teachers are being bombarded with many messages about assessment: use a variety of alternative techniques; thinking is more important than skill mastery; attitudes and beliefs count; prepare students to work in groups; profiles are more valid than single grades; and so on. The teachers, being professionals, have to decide what changes to implement in their own classrooms and how to do it. Making changes is a personal and risky journey. It requires courage, especially in a kiasu culture. The following from Measuring what counts (1993) sums up the challenge to teachers very neatly:

"The destination for the voyage of reform is well-known: every student must learn more mathematics. All educational actions must support this goal, and assessment is no exception." (p.141)
Teachers are the captains, charged with the front-line responsibility of providing high-quality mathematics education to all students. (p.140)

References


