The Mathematics Educator
1996, Vol. 1, No. 2, 159-171

Problem Posing As Assessment: Reflections And Re-constructions

Shukkwan S. Leung

Abstract

The study is an attempt to introduce assessment ideas, such as problem posing, in an elementary teacher-training programme in Taiwan. Here, problem posing is restricted to mean the posing of problems for a mathematics test. During teaching practice, student teachers self-assessed their own tests and provided written accounts on their reflections and re-constructions. It was found that student teachers, when reflecting, were aware of strengths and weaknesses of problems they initially posed. In general, student teachers tended to re-construct initially posed problems according to mathematical complexity. In particular, they revised instructions and formats of those problems. There were also instances when they chose to alter scoring while keeping posed problems intact.

Statement of the Problem

According to the recent document on assessment standards by the National Council for Teachers of Mathematics (NCTM) in the United States, an assessor can be a student, a teacher, a district committee, a supervisor or a test publisher (NCTM, 1995, p.5). Among these alternatives, class teachers are in the best position to assess students' performance (NCTM, 1991, 1995). Therefore, in teaching practice, it is appropriate for instructors to ask student teachers to be assessors of their school pupils, to make up their own tests rather than to use tests from outside sources. In doing so, the instructors act as assessors of student teachers: assessing student teachers' ability to pose problems for a test and finding out what mathematics contents and thinking processes they valued.

To some extent, learners as assessors can be a complement to teachers as assessors. In fact, "research has suggested that students who are able to evaluate their own thinking and learning process often have higher achievements" (Kulm, 1994, p.73). Therefore, in teaching practice, it is also appropriate to ask student teachers to do self-assessment. For example, they can self-assess their making-up of a test. This self-assessment means reflecting on one's process and one could collect reflective thoughts in journals. Journal writing, according to Castle and Aichele (1994), is a means to illustrate teachers' reflective insights, which contribute to professional knowledge used to make autonomous decisions. However, in students' self-assessment, the instructor assumes new roles: fostering students' own thinking (e.g. providing thought questions or checklists) and collating students' thoughts in class discussions (Kenny & Silver, 1993). Furthermore, the instructor can also self-assess, thus, finding out the possible impacts of a course in the teacher-training programme.

The aim of this study was to trace, from student teachers' journals, their reflections and re-constructions on their own problem posing. Here, problem posing means the posing of problems for a mathematics test. The main focus is on how student teachers posed problems for a test and how they self-assessed their tests. This study is feasible in a teachers college, where student teachers fulfill course requirements and practice teaching simultaneously.

Methodology

The School Calendar and The Course

In Taiwan, a teachers college prepares elementary school teachers. At the fourth year in a teachers college, all students do student teaching while finishing course requirements. The course, Elementary Mathematics Curriculum, is a compulsory year course for all fourth year students electing Mathematics. However, the course is interrupted in the Spring term, when students are off for a four-week teaching practice in the elementary schools.

In the 1993 school year, there were 13 fourth year students who took Elementary Mathematics Curriculum. The course contents were developed by the instructor. During the Fall term (October through January), issues on curriculum standards (NCTM, 1989; Education Department, 1993) were discussed and examples in the local textbook were used. Specifically, the course emphasized problem solving and problem posing activities. In doing problem posing, students posed problems individually and then they helped each other to revise and edit. In the Spring term (March through June), alternative assessments issues were discussed: part marks by SCME[1]; scoring rubrics by CAP[2] (Ridgway, 1988; Stenmark, 1989); realistic tests in the MORE[3] project (Van den Heuvel-Panhuizen,

[1] SCME, Shell Centre for Mathematical Education in England
[2] CAP, California Assessment Program in the United States
[3] MORE, One project of the Realistic Mathematics Education in the Netherlands

1990); and, alternative assessments like journals or peer evaluation in MCTP[4] (Clarke, 1988). During teaching practice, students completed an assignment and kept journals. After teaching practice, when there were two more class meetings, the instructors collated ideas from students' journals and conducted seminar discussions. The seminar discussions enabled a display of students' reflections and re-constructions of tests while allowing peer evaluation. In peer evaluation, each group assessed tests made by peers of another group. They helped by identifying strengths and suggesting revisions.

The Assignment

The instructor designed an assignment for students to complete during student practice. The assignment is a part of the requirements of the course. Each student was to make up a written test, to administer it and to do grading. Then the student was to reflect on the test and specifically re-construct the test. Specifically, the instructors provided two thought questions. The first one is on how they assessed pupils' performance, "If the test is designed properly then what does it tell you about pupils' performance?" The second question is on self-assessing the make-up of the test, "If you were to use the test again how would you revise it and why?" Note that students already had a habit of keeping assignments in journals. This assignment only constitutes one part of the journals.

Design of Study

Using students' journals as data, the investigator addressed multiple assessments by different assessors. Figure 1 depicts the overall assessment framework of the investigator or the design of this study. The main purpose is to see how student teachers posed problems for a mathematics test and how they would revise the tests after administering it once during teaching practice. The ancillary purposes are: to study how student teachers assessed pupils' performance and to see the possible impacts of the course. Due to limitations of space, results corresponding to ancillary purposes will be reported briefly.

---

[4] MCTP, Mathematics Curriculum and Teaching Program, Curriculum Development Centre in Australia

Figure 1. The Overall Assessment Framework.

Assessor/Learner	Instructor	Student Teacher	Pupils
Instructor	Impact of Curriculum Course	POSING PROBLEMS FOR A MATH TEST	
Student Teacher		POSING PROBLEMS FOR A MATH TEST	Assessing Mathematics Performance

Data Analysis

In the thirteen tests, different mathematics contents were covered. Of all, five were on the multiplication and division of fractions; two on rounding decimals; another two on cubes and cuboids; and the remaining four were on averages, multiples, time, and miscellaneous topics respectively. To bring a central focus to this paper, the investigator will only discuss examples from tests on multiplication and division of fractions. There were two levels of analysis: test level (5 sets) and problem level (18 items).

The Instructor as Assessor

The instructor assessed coverage and complexity of tests. Coverage here means whether or not a variety of contexts are included. Complexity means whether or not problems of varying degree of challenges are complied. In complexity, the instructor addressed to two dimensions: problem type and mathematics structure. By problem type, the instructor used Butts' classification (1980) to partition problems as Recognition, Algorithm, Application, Open-search or Problem Situation category. By mathematics structure, the instructors referred to Vergnaud's multiplicative structures (1983): Isomorphism Of Measures, Product Of Measures, and Multiple Proportions. Subsequently, problems in Isomorphism of Measures category were further classified as Multiplication, Quotitive Division, Partitive Division, or Rule-of-three (in ascending order of complexity) while problems in Product Of Measures category were further categorized as Cartesian Products and Rectangular Area. Finally, Multiple Proportions category refers to the multiplication of multiple quantities. Vergnaud's classification thus included seven categories altogether.

Student Teachers as Assessors

This part of analysis included both having student teachers as assessors of themselves and of their peers. Since the instructor purposely made student assessment open and that student journals were free format, the analyses of students' reflections and re-constructions were simply open and descriptive. No particular scheme of analysis was pre-defined. This part of analysis was done by referring to students' directions of changes, as given in their journals.

Results and Discussions

In the assignment, two thought questions were given by the instructor. The first one was on how they assessed pupils' performance, "If the test is designed properly then what does it tell you about pupils' performance?" The second thought question was, "If you were to use the test again how would you revise it and why?" By referring to answers to these two questions, the instructor wanted to assess students' problem posing and as well, see how student teachers reflected upon their own posing and re-constructed the tests. Finally, the instructor could self-assess and attempt to find the impact of the course from students' problem posing. This part of analyses will be reported separately, with additional analyses on other parts of students' journals. This assignment is only one part of students' journals.

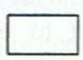
There were altogether five tests, comprising a total of 18 problems on multiplication and division of fractions. On the analyses of five sets of tests, the presentation of results was on reflections and re-constructions. It began with Adrian as an example on assessment on one person's problem posing with multiple persons as assessors. Then it would close with collated thoughts on how the other four students self-assessed their tests. The analyses on 18 problems followed. The instructor reported on what was important to be tested in multiplication and division of fractions. This was done by classifying the 18 problems according to problem type and multiplicative structure.

Multiple Assessors of Adrian's Problem Posing

The Instructor as Assessor

The instructor referred to Adrian's test and assessed coverage and complexity. It was found that Adrian followed the general rule-of-thumb on coverage and complexity. On coverage, she posed five problems with a variety of contexts ranging from books, area, weights, and algorithm. On complexity, the five problems she posed were numbered in ascending order of difficulty according to Butts (1980): the first three are Application problems, the fourth is an Open-search problem and the fifth is a Problem Situation. Among the first three Application problems, she ordered them according to the complexity of multiplicative structure by Vergnaud (1983): a multiplication problem of isomorphism of measures preceded the problem on rule-of-three. Figure 2 is her actual test paper translated in English. Her comments after self-assessment, which was absent in pupils' test paper, were typed in *italics*, on the right side of Figure 2.

Figure 2. Adrian's Test.

Test Item	Adrian's Comments
1. A book is $3\frac{1}{2}$ cm thick. If 5 are piled together, what is the total thickness?	*Applications, numbers and wording alright, easy to read, daily life, a basic question*
2. The area of a triangular piece of land is $54\frac{3}{4}$ m^2, the base is 9m, what is the height?	*It is a challenging word problem*
3. A steel rod of $1\frac{1}{2}$ m long weighs 3 kg. What is the length of a steel rod weighing 1 kg? Weighing 5 kg?	*Test if they know when to multiply and when to divide*
4. A piece of cake was divided equally among four persons. How would you cut the cake? Refer to the diagram and show how you would cut the cake on the diagram. 	*Let pupils draw what they would perform (horizontal, vertical, others)*
5. Make up an application problem that matches the mathematics expression "$45\frac{4}{5} \div 9$".	*Generate a problem that need to divide a fraction by a whole number*

Student Teacher as Assessor of Oneself

The instructor was not the only assessor of Adrian's test. Adrian self-assessed her own test as well. In her journal, Adrian explained what key ideas she meant to assess (Figure 2). For example, next to item number 1 she wrote, "Applications". Then, in number 3, she wanted to see if pupils knew when to multiply and when to divide. Also in her journal but separate from the test paper, she indicated her way of scoring. Her criteria for scoring included both computation and reasoning. Under each problem, she explained how she gave partial credits according to her criteria.

When she was asked to assess mathematics performance of her pupils, she made judgment according to the percentage of pupils who correctly answered a problem. Next, she tallied the scores of her pupils (n=44) to get an idea of how scores were distributed. The modal class was 70% to 80% (13 pupils). There were three pupils who scored 100%, and the lowest score was in the 30% to 40% range. Occasionally, she commented on the beliefs of her pupils, "Pupils lost confidence if the answers are large numbers."

In her reflections, she commented on the whole set of problems, "either the problems were too difficult or the pupils were weak". An item-by-item evaluation followed. For the book problem (no. 1), she evaluated it as a basic problem. She commented, "It was solved by most pupils but, three of them multiplied '5' by the denominator instead of numerator." From her account, a lot of pupils could not solve the area problem (no. 2). She suggested that the formula for finding the area of triangle was difficult. She was pleased with the weight problem (no. 3), because she could find out if her pupils were confused over multiplications and divisions. Next, under number 4, she put, "Pupils were not able to explain how they solved the problem." Finally, she commented and gave examples of pupils-generated problems to item number 5. "I found some of these problems not making any sense. For example, Tom walks from home to the train station for 9 hours, covering a distance of $45\frac{4}{7}$ cm. How many cm were covered in one hour?"

In her re-construction of test, she changed number 2 and number 4 of the five problems she originally posed. Next to number 2, she wrote, "I would simplify the problem by changing the word 'triangle' to 'rectangle', because pupils found it difficult." Then, next to number 4, she put, "Since pupils were not able to explain how they cut the cake into 4 parts on drawing, I need to change the problem but I cannot think of how to change it right now."

Student Teachers as Assessors of Peers.

Four students evaluated Adrian's test. They considered strengths of Adrian's test as: creative, realistic, and included problems with varying degree of difficulty. They suggested to her to include simpler problems.

Collated Thoughts of Other Students

As a whole, the four students showed concern for coverage, complexity, format, and scoring. Below are example illustrations of their concern given in their journals. Interestingly, the investigator found that comments in the other eight student journals, which reported on tests on other topics, also fell into these four categories.

Coverage

"I have in mind covering topics like length, area, capacity, time and writing story problems" (Beverly) and "I purposely included different contexts" (Dawn).

Complexity

"This is a straight forward problem. Therefore, I put it as the first problem in the test" (Connie) and "I should put the most difficult problem as the last problem in the test" (Edward).

Format

"I thought we can ask questions in context and numbering them will make them apparent. However, pupils missed the questions completely. Now I see that it is not clear enough and that the text of information should appear before the actual questions. Also, it is best for one question to appear on one line" (Beverly).

Scoring

"I would interchange the points carried by the first problem on solving (6 points) and the third on posing (8 points). The problem on solving should carry 8 points instead. It is more important than the one on posing without solving. The problem on posing should only carry 6 points" (Connie).

<u>What is important to be tested?</u>

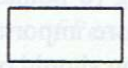
The analyses at problem level, which were completed by the instructor, followed the analyses at test level. The 18 problems of five tests suggested what in multiplication and division are important to be tested, according to student teachers. As a result, these student teachers preferred to ask application problems and problems that were isomorphism of measures in structure. The following two figures depict the counts and empirical examples in various categories. In Figure 3, we can see that all five problem types of problems identified by Butts (1980) were posed by these student teachers. Most of the problems they posed were application problems (10 out of 18) and one of them was Open-search problem. There were 4 problem situations altogether. Interestingly, they were all on problem posing, which was thoroughly discussed during class meetings in the Fall term.

Figure 3. Empirical Problem Type Examples and Counts (N=18).

<u>Recognition</u> (n=2) The answer to 6÷7×3 is the same one of the following. Which is it? a) $\frac{3}{6}\times 7$ b) $\frac{7}{6}\times 3$ c) $\frac{6}{7}\times 3$ d) $\frac{6}{7}\times\frac{1}{3}$
<u>Algorithm</u> (n=1) There are 8÷6×12 kg kiwi and $\frac{5}{3}\times 15$ kg apples. When put on a balance, which type of fruits is heavier?
<u>Application</u> (n=10) A piece of rod is $\frac{1}{4}$ m long. When three pieces were put together, how long is the rod?

Open-search (n=1)

A piece of cake was divided equally among four persons. How would you cut the cake? Refer to the diagram and show how you would cut the cake on the diagram.



Problem Situation (n=4)

Pose a problem to go with the mathematics expression "$1\frac{5}{18} \times 12$"

After analysing the problems according to Butts' classification, the instructor organised these 18 problems according to Vergnaud's multiplicative structures. Figure 4 shows empirical examples of posed problems with signified multiplicative structure. This time, 10 out of 18 were on Isomorphism of Measures and 2 were on rectangular area. There was none on cartesian products nor multiple proportions. The remaining 6 were two items on computation and four on posing problems to match a mathematics expression. Therefore, these 6 did not resemble any multiplicative structure.

Figure 4. Empirical Multiplicative Structure Examples and Counts (N=18).

Isomorphism of Measures (n=10)

Multiplication (n=4)

A book is $3\frac{1}{2}$ cm thick. If 5 are piled together, what is the total thickness?

Quotitive division (n=1)

Several balls weigh $2\frac{2}{3}$ kg. If each weighs $\frac{1}{3}$ kg, how many balls are there?

Partitive division (n=4)

Half a pizza is divided equally among three children. How much did each receive?

Rule-of-three (n=1)

A steel rod of $1\frac{1}{2}$ m long weighs 3 kg. What is the length of a steel rod weighing 1 kg? Weighing 5 kg?

Product of Measures (n=2)
Cartesian product (n=0)
Rectangular area (n=2)
The length of a rectangle is 8 cm and its width $5\frac{3}{4}$ cm. Find the area.
Multiple Proportions (n=0)
Not Applicable (n=6)
Computation (n=2)
$2\frac{2}{5} \times 13$
Problem Posing (n=4)
Write a story to match $5\frac{1}{2} \times 4$.

Implications

In this study, we share an assessment idea in an elementary teacher-training programme. This assessment exercise demonstrates how an instructor in a teachers college can teach and research simultaneously. In fact, it has implications on teachers' development in addition to assessments. Good teachers are able to create meaningful mathematical tasks (NCTM, 1991), and posing problems properly can motivate students to solve problems (Butts, 1980). Also, teachers should take an active role in their own professional development by accepting responsibilities for reflecting on learning and teaching individually and with colleagues (NCTM, 1991, p.168). However, the techniques in posing problems are many. Teacher-training programmes should allow more room for practicing technique, reflecting on practices, and commenting on peers' work. For a beginning teacher, this ongoing progress cannot be assumed without having them actively engage in such activities when they are still in the college. For example, in this study, a student teacher could at least reflect on what she did and put in her journal, "I need to change the problem but I cannot think of how to change it right now."

References

Butts, T. (1980). Posing problems properly. In Stephen Krulik and Robert E. Reys (Eds.), *Problem solving in school mathematics* (pp.23-33). Reston, Virginia: National Council of Teachers of Mathematics.

Castle, K., & Aichele, D. B. (1994). Professional development and teachers autonomy. In Douglas B. Aichele and Arthur F. Coxford (Eds.), *Professional development for teachers of mathematics* (pp.1-8). Reston, Virginia: National Council of Teachers of Mathematics.

Clarke, D. (1988). *Assessment alternatives in mathematics*. Queensland, Australia: Watson Ferguson.

Education Department (1993). *Curriculum standards for national elementary schools in Taiwan* (pp. 91-132).

Kenny, P. A., & Silver, E. A. (1993). Student self-assessment in mathematics. In Norman L. Webb and Arthur F. Coxford (Eds.), *Assessment in the mathematics classroom* (pp. 229-238). Reston, Virginia: National Council of Teachers of Mathematics.

Kulm, G. (1994). *Mathematics assessment: What works in the classroom*. San Francisco, California: Jossey-Bass Inc., Publishers.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, Virginia: Author.

National Council of Teachers of Mathematics (1991). *Professional standards for teaching mathematics*. Reston, Virginia: Author.

National Council of Teachers of Mathematics (1995). *Assessment standards for school mathematics*. Reston, Virginia: Author.

Ridgway, J. (1988). *Assessing mathematical attainment*. Berkshire, England: NFER-NELSON Publishing Company Ltd.

Stenmark, J. K. (1989). *Assessment alternatives in mathematics*. Berkerly, California: EQUALS, University of California.

Van den Heuvel-Panhuizen, M. (1990). Realistic arithmetic/mathematics instruction and tests. In K. Gravemeijer, M. van den Heuvel, and L. Streefland (Eds.), *Contexts, free productions, tests, and geometry in realistic mathematics education.* Utrecht, The Netherlands: CD-b Press, Freudenthal Institute.

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.)., *Acquisition of mathematics concepts and processes* (pp.127-174). New York: Academic Press.