The Effect of Peer-Assessment on the Attitudes of Pre-Service Elementary and Middle School Teachers about Writing and Assessing Mathematics

Cheryl Beaver
Western Oregon University
Mathematics Department
345 N. Monmouth, Ave
Monmouth, OR 97361
beaverc@wou.edu

Scott Beaver
Western Oregon University
Mathematics Department
345 N. Monmouth, Ave
Monmouth, OR 97361
beavers@wou.edu

Abstract

Through a study conducted in a core Foundations of Mathematics course at Western Oregon University, the authors investigate the thesis that peer-grading helps future elementary and middle school teachers improve their own attitudes about writing and assessing mathematics. Study participants were asked to provide scale responses to a series of questions regarding their perceptions of their ability to write mathematics and to assess written mathematics, before and after the course sequence. Students in the experimental group were asked to do a sequence of scored peer grading exercises during the course, while those in the control group were not. Statistically significant positive changes are noted in a variety of the experimental groups' perceptions, in particular the beliefs that writing about mathematics will help them learn the topic better, that peer-assessment can help them increase their own depth of knowledge about the topic, and that they will eventually become good at writing sentences and paragraphs about mathematics. The full results of the study are presented, along with a representative sample of students' open comments. .

Keywords: Peer-assessment, pre-service teachers, writing mathematics

Introduction

The role of assessment in a classroom is more than just a means by which to assign a grade. Assessment has evolved to play a role in student learning as well (Boud, 1990; Dochy & McDowell, 1997; Zevenbergen, 2001). In particular, self- and peer-assessment can help students develop not only critical thinking and analysis skills, but content-related and reflective skills. This study sought to provide evidence for the hypothesis that the use of peer-assessment on written mathematics assignments using a scoring rubric could help improve the attitudes of pre-service teachers toward writing, explaining, and assessing mathematics.

The learning and comprehension of mathematics is, obviously, crucial to prospective mathematics teachers. Effective teaching of mathematics requires the teacher to both broadly and deeply understand not only the material they will be teaching, but also the mathematical concepts that their students will be exposed to in future grades as well as what they have encountered in previous grades (Conference Board of the Mathematical Sciences, 2001, p.7). College level foundational mathematics courses are essential components in the effort to achieve this goal in the pre-service teachers' education. Often, it is in these classes that students solidify their understanding of the material they learned in elementary and middle school and begin to understand the threads that tie the material together between grades.

The ability to communicate effectively is one key to the understanding of mathematics – particularly for pre-service teachers who will later teach and assess the learning of their students. Indeed, Communication is one of the five Process Strands cited in the Principles and Standards for School Mathematics (National Council of Teachers of Mathematics [NCTM], 2000). The standard states

> When students are challenged to communicate the results of their thinking to others orally or in writing, they learn to be clear and convincing. Listening to others' explanations gives students opportunities to develop their own understandings….Students who have opportunities, encouragement, and support for speaking, writing, reading, and listening in mathematics classes reap dual benefits: they communicate to learn mathematics, and they learn to communicate mathematically. (NCTM, 2000, p. 60).

In "Writing to Learn Mathematics," Countryman finds that writing about mathematics helps students to "become aware of what they know and do not know, can and cannot do", "connect their prior knowledge with what they are studying", "summarize their knowledge and give teachers insights into their understanding", "raise questions about new ideas", "reflect on what they know" and "construct mathematics for themselves" (1992, p.7). One avenue for the encouragement of mathematical writing is to require students write about the salient points in their solutions to word problems. Countryman goes on to say that, among other things, word problems in particular give students the opportunity to "focus on their own thinking" and that "conceptions and misconceptions will be revealed as students describe their explorations of a problem" (1992, p. 57). Carefully written out details about the solution to a problem provide a clearer window into a student's understanding than a simply stated solution. Detailed writing is an avenue by which students can learn more deeply about the mathematics than just firming up their superficial understanding of the procedure or algorithm for solving the problem.

Students often come to college inexperienced and ill-prepared to write about mathematics. Many students have a perception of ambiguity regarding what constitutes a mathematically rigorous argument in support of their answer. One way to clarify expectations is to make use of a scoring rubric. Scoring rubrics help students to clearly identify key criteria common to a wide variety of problem solutions. Repeated use of generalized rubrics can help students learn to discern the components of good problem-solving and of well-written mathematical solutions. Such rubrics ultimately play a dual role as both a method of scoring and a learning tool for the classroom, especially if the students learn to use the rubrics themselves in self- or peer-assessment (Arter, 1993; Mehrens, Popham & Ryan, 1998).

At Western Oregon University (WOU) all K-8 pre-service teachers are required to take a sequence of three foundational mathematics classes: Foundations of Elementary Mathematics I, II, & III. Each course is 4 (quarter) credit hours and meets for 5 hours per week. As part of the coursework students are required to solve several "Problem of the Week" (POW) assignments. POWs are problems requiring significantly more analysis and synthesis than typical homework problems. For their solutions to POWs, students are asked to clearly show their process and write sentences and paragraphs supporting and explaining their work. POWs are graded using the Oregon State Official Mathematics Scoring Guide 2000-2008 (hereafter OSMSG - Appendix A).

The guide is a generalized rubric that was used in the state of Oregon to score written mathematics works from 2000-2010.

The authors' experience with the POW assignments prior to the study discussed here showed that many students had a negative attitude toward the requirement that they write extensively about their mathematical solutions. The decision to assign the POWs was made in the hope that students would be compelled to think more deeply about the mathematics and improve their communication and justification skills, but the results were poorly written solutions that often made little sense and bore all the hallmarks of work completed solely for the sake of completion. Further, even with the aid of a scoring rubric, many students were still unsure what constituted a good quality written response. Based on comments about the POWs from students, we felt that their attitude about writing and their perception of its usefulness in mathematics (or lack thereof) played a role in the poor work that many students submitted. White, Way, Perry and Southwell (2005) in "Mathematical Attitudes, Beliefs and Achievement in Primary Pre-service Mathematics Teacher Education" state that "The experiences and beliefs of pre-service teachers influence the formation of attitudes and these, in turn, influence their classroom practices and beliefs." It is important to convince students that writing about mathematics can increase learning and understanding of mathematics so they themselves can become better communicators of mathematics and pass that skill on to their own students.

We suspected that the poor outcomes and attitudes about the POW assignments didn't necessarily reflect a lack of understanding of the particular problems, but that students' negative attitudes about the value of writing mathematics was preventing them from performing at their best when trying to communicate that knowledge. As such, the deeper learning that can arise from writing assignments was rendered inaccessible. Peer-assessment can be a tool for increasing student performance (Sluijsmans, 2002). Our belief was that if students could read each others' written solutions and take part in the assessment process they would experience firsthand the lack of clarity and understanding demonstrated in their peers' mathematical writings. They may then in turn understand that this lack of clarity would be reflective of that in their own writing. Further, by repeatedly using the scoring rubric themselves they would better understand the essential components of writing a clear solution. We also hoped that this would cause them to take more care in their own mathematical writing, particularly given that some of their peers would be reading their work. We further believed that as students became better communicators in mathematics they would then change their attitude and beliefs about the utility of writing mathematics. Many of these hypotheses about what we believed are difficult or impossible to measure directly. Hence our study sought only to measure whether the students' self-reported attitudes about mathematics writing, learning, and assessment were changed if they participated in peer-grading experiences. More specifically, we sought to answer the following research questions:

- Can peer-grading exercises help allay pre-service teachers' concerns about teaching mathematics to elementary or middle students?
- Can peer-grading exercises help pre-service teachers improve their perceived skills at or enjoyment of writing about mathematics?
- Can peer-grading exercises help pre-service teachers improve their perceived comfort level regarding their current or eventual ability to assess and critique written mathematical work done by
    - their peers?
    - elementary or middle school students?

- Can peer-grading exercises improve pre-service teachers' perceptions of their understanding of the mathematical topic at hand?
- Can peer-grading exercises improve pre-service teachers' confidence in their current mathematical ability and in their confidence in their ability to improve their mathematical skills in the future?
- Can peer-grading exercises improve pre-service teachers' confidence that mathematical assignments can be fairly graded?
- Can peer-grading exercises improve pre-service teachers' confidence in the proposition that by grading mathematics assignments completed by peers, or by elementary/middle school students, they can improve their mathematical abilities?
- Can peer-grading exercises improve pre-service teachers' confidence in their ability to explain, in writing or verbally, a mathematical topic with which they are already comfortable?

Participants

The participants in the study were 58 pre-service teachers who completed one of four sections of the Foundations of Elementary Mathematics I course at WOU. All participants completed the same four POWs. Participants were divided into experimental group participants (n_1=30) and control group participants (n_2=28). To be consistent in assessment within a given section, and to avoid the organizational problems associated with mixing grading and non-grading students in a given section led us to choose the experimental groups to be entire sections of classes, rather than randomly assigned students across the sections. The demographic makeup across sections was largely uniform in terms of student ages, academic background, concerns about teaching mathematics (as evidenced by questions 1 and 2 below) and gender. The courses were taught by the authors – each taught one experimental group and one author taught two control groups. The structure of the courses and assignments were the same in all class sections.

Method

After each POW was due, the students' work in the experimental group was collected and two copies of each paper was made with the name removed (we were aware that students might recognize each others' handwriting, but judging by the scores this apparently did not cause a problem). Each student in the experimental group received two different POWs written by their peers (pseudo-randomly selected) and then used the OSMSG as a rubric to assess the POWs. The rubric has five components: conceptual understanding, process and strategies, verification, communication, and accuracy. Experimental group participants were assessed, as part of their course homework grade, on their adherence to OSMSG and the thoughtfulness of comments they wrote on POWs they graded. To ensure at least some level of thoughtfulness, as part of their POW grading duties experimental group participants were required to write comments for each part of the scoring rubric (except accuracy) with at least one positive comment overall and at least one comment indicating a need for improvement in that category (Appendix B). Ultimately each student in the experimental group received two scores: one score for their POW work and one score for their POW grading work. In the case of the score for their work on the POW, the instructor first graded the POW him/herself and then compared that score to the score given by the student peer-grader. In the case of a conflict, the instructor's score had priority and the instructor changed the grade on the student sheet with an explanation for the change – this removed the concern by some of the students that they might not be able to grade correctly or be fairly graded. The scoring grade was based on two components (1) the adherence to the OSMSG when assigning grades and (2) the completeness of the previously mentioned required comments.

The Instrument

The thesis that peer-grading helps future elementary teachers improve their attitudes and perceptions about writing and assessing mathematics was investigated using hypothesis testing on a bank of 17 questions with Likert-type scale responses given to each the experimental and control group. An instrument was developed consisting of an entrance survey with seventeen items and an exit survey containing the same items along with two additional questions for all and three additional follow-up items for experimental group participants only. Participants answered each question using a five-point Likert-type scale (1=strongly disagree; 5 = strongly agree). During the first class meeting we distributed the entrance survey to participants. The exit survey was distributed during the last class period of the term.

The three additional items in the exit survey for experimental group participants measured each participant's belief in the idea that grading the POWs helped improve (1) their grading/critiquing abilities, (2) their mathematics writing skills, and (3) their knowledge and understanding of the POW topic.

Hypothesis Testing

Since the study was designed to detect improvements in students' attitudes toward their mathematical writing skills, a one-tailed test was chosen. For each item in the surveys we used the Student t-test with unpooled variances to test the null hypothesis that the mean difference between the paired scale responses in the entrance and exit surveys for the experimental group, denoted EMD, equals the mean difference between the paired entrance and exit responses for the control group, denoted CMD, against the alternative hypothesis that EMD was more "favorable" than the CMD (generally this meant EMD > CMD). The p-values are reported for each question in the survey. Generally, a significance level of $\alpha = .10$ is used, but we avoid making strong claims either way for p-values close to that threshold. In some cases the entrance and exit surveys could not be paired (e.g. a student took an entrance survey, but no exit survey or vice versa). These surveys were not included in the analysis or count of the 58 participants.

Results and Interpretations

The questions common to each of the entrance and exit surveys are listed in this section, along with the values of the overall average score in the exit survey by the experimental group for each item (EM) and by the combined control group (CM), the EMD and CMD and the associated p-values. Again, note the p-value is a measurement on the mean *change* in response values between the entrance and exit surveys for participants in each class, not a comparison of the mean values on the exit surveys. Recall that since the surveys were administered before and then after the Foundations of Elementary Mathematics courses, even in the control group one should expect a difference in the scale responses. We sought to determine whether the scale responses differed more in the experimental group.

Concerns about teaching

1. I have concerns about teaching mathematics to elementary school students.

EM exit = 2.36667	CM exit = 2.53571	
EMD = -0.4	CMD = -0.28571	p = .36

2. I have concerns about teaching mathematics to middle school students.
EM exit = 3.13333 CM exit = 3.18519
EMD = -0.26667 CMD = -0.03704 p = .25

Since items 1 and 2 are phrased negatively, both EMD and CMD are negative indicating that students' concerns about teaching mathematics to these students went down after the course. The difference between EMD and CMD in these two questions is not statistically significant. Thus, one limit on the efficacy of the peer-grading exercise is that it apparently has little or no effect on pre-service teachers' concerns regarding the prospect of teaching mathematics to elementary or middle school students. This was not too surprising; the anxiety which may arise from knowing that eventually one may find oneself explaining mathematics to a group of students is not likely to be easily allayed by an exercise in grading. Rather, we believe that the completion of the course itself was likely responsible for the marginal drop in concerns.

Writing about mathematics

3. I enjoy writing sentences and paragraphs about mathematical concepts.
EM exit = 2.96667 CM exit = 3.07143
EMD = 0.23333 CMD = 0.53571 p = .15

4. I am good at writing sentences and paragraphs about mathematical concepts.
EM exit = 3.80000 CM exit = 3.21429
EMD = 0.76667 CMD = 0.5 p = .12

5. I believe I will eventually become good at writing sentences and paragraphs about mathematical concepts.
EM exit = 4.30000 CM exit = 3.96429
EMD = 0.6 CMD = 0.0 p = .011

One might hypothesize that the high p-value of .15 for item 3 is a result of many students' deeply-ingrained bias toward the idea that mathematics is not enjoyable. Though the p-value for item 4 was .12, a p-value so close to the threshold offers some likelihood that several pre-service teachers in the experimental group felt better about their mathematical writing abilities after the exercises. We conclude from item 5 is that it is overwhelmingly likely that the peer-grading exercises give students hope that they will acquire sufficient skill in mathematical writing as their education proceeds.

Assessing written assignments

6. I now feel comfortable grading/critiquing written math assignments from elementary or middle school children.
EM exit = 3.88333 CM exit = 3.25000
EMD = 1.13333 CMD = 0.464286 p = .012

7. I believe I will eventually become comfortable grading/critiquing written math assignments from elementary or middle school children.
EM exit = 4.50000 CM exit = 4.17857
EMD = 0.36667 CMD = -0.01786 p = .093

8. I feel comfortable grading/critiquing written math assignments from peers.

EM exit = 4.06667 CM exit = 3.17587

EMD = 0.9 CMD = 0.5 p = .053

9. I believe I will eventually become comfortable grading/critiquing written math assignments from peers.

EM exit = 4.43333 CM exit = 3.92857

EMD = 0.53333 CMD = 0.07143 p = .044

We expected that the peer-grading exercises would help make students more comfortable with the idea of grading written mathematics assignments submitted by elementary or middle school students, and the p-value in item 6 confirmed this. A p-value of .093 for item 7 is not surprising since it regards students' future expectations, rather than their current perceptions. The p-value for item 8 indicates that a strong likelihood that the grading exercises helped make experimental group participants more comfortable with assessing each others' assignments, and the lower p-value from item 9 seems to show that this effect should be attainable even to those who were not yet comfortable with such assessment at the study's conclusion.

Writing for understanding

10. I believe that writing sentences and paragraphs about a mathematical concept will help me understand the topic better.

EM exit = 4.13333 CM exit = 4.10714

EMD = 0.38333 CMD = -0.21429 p = .024

The results from item 10 show that pre-service teachers can expect improved understanding of mathematical topics if they write paragraphs and sentences; however, the attitudes of those in the experimental group went up by a statistically significant margin. We note that although the average score in the exit survey for this item is similar for both groups, for the control group this actually represents a decrease in the average score from the entrance survey while for the experimental group it indicates an increase in average score.

Mathematical confidence

11. I am confident in my mathematical abilities so far.

EM exit = 3.93333 CM exit = 3.67857

EMD = 0.73333 CMD = 0.375 p = .11

12. I believe that my mathematical abilities can improve.

EM exit = 4.63333 CM exit = 4.60714

EMD = 0.5 CMD = 0.17857 p = .077

Though we cannot conclude that peer-grading exercises help improve students' current confidence in their mathematical abilities, the p-value for item 11 indicates that more study is warranted; this is supported by the result from item 12 which indicates that having completed the peer-grading exercise appears to associate with improved expectations regarding students' math skills. In a future study one might ask experimental group participants a battery of questions designed to detect what skills pre-service teachers might need to improve their mathematical abilities, and whether those needs could be met in exercises like peer-grading.

Fair assessment

13. I believe mathematics assignments are usually graded fairly.
    EM exit = 4.20000 CM exit = 4.08929
    EMD = 0.33333 CMD = 0.51786 p = .51

14. I believe mathematics assignments can be graded fairly.
    EM exit = 4.36667 CM exit = 4.50000
    EMD = 0.33333 CMD = 0.25 p = .37

Based on the results from items 13 and 14, the grading exercises appeared to have no effect on participants' perception of whether or not math assignments are, or can be, graded fairly. However, since the means of the answers on the exit survey were in all cases higher than on the preliminary survey, we might conjecture that the general experience with the grading rubric itself on the POWs had a positive effect on the student's perceptions about mathematical assessment.

Learning mathematics through assessment

15. I believe that by grading mathematics assignments of elementary/middle school students I can improve my math skills.
    EM exit = 4.24138 CM exit = 4.28571
    EMD = 0.31035 CMD = -0.03571 p = .098

16. I believe that by grading mathematics assignments of my peers I can improve my math skills.
    EM exit = 4.41379 CM exit = 4.35714
    EMD = 0.48276 CMD = -0.03571 p = .026

The very low p-value for item 16 suggests that the act of peer-grading did help students believe that reading the work of their peers could help them improve their own mathematical skills. A p-value of .098 for item 15 is below but quite close to the threshold. Though we can conclude that the exercises improved pre-service teachers' expectations for improvement of their own math skills by grading students' assignments, more study may be warranted. The disparity between the p-values for items 15 and 16 might be explained by the fact that experimental group participants actually graded each others' work, but neither group had graded elementary or middle school students' work. It is notable that in all cases the mean scores were high indicating that it is likely that the students believe they can learn mathematics through assessment.

Explaining mathematics

17. Pick a topic in mathematics that you are very comfortable with – for example, adding one to a number.
    a. I am comfortable explaining this topic to someone else verbally.
    EM exit = 4.33333 CM exit = 4.48000
    EMD = 0.62963 CMD = 0.28 p = .11

b. I am comfortable explaining this topic to someone else in writing.

EM exit = 4.20370 CM exit = 4.32000

EMD = 0.79630 CMD = 0.48000 p = .11

The results from item 17 do not permit us to conclude that the exercises improved pre-service teachers' perceptions of their ability to explain, verbally or in writing, a concept with which they were already quite comfortable. But again because the p-value is close to the threshold, we believe that more study is warranted regarding this topic. We actually expected a much higher p-value since we specifically designed the item so that participants could choose a topic they already understood well. In designing this question, we were concerned about the potential effect from including that specific example, but in the end decided to include the example to ensure uniform interpretation of the question for all participants. In future research more care will be taken in the design of this type of question, perhaps offering a broad choice of topics and categorizing responses by the topic chosen.

Mathematical writing and learning through POWs

Problems 18 and 19 were not included in the pre-survey because students had not yet completed any POW assignments, so we report only the experimental mean (EM) and control mean (CM) from the post-survey. In this case we base the p-value on the differences in the mean scores rather than the difference in the mean of the change in paired scores.

18. I believe that the Problems of the Week (POWs) helped improve my skills in writing about mathematics.

EM = 4.16667 CM = 3.71429 p = .048

19. I believe that the POWs helped me improve my knowledge and understanding of the topic of the POW.

EM = 4.1 CM = 3.71429 p = .094

Since the purpose of the POWs was to help students deepen their understanding and improve their writing skills we were pleased that in all cases the mean value was above the midpoint of the scale. The p-value in both cases was below our threshold which further suggests that reading and assessing the work of their peers deepened their realization that writing in depth about mathematics problems helps improve both one's mathematics writing ability and content knowledge.

Peer-Assessment as a tool for learning

We did ask questions specific to the experimental group, with the purpose of determining the strength of the participants' beliefs that peer-grading POWs improved grading ability, mathematical writing skills, and understanding of the topic of the POW itself.

20. I believe that grading POWs helped me improve my grading/critiquing abilities.

Average = 4.266666666

21. I believe that grading POWs helped me improve my own skills in writing about mathematics.

Average = 4.083333333

22. I believe that grading POWs helped improve my knowledge and understanding of the topic of the POW.

Average = 4.150

The fact that the averages were above 4 on a 5-point scale indicated to us that the peer-grading played a meaningful role in the class and in the students' perception about the value of the peer-grading in their education of assessing, writing, and learning mathematics.

Comments from Students

After each grouping of questions on the survey, students were asked for open-ended comments. In both the experimental and control group there were students who felt the POWs were useful and those who disliked them. As we had hoped, of those who made comments on the post survey in the experimental group, overwhelmingly more students made positive comments than negative. In the experimental group, 16 students wrote (what we felt were) positive comments about the POWs, while only 2 made negative comments. In the control group, 10 students made positive comments and 10 students wrote negative comments.

In addition to being asked for comments after each group of questions, in the exit survey students were asked: "Please comment in general about what you liked/disliked about the POWs and their purpose." A sample of the comments are given below; the EG or CG after the comment indicates that the student was in the experimental group or the control group, respectively.

Positive comments

"POWs gave me no other choice then [sic] to understand what I was doing. I had to understand it <u>and then</u> be able to explain it." EG

"I <u>didn't</u> like how hard they were. I <u>didn't</u> like how stressful they were. I <u>did</u> like how they made us think. I <u>did</u> [like] how they taught us to explain somewhat difficult concepts, & then how to do them. It helped me better understand." EG

"The POWs were hard, but they honestly were what taught me about the section we were discussing" EG

"POWs were a pain at times, but they definitely forced me to explain the topic more thoroughly. This improved my knowledge and understanding of all the POW topics." CG

Negative comments

"I didn't understand most of the POW's and I feel they weren't very useful for me." EG

"I really struggled with several of these POWs and I didn't feel like they helped me improve my math skills. It would also take me hours to get one done." CG

"I do not think that the POWs were any good at all because they just confused me a lot more."CG

Comments about peer-assessing

Students in the peer-grading group were asked to comment on the grading experience. Some of their comments follow:

"Grading my peers POWs I was able to see the same problem I did in different ways that were all right."

"There are some that I wish I could hand back & tell them to submit a completed assignment. I think that's worse than not doing it at all."

"It always helps me to see how others go about the same problems I attempt to solve. When I graded my peer's assignments, I learned many things and noted what I could have done to improve my score/understandings."

Overall, the comments from the experimental group seemed to indicate that the biggest gain students saw in the peer-grading process was an increased awareness in the variety of ways to solve problems. Others did notice, as we had hoped, that their peers' solutions lacked clarity and were sometimes difficult to score. Whether this awareness helped them to improve the clarity in their own writing we did not analyze.

Conclusions

This study utilized the act of peer-grading word problem written solutions using a scoring guide as a tool for analyzing its effect on pre-service teachers' perceptions about understanding about learning, writing, and assessing mathematics. There were statistically significant positive changes in the experimental groups' perceptions that: writing about mathematics will help one learn the topic better; that peer-assessment can help one increase one's own depth of knowledge about the topic; and that they will be able to assess their future students well. The experimental group had the most significant increase over the control group regarding the perception that they will eventually become good at writing sentences and paragraphs about mathematics. We believe that this may be due to the fact that students recognized that they were learning through the process and gained some confidence in their ability to further their own understanding. In our opinions, the study provides evidence, some strong, in favor of the hypothesis that peer-assessment can be a tool in learning and softening attitudes of pre-service teachers about the utility of writing mathematics. Improving the perceptions of pre-service teachers will increase the likelihood that they will pass on to their own students a more positive attitude about writing to learn mathematics.

The Common Core State Standards (CCSS) that have recently been adopted by most states has as three of the Standards for Mathematical Processes: *Make sense of problems and persevere in solving them, Construct viable arguments and critique the reasoning of others* and *Model with mathematics.* (CCSS, 2010, p.6). Teachers will be expected to skillfully guide their students to learn these practices. Pre-service teachers must have as part of their education a familiarity with a variety of paths that can be taken to solve a given problem, they need to anticipate the directions in which a mathematical discussion might go, and be able to lead a student to develop and refine their own thoughts into a correct answer or to correct erroneous thought paths. Solving word problems and writing about mathematics through POWs can contribute to this goal. In addition, this study indicates that the process of peer-assessment can also be useful in widening the toolbox from which students model problems, giving them experience in recognizing viable arguments, and helping them to gain confidence in critiquing the reasoning of others.

References

Arter, J. (1993). *Designing scoring rubrics for performance assessments: The heart of the matter.* Paper presented at the annual meeting of the American Educational Research Association. Atlanta.

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education, 15* (1), 101-111.

Conference Board of the Mathematical Sciences. (2001). *The Mathematical Education of Teachers*. Providence RI and Washington DC: American Mathematical Society and Mathematical Association of America.

Common Core State Standards for Mathematics (2010): http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Countryman, J. (1992). *Writing to learn mathematics: Strategies that work.* Portsmouth, NH: Heinemann.

Dochy, Filip, & McDowell, Liz. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, *23* (4), 279-298.

Mehrens, W. A., Popham, J. W. & Ryan, J.M. (1998). How to prepare students for performance assessments. *Educational Measurement: Issues and Practice*, *3* (3), 18-22.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: The National Council of Teachers of Mathematics, Inc.

Sluijsmans, D. (2002). Establishing learning effects with integrated peer assessment tasks. *The Higher Education Academy*. Retrieved from http://www.palatine.ac.uk/files/930.pdf.

White, A., Way, J., Perry, B. & Southwell, B. (2005/2006) Mathematical attitudes, beliefs and achievement in primary pre-service mathematics teacher education. *Mathematics Teacher Education and Development*, *7*, 33-52.

Zevenbergen, Robyn. (2001). Peer assessment of student constructed posters: assessment alternatives in preservice mathematics education. *Journal of Mathematics Teacher Education 4*, 95-113.

Appendix A
Scoring Rubric

2000 - 2008 Mathematics Problem Solving Official Scoring Guide 2000 – 2008

	CONCEPTUAL UNDERSTANDING *Interpreting the concepts of the task and translating them into mathematics* WHAT?	PROCESSES & STRATEGIES *Choosing strategies that can work, and then carrying out the strategies chosen* HOW?	VERIFICATION *In addition to solving the task, identifiable evidence of a second look at the concepts/ strategies/ calculations to defend a solution* DEFEND!	COMMUNICATION *Using pictures, symbols, and/or vocabulary to convey the path toward the identified solution* THE CONNECTING PATH!
6	The translation of the task is enhanced through connections and/or extensions to other mathematical ideas	Elegant, complex and/or enhanced mathematical processes / strategies used to solve the task are completed	The review is related to the task, and enhanced, possibly by using a different perspective as the defense	The connecting path is enhanced (e.g., graphics, examples) allowing the reader to move easily and make connections from one thought to another
5	The translation of the task into mathematical concepts is thoroughly developed	Pictures, models, diagrams, and/or symbols used to solve the task are thoroughly developed	The review is a thoroughly developed look at the concepts/ strategies/ calculations in relation to the task	The path connecting concepts, strategies, and/or verification toward the identified solution is thoroughly developed
4	The translation of the task into adequate mathematical concepts using relevant information is completed	Pictures, models, diagrams, and/or symbols used to solve the task are complete	The review is completed (concepts/ strategies/calculations), and supports a solution	The path connecting concepts, strategies and/or verification toward the identified solution is complete
3	The translation of the major concepts of the task is partially completed and/or partially displayed	Pictures, models, diagrams, and/or symbols used to solve the task may be only partially useful and/or partially recorded	The review is partially completed, partially recorded, and/or partially effective	The path connecting concepts, strategies and/or verification toward the solution is partially complete, and/or partially displayed with significant gaps that have to be inferred
2	The translation of the task is underdeveloped or sketchy	Pictures, models, diagrams, and/or symbols used to solve the task are underdeveloped or sketchy	The review is underdeveloped or sketchy (e.g., focusing only on its reasonableness)	The path connecting concepts, strategies and/or verification toward a solution is underdeveloped or sketchy
1	The translation of the task uses inappropriate concepts or is minimal or not evident	Pictures, models, diagrams, and/or symbols used to solve the task are ineffective, minimal, not evident, or may conflict with their solution	The review is ineffective, minimal, inappropriate and/or not evident	The path connecting concepts, strategies and/or verification toward a solution is ineffective, minimal or not evident

Accuracy:

5) The answer given is mathematically justifiable and supported by the work.	4) The answer given is adequate or it may contain a minor error, but no additional instruction in the key concepts appears necessary.		1) The answer given is incorrect, incomplete or correct but conflicts with the work.

For use during the 00 - 08 statewide assessments Office of Assessment and Evaluation
Oregon Department of Education **Wording may be refined based on samples of student work** 2004 - 2008

Appendix B
Peer Grading Comments Sheet

GRADER: To get full credit for grading, you must fill out the OSMSG AND write comments in every section (Conceptual Understanding; Process & Strategies, Verification, Communication). Further, at least one of the comments must be under the "I liked this…" subsection and at least one under "Something you could have done better" subsection.

Conceptual Understanding
I liked this about your POW:____
____
____
Something you could have done better:__________________________________
____
____

Process & Strategies
I liked this about your POW:____
____
____
Something you could have done better:__________________________________
____
____

Verification
I liked this about your POW:____
____
____
Something you could have done better:__________________________________
____
____

Communication
I liked this about your POW:____
____
____
Something you could have done better:________________________________
____
____