Beyond the Right Answer: Exploring How Preservice Elementary Teachers Evaluate Student-Generated Algorithms

Tracie McLemore Salinas

Tasks regularly completed by elementary teachers reveal the mathematical nature of their work. However, preservice teachers demonstrate a lack of depth of mathematical thought. This study investigated the criteria preservice teachers intuitively used to evaluate algorithms. The intent was to use that knowledge as a foundation for modeling mathematical habits of mind for similar tasks. Journal writings and notes from in-class discussions were collected over three semesters of an introductory course for future teachers. Data were analyzed to discover dominant criteria used by preservice teachers to evaluate student algorithms. Four criteria, namely efficiency, generalizability, mathematical validity, and permissibility, were routinely used by preservice teachers.

Introduction

Paper-and-pencil algorithms are important tools that equip students for computational fluency. Before algorithms become mechanical procedures for students, their use should involve conceptual knowledge as well as procedural skill (Ashlock, 2006). Hence, teachers are encouraged to allow students to explore and create algorithms before the traditional algorithms are introduced. In Principles and Standards for School Mathematics, the National Council of Teachers of Mathematics (2000) suggested that “when students compute with strategies they invent or choose because they are meaningful, their learning tends to be robust - they are able to remember and apply their knowledge” (p. 86). Reasoning and justification are both inherent in the invention of procedures (Kilpatrick, Swafford, & Findell, 2001). Thus, children use and reveal their own construct of understanding with the procedures that they create (Baek, 1998).

The importance of providing time for exploration and creation of non-traditional algorithms is emphasized in mathematics methods texts for preservice teachers (e.g., Cathecart, Pothier, Vance, & Bezuk, 2006; Van de Walle, 2004). Preservice teachers are encouraged to allow students to generate their own notation and algorithms for strategies they create after exploring with manipulatives. The understanding revealed by students should then be channeled into a logical path to the traditional algorithms, resulting in conceptual and procedural understanding. This sounds deceptively simple to many preservice elementary teachers; however, the mathematical and pedagogical understanding required for this navigation is actually quite rich.

In a discussion on elementary teacher preparation, the Conference Board of the Mathematical Sciences (CBMS; Conference Board of the Mathematical Sciences, 2001) presents a vignette of third-grade students investigating a variety of strategies for multiplying. The vignette supports the CBMS depiction of elementary mathematics as an intellectually rich and challenging field, requiring the development of new mathematical habits, strong connections among mathematical topics, and skills in mathematical justification. In the vignette, the teacher elicits methods of solving a word problem requiring multiplication from five students. The teacher then mines the responses for evidence of correct thinking, but she also must recognize and investigate sources of error, determining which elements to use as the foundation for additional exploration.

The depth of mathematical understanding required to evaluate students’ methods of performing operations is evidence of the mathematical nature of elementary teaching. The implications of the teacher’s understanding of student-generated methods are so profound that Ball, Bass, and Hill (2004) suggest “no pedagogical decision can be made prior to asking and answering this question . . . ‘What . . . is the method, and will it work for all cases?’” (p. 7). Unfortunately, research has demonstrated that teachers exposed to student-generated algorithms routinely look to procedural steps, not to reasoning, in evaluating the correctness of student work (Ball, 1998; Simon, 1993).

Campbell, Rowan, and Suarez (1998) proposed three criteria for evaluating student-invented algorithms: 1) efficient procedures, 2) mathematically valid procedures, and 3) generalizable procedures.
Furthermore, Kilpatrick, Swafford, and Findell (2001) suggested four characteristics of algorithms including: 1) transparency, 2) efficiency, 3) generality, and 4) precision (2001). These lists share a tendency toward the habits of mind demonstrated by mathematicians. Preservice teachers, however, often demonstrate a lack of what Seaman and Szydlik (2007) refer to as “mathematical sophistication”. In other words, preservice teachers tend to have an orientation so far removed from that of the mathematical community as to seriously inhibit their mathematical understanding. Finding similar underlying structures, valuing sense making, and using counterexamples are all practices that serve a teacher well in evaluating student-generated algorithms and are in keeping with the habits of mathematicians.

It makes sense, then, that if preservice teachers are to engage successfully in the mathematical tasks of elementary teaching, their preparation should support the development of appropriate mathematical habits of mind (Borko, Eisenhart, Brown, Underhill, Jones, & Agard, 1992) and perhaps an introduction to the community of mathematicians (Seaman et al., 2007). Moving students towards new mathematical habits of mind requires knowing where they begin the journey (CBMS, 2001). Mathematics educators assume that at that origin lie some appropriate mathematical and pedagogical inclinations on which teachers should be encouraged to act (Ball, 1998).

**Algorithms in Teacher Preparation**

As a mathematics teacher educator, I regularly incorporate activities that simulate teacher tasks into my courses. One of these tasks is evaluating student-generated algorithms. Preservice teachers in my methods and content courses often seem astounded that alternate algorithms exist or that students might create or adapt their own. Typical textbooks for these courses usually demonstrate alternate algorithms, but often they only tell about the procedures rather than encouraging the reader’s discovery of them. Because the algorithms presented tend to be correct, preservice teachers have little opportunity for actually evaluating correctness and identifying sources of error.

As a result, I supplement our texts by demonstrating student-generated algorithms that I have collected through years of working with teachers and student teachers. I challenge my preservice teachers to explore, justify, and counter algorithms through journal writings and classroom discussions. Through their journal writings, I often see evidence of relatively consistent thinking among students. Consequently, I used this information to investigate the thinking preservice teachers bring to the task of evaluating algorithms.

**Methodology**

This study investigated the knowledge that preservice teachers bring to an initial mathematics course for future elementary teachers with regard to evaluating algorithms. Two research questions guided the study:

1) What criteria do preservice teachers tend to use to evaluate algorithms?
2) Do the criteria that preservice teachers rely on vary depending on the nature of the algorithm?

**Setting**

Preservice teachers participating in this study were enrolled in an undergraduate course for future elementary teachers. For many, the course constituted one of the first education-related courses and the first mathematics education-related course of their programs. Content covered in the course included measurement, data analysis and probability, and algebraic thinking. At this university, preservice teachers do not complete a field experience during the course but are engaged in activities similar to teachers’ professional development experiences. Thus, they often attend workshops or conferences; some choose to also observe K-8 classrooms or to interview K-8 teachers. Although the course is taught within the Department of Mathematical Sciences, it is a combination of content and methods. Students use manipulative kits regularly and are expected to justify their work and to use multiple representations in classroom presentations and discussions.

**Participants**

Data were collected over three semesters, allowing 61 students to participate of 69 enrolled in the course. Most students were in the first two years of their undergraduate programs, but about a third of the students were in their third year. Seven of the 61 participants were male.

**Data Collection**

I relied on journal-writing assignments as the primary source of data about student thinking. Students completed a variety of journal writings, including responses to problems in class, reflections on activities, and investigations of teacher tasks. All students were required to complete each writing assignment. In order to explore preservice teachers’ evaluation of student-generated algorithms, I also made notes following in-
class discussions of algorithms from journal assignments and other algorithms mentioned in class.

For the journal assignments that focused on algorithms, I provided a real example of a student-generated algorithm introduced either through explanation or video clip. Preservice teachers then spent some time talking in small groups about the algorithm before beginning to write explanations in their journals. Journal writings were to address the algorithm observed, and preservice teachers were to consider and justify whether they would allow a student to use that algorithm in class. I collected, reviewed, and responded to the writings then provided participants at least one opportunity to edit their responses and resubmit. Eventually, I brought the algorithm back into the classroom for a group discussion. Some of the algorithms provided worked consistently and would be considered mathematically correct; others had only one or neither of these two characteristics.

It is important to note that for each algorithm, initial journal responses were collected and reviewed prior to any instruction in the topic. The introduction of the algorithm into the classroom for group discussion was intended to coincide with the appropriate instruction or to enrich the concurrent instruction. I encouraged students to generate the discussion as much as possible, even if our work took pieces of several class periods. As much as possible, I let the class as a group determine if the procedure was correct, why it was or was not, and whether it was a method that participants would allow in their own classrooms. This process might include a week of individual journal writings and responses followed by portions of three or four class sessions in which the preservice teachers attempted to present their own thinking to their peers. The preservice teachers could use chalkboards and manipulatives to demonstrate their thinking through examples or counterexamples.

**Data Analysis**

After I collected the journal writings and notes, I organized student responses for each sample algorithm in two ways, first by mathematical content and then again by the nature of the responses in spite of mathematical content. This allowed me to consider how the content influenced the students’ responses and to look for similarities in their approaches to the algorithms in spite of content. I reviewed journal writings and my notes from class discussions to determine what criteria preservice teachers seemed to be using in their evaluation processes. I recorded the criteria and grouped them by similarity. As the clusters of similar criteria developed, I created a name for each group that best represented the criteria. For example, as preservice teachers wrote or explained that a student-generated algorithm included “an extra step,” “too many steps,” or a step that “complicates the process,” I grouped these together and considered them to support the criterion of efficiency. I attempted to produce a dominant list that was exhaustive but mutually-exclusive of student responses. I then checked each criterion to see with what frequency it appeared and in which formats – journals and/or in-class discussions. All resulting criteria explored in this paper appeared in a minimum of two-thirds of journal responses and were raised in at least half of in-class discussions. This ensured that the criteria discussed were those used most by preservice teachers.

**Two Examples of Algorithms**

Although a number of algorithms were explored in each semester, I found that two journal responses in particular demonstrated the teachers’ thinking. The two algorithms, one on multi-digit multiplication and the other on division of fractions, are presented here.

An example of a simpler algorithm that most preservice teachers understood well was multiplying a three digit number by a three digit number in which the only non-zero digit is in the hundreds place. For example, preservice teachers were provided 287 x 400. Using a traditional method, one might proceed as shown in Figure 1. However, a student in a local elementary classroom had discovered that he could obtain the same product by performing the multiplication in two steps, as shown in Figure 2. Of course, the student’s method produces the same result every time and is simply the process of breaking one multiplicand into two factors.

```
287
x 400
000
+ 114800
114800
```

Figure 1: Multiplication Using the Traditional Algorithm

```
287 x 400
287 x 100 = 28700
28700 x 4 = 114800
```

Figure 2: Multiplication Using a Student’s Method

The most troublesome algorithm overall for preservice teachers was an example of division of a
fraction by a fraction. Preservice teachers’ difficulty with fraction concepts is well-documented (Ball, 1990b; Ma, 1999; Zazkis & Campbell, 1996). Consequently, I expected preservice teachers to be somewhat procedure-oriented in their approach, but I was surprised by how strictly procedural their responses were.

This example was originally brought to me by a preservice teacher completing a service-learning requirement. The teacher she was assisting explained to her students that “you never, never, never touch the first fraction.” Instead, she demonstrated that you “flip” the second fraction and multiply across. After several examples, she assigned some exercises to the students. While walking around the classroom, she observed one student who had decided, in spite of her warnings, that he would in fact “touch the first fraction.” He “flipped” the first fraction (dividend) to obtain its reciprocal, multiplied it by the second fraction (divisor), and promptly inverted his solution as shown in Figure 3.

\[
\frac{3}{4} + \frac{1}{2} = \frac{4}{3} \times \frac{1}{6} = \frac{4}{6} = \frac{2}{3}
\]

Figure 3: Division of Fractions Using a Student’s Method

Mathematically, there is the obvious question of whether the student’s method of working the problem will produce the correct answer each time. Replacing the numerals of the problem with variables, it is simple to see that the student’s procedure will result in the correct answer. (See Figure 4).

\[
\frac{a}{b} + \frac{c}{d} = \frac{a}{b} \times \frac{d}{c} = \frac{ad}{bc}
\]

Traditional

\[
\frac{b}{d} + \frac{c}{d} = \frac{b}{d} \times \frac{c}{d} = \frac{bc}{cd} = \frac{ad}{bc}
\]

Student -Created

Figure 4: Comparing the Traditional and Student -Created Algorithms

However, anyone familiar with fractions will quickly realize that \(\frac{bc}{ad}\) will not necessarily equal \(\frac{ad}{bc}\) as the student’s work suggests. Figure 3 is an example of this. Digging more deeply, one may question if these two fractions are not equal, are the other components separated by equal signs in fact equal? In the student’s algorithm, inverting the dividend produces a multiplicative statement that is not equivalent to the initial problem. There are obvious errors in the student’s use of equal signs. The initial division statement cannot possibly be equal to both and . Similarly, the multiplicative statement cannot possibly be equal to both and . However, it is impossible from his work alone to determine how the student may have rationalized the differences in his representation and his conceptual understanding or whether he was aware of any inconsistencies at all. This example algorithm was one that produced a correct solution every time but failed to make mathematical sense.

Results

All student journal responses were first reviewed to find similarities in the criteria preservice teachers seemed to select for evaluating student-generated algorithms. Four primary criteria emerged from the preservice teachers’ writings and discussions: efficiency, generalizability, mathematical validity, and permissibility. On a positive note, three of the four criteria they chose echoed those proposed by the mathematics education community, namely efficiency, generalizability, and mathematical validity. On the other hand, preservice teachers seemed to use these criteria quite superficially. For example, preservice teachers applied the generalizability criterion by simply trying a few examples and not by using a more appropriate approach, such as replacing numbers with variables and continuing the investigation. Additionally, preservice teachers added the fourth criterion, permissibility, which seems to demonstrate their lack of personal authority in evaluating mathematics.

Efficiency

In the minds of the preservice teachers, the inefficiency of student-generated algorithms was less a problem of inelegance and more an opportunity for student misunderstanding. For instance, in algorithms that broke numbers apart into their factors, such as in Figure 2, preservice teachers saw the extra step as a potentially confusing one. An example of a typical preservice teacher explanation is:

Although this would work, I would praise [the] student for being creative and finding that method, but I would then discourage the student from using this method. I [sic] reasoning for this is that it creates an extra step and when the student gets into higher level maths it could become confusing and create extra work to go through.

This response followed the introduction of an algorithm in which the “extra step” was actually a
mathematically correct one. Typically, the responses to mathematically correct algorithms with “extra steps” were not very different from ones with extra steps that were mathematically incorrect. It appears that preservice teachers looked at extra steps in light of efficiency alone, not mathematical validity. For instance, in the fractions example demonstrated in Figure 3, preservice teachers tended to look at the final flip as an extra step. One preservice teacher commented that “It can really make it complicated. What happens when you forget to flip the second time?” In classroom discussions that followed the examination of this algorithm, when I asked students why they flip the first time, only one student in all of the class sections was able to provide an explanation beyond the procedure itself. Most only echoed the teacher who instructed her students that “you never touch the first fraction.”

**Mathematical Validity**

Mathematical validity seemed to be secondary to efficiency in preservice teachers’ minds. In fact, for most students, the determination of mathematical validity involved simply trying to recall a rule that related to a particular step. They relied on procedural types of explorations, not conceptual ones. For most preservice teachers, concern arose when a student did something “without having any process to do it.” As one preservice teacher stated, “ . . . in mathematics every thing is done for a reason. Every thing taught in math has a precise way of doing different concepts.” If the preservice teacher could not recall the reason, the algorithm was considered flawed. For the fraction division example, one preservice teacher explained, “ . . . that is no mathematical rule that gives us the right to flip the end. . . everything you do in math and every step has a rule that was created out of logical reasoning and this doesn’t have a reason.” In all of the course sections considered for this study, only four preservice teachers regularly explored the validity of student-generated algorithms by using models, drawings, or other methods that moved beyond procedure and rules.

**Generalizability**

Interestingly, the four preservice teachers mentioned above were also the most successful at determining the generalizability of student-generated algorithms, although well over half of the preservice teachers did discuss the need to do so. Those preservice teachers who were able to accurately evaluate the algorithms with regard to generalizability were able to explain their thoughts clearly. For example, with the multiplication problem from Figure 1, a number of students discovered that the four hundred was being broken into four groups of one hundred. As one student explained in class:

I like this idea of doing this problem. To me as a teacher this shows that the child knows what he/she is doing. It proves they not only know the answer but that they know exactly how to get it. He/she understands the reasons why you put the two zeros at the beginning of the problem instead of just place holders. The student sees 400 as 4 hundreds instead of just a number with zeros.

**Permissibility**

For many preservice teachers, their search for mathematical validity resulted in little more than a search for permission granted by some rule somewhere. However, there is another level of finding permission that seemed strong enough in student responses to become its own criterion. This second idea of permission results from the preservice teachers beliefs’ about what students are permitted to do mathematically in the classroom. That permission is assumed to be granted by the classroom teacher or even future classroom teachers. Preservice teachers demonstrated the idea of permissibility when they explained their concerns for a child using an alternate method in a future classroom. As one preservice teacher wrote:

I would have no problem with one of my students doing the problem this way. As long as it worked each time, I would encourage new and creative ways to solve problems so that they don’t get used to doing the same methods all the time. . . I would, however, still like them to learn the problems both ways so that they can be prepared for future math classes that might need a more traditional way of solving the problem.

Not all responses were quite so supportive of building a repertoire of traditional and alternative algorithms. Many preservice teachers agreed with their classmate who asked, “What happens when they have a teacher next year who doesn’t let them do it this way?” Others chimed in that it was important to discourage alternate algorithms in favor of the traditional ones for fear that the method might confuse other students, make grading more difficult, or cause the child to be singled out by a future teacher as having misunderstood.

**Summary**

The four criteria – efficiency, mathematical validity, generalizability, and permissibility – were
gleaned from responses for algorithms of several types and for different operations. These included the student-generated algorithms as well as those typically presented in textbooks, such as the lattice algorithm. Separating responses by mathematical content, such as whole number multiplication versus multiplication of fractions, resulted in very few differences in the criteria. Preservice teachers applied the same four primary categories across the algorithms.

A striking difference, however, was in the level of confidence that appeared in the responses. For instance, many more of the responses to algorithms dealing with fractions were met with ambiguous replies that would provide the teacher time to reflect on the student-generated algorithm. Unfortunately, while the preservice teachers explained that this is what they would do in the classroom, very few then proceeded to investigate the algorithm in their journals. Others explained what they might try as they considered the algorithm, such as asking if it would work every time or suggesting that more examples were needed; however, they did not demonstrate the ability to do so. Thus, particularly with algorithms that involved fractions, preservice teachers did not address the mathematical content directly in their journal responses. This suggests that they have less confidence in the mathematical content related to algorithms applied to fractions.

Discussion

The tendency of preservice teachers to consider efficiency, generalizability, and mathematical validity suggests that the foundation exists upon which to model a more mathematically rigorous evaluation of student-generated algorithms. It is likely that underlying preservice teachers’ superficial use of these criteria is their inexperience in the mathematical thinking required to properly investigate student work. Modeling the habits of mind of mathematicians offers a way to counter this (Seaman & Szydlik, 2007). On the other hand, preservice teachers seemed to disregard the need to investigate student work if the solution produced was correct. However, we have seen that there are algorithms that produce the correct solution but have seriously flawed conceptual bases or representations. Encouraging preservice teachers to investigate beyond the final answer is vital to teaching them to link conceptual understanding to the use of algorithms; challenging them with algorithms that are numerically correct but mathematically incorrect may motivate such investigations. From my own experience, preservice teachers enjoyed exploring algorithms with which they were not familiar and in doing so, observed the depth of mathematical thinking that is required of common teacher tasks.

Permissibility, on the other hand, is an idea that we may wish to discourage. While we do not want to encourage children to torment their teachers with additional algorithms that require dozens of steps simply because they “work,” we should want children (and teachers alike) to develop their own sense of self as an authority in mathematical thinking. Investigating mathematical authority, Schoenfeld (1994) observed that most college students “have little idea, much less confidence, that they can serve as arbiters of mathematical correctness, either individually or collectively” (p. 62). It is little wonder then that preservice teachers tend to consider themselves as receivers of mathematical knowledge and understanding rather than consumers, evaluators, and generators of it. Perhaps the idea of permissibility is somehow reflective of the tendency of preservice teachers to require verification by outside authority for mathematical validity rather than relying on their own reasoning. Newborn (1999) investigated the relationship of preservice teachers’ loci of authority and their ability to reflect deeply on mathematics teaching and learning. She found that a setting that promotes inquiry may assist in ushering them from requiring an external locus of authority to becoming confident in their ability to provide authority. Furthermore, preservice teachers should also be made aware of the difference between their idea of “permissibility” in the sense of mathematical “rules” and the deeper notion of mathematical validity. Here again, we see the importance of offering preservice teachers more mathematical ways of thinking.

Preservice teachers who did not successfully determine the generalizability, or lack thereof, for a particular algorithm failed to do so because of one of two errors. They either tried sample problems that were too limited in the types of numbers or they looked too superficially at the written representation of the procedure and not at the underlying thinking. The first of these errors likely results from mathematical inexperience. The second error, however, suggests that preservice teachers saw algorithms less for what they represented and more for what was written on the paper. In investigations that occurred in class, it was apparent that preservice teachers responded to the question of whether an algorithm works by considering initially whether it produced the correct solution. If it did, most preservice teachers stopped the evaluation process, convinced that there was nothing else to consider. This observation fits with the difficulties
Preservice teachers had in evaluating an algorithm that numerically produced the correct solution but was flawed mathematically.

With regard to the differences that appeared when reviewing results based on content, it is not surprising that fractions topics were most challenging for preservice teachers. The difficulties that preservice teachers have with fractions and division-related content is well documented (Ball, 1990; Ma, 1999; Simon, 1993; Tirosh, 2000). Teachers’ self-confidence in their ability to understand student work may play a role in their decision to investigate students’ claims mathematically (Ma, 1999). Thus, while introducing preservice teachers to mathematical habits of mind through activities such as evaluating student work, it is vital to maintain a focus on further deepening content knowledge so that the confidence to carry out that evaluation is also developed.

Exploring alternate algorithms, along with the traditional algorithms, may also offer opportunities for rich discussions in K-12 mathematics classrooms (Baek, 1998; Mingus & Grassl, 1998). If teachers are not equipped with the mathematical knowledge necessary to evaluate student work, they are less able to lead a class through an investigation that is anything but shallow. A substantial mathematical preparation of teachers is necessary so that this opportunity for mathematical investigations is not missed.

Conclusion

Preservice teachers demonstrate an intuitive understanding of the need to check whether an algorithm is generalizable and efficient. Their concern with mathematical validity, though superficial, is at least present. Modeling mathematical thinking in tasks of this nature may assist in developing the habits of mind necessary for successfully completing the mathematical tasks of elementary teaching. In addition, a combination of exposure to the way mathematicians think and experiences with students may lead preservice teachers to discover themselves as sources of mathematical authority, particularly in their own classrooms. This level of independence and competence seems necessary to achieve the kinds of classroom necessary for developing students’ mathematical understanding.

References


