Isotani, S., Adams, D., Mayer, R.E., Durkin, K., Rittle-Johnson, B., & McLaren, B.M. (2011). Can erroneous examples help middle-school students learn decimals? Accepted for full paper presentation at the *Sixth European Conference on Technology Enhanced Learning: Towards Ubiquitous Learning (EC-TEL-2011)*.

Can Erroneous Examples Help Middle-School Students Learn Decimals?

Seiji Isotani[1], Deanne Adams[2], Richard E. Mayer[2], Kelley Durkin[3], Bethany Rittle-Johnson[3], and Bruce M. McLaren[1]

[1] Carnegie Mellon University, U.S.A.
[2] University of California, Santa Barbara, U.S.A.
[3] Vanderbilt University, U.S.A.
{sisotani@icmc.usp.br, deanne.adams@gmail.com, mayer@psych.ucsb.edu, kelley.l.durkin@vanderbilt.edu, bethany.rittle-johnson@vanderbilt.edu, bmclaren@cs.cmu.edu}

Abstract. This paper reports on a study of learning with erroneous examples, mathematical problems presented to students in which one or more of the steps are incorrect. It is hypothesized that such examples can deepen student understanding of mathematics content, yet very few empirical studies have tested this in classrooms settings. In a classroom study, 255 6th, 7th, and 8th graders learned about decimals using a web-based system under one of three conditions – erroneous examples, worked examples, and partially-supported problem solving. Although students' performance improved significantly from pretest to posttest the learning effect for erroneous examples was not better than the other conditions, and unlike some earlier empirical work, the higher prior knowledge students did not benefit more from erroneous examples than from worked examples or problem solving. On the other hand, we were able to identify certain key decimal misconceptions that are held by a high percentage of students, confirming earlier mathematics education studies. Also, the incidence of misconceptions declined over the course of the lesson, especially for the worked example group. Overall, these results could indicate that erroneous examples are simply not as effective for learning as we (and other) researchers hypothesize. The results could also indicate that the manner in which erroneous examples were presented to the students in this study somehow missed the mark in promoting learning. It is also possible that erroneous examples, like some other e-learning techniques, do not work as well in classroom as they do in a laboratory setting. We discuss these possibilities and how we are redesigning the treatments to be more focused and appealing to learners for a subsequent study.

Keywords: erroneous examples, decimal, math education

1 Introduction

A U.S. National Math Panel Report emphasizes the importance of students mastering decimals [1]. Although conceptual understanding of decimals is critical for most of later mathematics, the Panel reports that in general, students receive very poor preparation in decimals. In fact, it is well documented that students often have difficulty understanding and mastering decimals [2, 3, 4]. Indeed, even adults are known to have trouble with decimals [5].

One way to possibly remedy this situation is to present students with an approach that falls outside of the classroom norm: erroneous examples. An erroneous example is a step-by-step description of how to solve a problem in which one or more of the steps are incorrect. Erroneous examples are seldom used in classrooms or empirically investigated as a means for teaching students mathematics. On the other hand, some researchers have argued that confronting students with mathematical errors can be valuable; particularly when students are sufficiently prepared to deal with errors. Further, some empirical research has demonstrated that erroneous examples can particularly benefit higher-prior knowledge learners in mathematics [6, 7].

To evaluate the effects of learning with erroneous examples in the domain of decimals we developed web-based instructional materials that help students learn by reflecting upon and self-explaining errors. We compared erroneous examples with more typical instructional materials, namely worked examples and problem solving, in a classroom study with 255 subjects from 6^{th}, 7^{th} and 8^{th} grades. Our hypothesis was that including erroneous examples would be a better instructional method to help students learn both cognitive and metacognitive skills and the positive effects would be particularly pronounced for higher-prior knowledge students.

The empirical findings of this work provide new insights into the controversial use of erroneous examples. With regard to learning gains, they suggest that erroneous examples can be as effective as, but not necessarily better than, more traditional approaches. However, these findings may also suggest that somehow our designed materials missed the mark in promoting learning.

In the following sections we discuss the materials, the study design and the results. We also present ideas for re-designing our materials for a subsequent study, developing a more interactive, more focused, and less verbose version of our initial erroneous examples materials.

2 Erroneous Examples

An erroneous example (ErrEx) is a step-by-step problem solution in which one or more of the steps are incorrect. Some theory and research in mathematics education has explored the phenomenon of erroneous examples and provides anecdotal evidence that studying errors can help student learning (e.g., [8]). For example, Borasi argues that mathematics education could benefit from the discussion of errors by encouraging critical thinking about mathematical concepts, by providing new problem solving opportunities, and by motivating reflection and inquiry. In an OECD report released in 2001, the highly-publicized TIMSS studies showed that Japanese math students outperform their counterparts in the U.S., as well as the rest of the western world, with a key difference cited that Japanese educators present and discuss incorrect solutions and ask students to locate and correct errors [9]. The critical point is that these theoretical analyses suggest that directly confronting and reflecting upon errors may lead to the eradication of those errors, similar to what has been shown in much of the learning research on misconceptions [10].

On the other hand, another view is that allowing errors to occur, or exposing learners to errors, hurts learning, presumably because errors are then more likely to be retrieved during later problem solving. For instance, B. F. Skinner's experiments, and the studies of other behaviorists, showed that the lack of a "penalty" when undesirable behavior was exhibited led to repetition of this behavior [11]. Many teachers are ambivalent about – or even hostile to – discussing errors in the classroom [12], quite possibly in reaction to the powerful impact that behaviorism has had, not just in academic research but in everyday education as well.

Thus, the question of how – and if – erroneous examples are beneficial to learning is still very much open and controversial. Despite this, there have been few controlled, empirical studies that have studied the effects of ErrExs on learning. For example, in the aforementioned study conducted by Grosse & Renkl [6], with 118 German university students in the domain of probability, it was identified that providing correct and incorrect solutions of problems fostered far transfer performance if learners had high prior knowledge; for learners with low prior knowledge, only correct solutions were beneficial. In another study with 87 subjects from the 3rd and 4th grade, Siegler [13] identified that students who self-explained both correct and erroneous (incorrect) examples learned mathematical equality better than those who explained only correct examples or who explained their own answers. According to Siegler, the learning benefit of explaining comes from strengthening correct knowledge about the domain and weakening incorrect knowledge. In a study with 93 middle school students in the domain of fractions, Tsovaltzi et al. [7] identified significant metacognitive learning gains for weaker (i.e., lower prior knowledge) students who learned from erroneous examples, as well as cognitive and conceptual learning gains for higher-prior knowledge students when additional help was provided.

We are aware of only a few other studies that have investigated the benefits of learning with erroneous examples [14, 15, 16]. Therefore, there is ample opportunity and need to conduct further empirical work to better understand the impact of erroneous examples on learning.

3 Design of Erroneous Examples in the Domain of Decimals

To design erroneous examples that address students' difficulties in learning decimals, we focused on the most common and persistent errors committed and misconceptions held by students. For example, students often treat decimals as if they are whole numbers (e.g. they think 0.25 is greater than 0.7, since 25 is greater than 7). Some students believe that adding a zero on the end of a decimal increases its magnitude, as with whole numbers. These misconceptions interfere with conceptual understanding of decimals and manifest themselves in various ways [4].

Through an extensive mathematics education literature review, covering 42 published papers and extending as far back as 1928 (e.g., [4, 17, 18, 19, 20, 21]), we found that most of the published papers address either a single misconception or a small set of related misconceptions. We evaluated the prior results and organized them into a taxonomy of decimal misconceptions as shown as reported in [22] and shown in Figure 1.

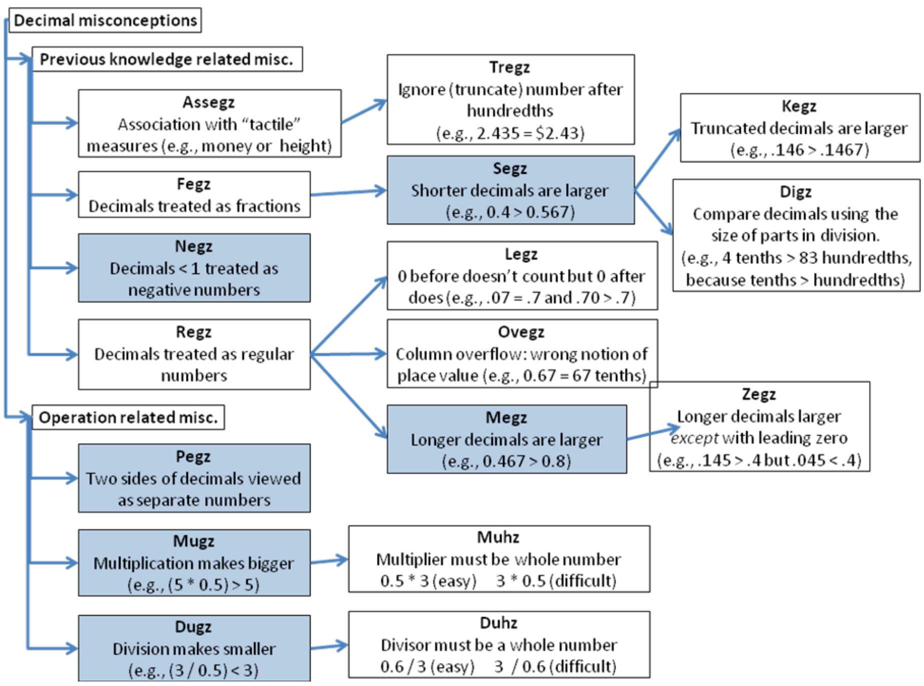

Figure 1. A portion of the decimal misconceptions taxonomy. The gray boxes are the misconceptions included in the study described in this paper.

To categorize and create meaningful clusters of misconceptions, we relied on the quantitative data reported by past researchers. Currently the taxonomy has 45 nodes and classifies the misconceptions according to (a) the prior knowledge the student may have had before learning decimals (e.g., whole numbers, fractions) and (b) the operations associated with the misconception. For example, many students who learn fractions before decimals believe that shorter decimals are larger than longer ones. This misconception occurs because students think that 0.3 is like 1/3 and 0.367 is like 1/367 and, thereby, 0.3 > 0.367. Conversely, students who only have been exposed to whole numbers are more likely to believe that longer decimals are larger (as per the example in the first paragraph of this section). With respect to decimal operations, the taxonomy includes misconceptions like "multiplication makes bigger."

By organizing and relating misconceptions we intended to provide a unique resource, useful both to ourselves as developers of decimal materials, and for other researchers. We also identified key decimal misconceptions that have been reported in different papers as highly frequent among middle school students. Here is a brief description of the seven most common misconceptions:

- **Decimals are like whole numbers** (we call this *regz – Regular numbers misconception*). A fundamental misconception that underlies other misconceptions. Students believe that decimals should be treated as whole numbers and, therefore, they ignore the decimal point [3, 19].

  - Example: 0.07 is equal to 7.

- **Longer decimals are larger** (we call this *megz – Mega numbers misconception*). The most frequently observed misconception. In this case, students think they can compare decimals by counting the digits of each decimal [23].
  - Example: 0.25 is greater than 0.7.
- **Shorter decimals are larger** (we call this *segz – Shorter numbers misconception*). Next to Megz, this is the most frequently observed misconception. Students use knowledge about fractions to evaluate decimals. Thus, when comparing decimals if you increase the number after the decimal point then the decimal will become smaller [4].
  - Example: 0.4 is greater than 0.567 because 0.4 is a fraction like 1/4 and 0.567 is a fraction like 1/567
- **Decimals less than 1 confused as negative** (we call this *negz – Negative numbers misconception*). After megz and segz, this is the next most frequently observed misconception, but its occurrence is far less frequent. Students believe that decimals starting with a zero are smaller than zero (usually students have not learned the concept of negative numbers) [5].
  - Example: 0.25 is less than zero
- **Two sides of decimals viewed as separate numbers** (we call this *pegz – Misconception on either side of the "peg"*). Along with negz, this is the next most frequently observed misconception, the first that involves operations. Students think that the numbers on the left side of the decimal point are completely separate from the right side. Thus, although they can add whole numbers correctly, they won't get the correct answer if they add decimals and need to "carry" a digit to the next column on the left of the decimal point [20].
  - Example: 1.5+3.9 = 4.14
- **Multiplication makes bigger** (we call this *mugz – Multiplication misconception*). One of the most common misconceptions with decimal operations. It builds upon other misconceptions (e.g. regz) since, for instance, the belief that decimals are like integers multiplication should always yield larger results. Students believe that the product of a multiplication should always be larger than the leftmost factor of the operation [24].
  - Example: 5 multiplied by 0.5 is bigger than 5
- **Division makes smaller** (we call this *dugz – Division misconception*). Along with mugz, one of the most common misconceptions with decimal operations. Students believe that the quotient of a division should always be smaller than the leftmost factor of the operation (dividend) [18].
  - Example: 3 divided by 0.5 is smaller than 3

These persistent misconceptions evident in students' decimal knowledge must be addressed so that students can master decimals and move on to more advanced mathematics. The erroneous examples developed in this work focus on 6 of the 7 misconceptions above (megz, segz, negz, pegz, mugz, dugz); the seventh misconception, regz, is encountered throughout all of the others, often the underlying cause of the others.

To give students the opportunity to practice their cognitive and metacognitive skills (vs. practicing procedures) the erroneous examples were designed to (a) prompt students to compare incorrect solutions with correct solutions; (b) guide them in self explaining why a solution is incorrect, and, finally, (c) prompt them to self explain how to correctly solve a problem. Such an approach offers the opportunity for students to be aware of common mistakes and learn from them. We also believe it will encourage critical thinking and motivate reflection and inquiry. This three-step process with erroneous is also consistent with Ohlsson's theory on learning from performance errors [25]. Ohlsson claims that learning from errors requires one to first perceive inappropriate actions or facts while performing a task. Then, one must understand why performed actions/facts are incorrect. And finally, the correct knowledge and skills are employed to fix the problem.

Figure 2 shows an erroneous example focused on the *mugz* (multiplication makes bigger) misconception.

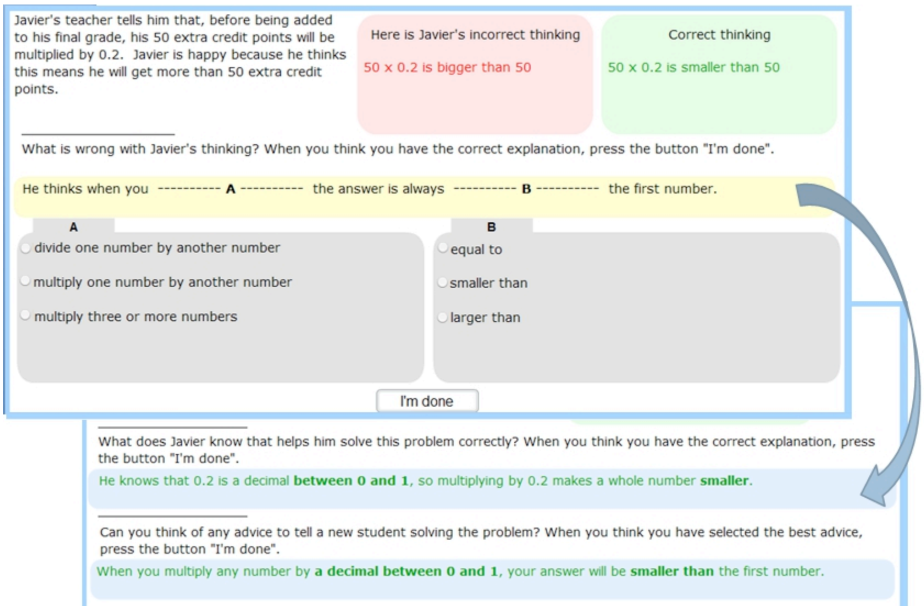

Figure 2. Example of ErrEx for *mugz* (multiplication makes bigger).

Initially students only see the problem statement (top-left) and the incorrect solution (red-box in the top-center). By clicking on a button ("Show me the answer") the correct answer appears in a green box on the top-right of the screen. Students are prompted to compare the correct and incorrect solutions and, using a multiple-choice menu, they next try to self explain the misconception by building a sentence that says why the first solution (in the red-box) is incorrect. If they build the sentence correctly the sentence turns green and a second question appears; otherwise the sentence turns red and the student must change to the correct sentence to continue. The second question asks students to give advice to another student with difficulties. The intention is to motivate the student to practice the correct knowledge in order to fix the erroneous solution (i.e., in the red box). Only after answering both questions correctly can the student proceed to the next problem.

The erroneous examples were developed using an intelligent tutoring system (ITS) authoring tool, known as CTAT [26]. CTAT supports rapid development of web-based learning materials and produces exercises that track and log students' interactions with the software. Thus, it is possible not only to see whether a student's answer is correct or not, but also to check the student's overall behavior in tackling a problem (e.g., how many times they get a particular item wrong; whether they "game the system" by trying all possible answers [27]).

4 Experiment

Due to the previous findings about the effects of erroneous examples, i.e., that erroneous examples can lead to positive learning effects, particularly for higher prior-knowledge learners, we determined that our experiment be guided by the following **hypothesis**: Erroneous examples will help students improve both their cognitive and metacognitive skills. With respect to metacognitive skills, the students will especially benefit by learning to diagnose bugs/errors in problem solving, a type of metacognitive practice that is not readily available, but for which the erroneous examples provide just such an opportunity. In turn, students will refine their own problem solving skills through such practice. Furthermore, prompting students to reflect on potential errors will lead them to more carefully and deeply reflect upon and improve their conceptual understanding and problem solving in the domain of decimals. The positive effects on cognitive and metacognitive skills will be particularly pronounced for higher- prior knowledge students, as they are better prepared than weaker students to handle the more complex and challenging erroneous examples.

4.1 Method

Participants. Two hundred and fifty-five (255) middle school students Pittsburgh, Pennsylvania, U.S., participated; 86 students from 6th grade, 125 students from 7th grade and 44 students from 8th grade. The study materials were used as a complement for normal class work and the eight participating teachers received the students' scores for all tests in addition to two reports summarizing students' and classroom's performance on each item of the tests (pretest, posttest, delayed posttest).

Materials and Procedure. The complete study design is shown in Figure 3.

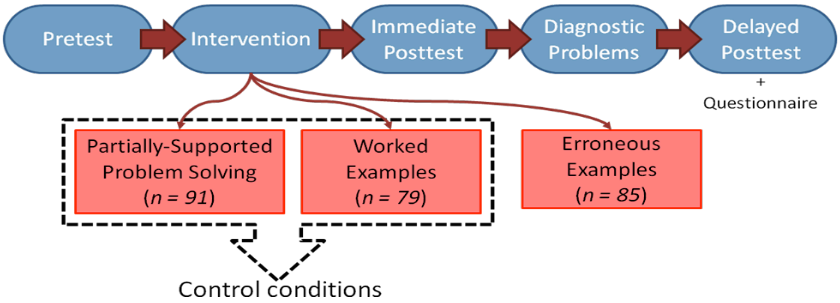

Figure 3. Study design. Overview of the sequence of activities.

All students were randomly assigned to one of three conditions: Partially-Supported Problem Solving (PS condition); Worked Examples (WE condition) and Erroneous Examples (ErrEx condition).

(1) *Partially-Supported Problem Solving (PS)*: In this condition, the student is presented with problems to solve. Incorrect answers turn red and correct answers turn green. After the student gets a problem wrong at least one time, a button is displayed on the screen with the label "Show me the Answer." From this point on the student can select the button to see the correct solution to the problem. (The combination of green/red feedback and the "Show me the Answer" button is why this condition is called "partially-supported" problem solving, rather than pure problem solving, in which no feedback or help would be provided.)

(2) *Worked Examples (WE)*: In this condition, the student was presented with a word problem and a *correct example* to solve it. Similar to Figure 1 the student was prompted to answer two self-explanation statements. However in this case the first and second questions help students to build sentences that explicitly show which concepts were used to solve the problem correctly;

(3) *Erroneous Examples (ErrEx)*: In this condition the student was presented with materials like those in Figure 1. That is, they were presented with a word problem and *both an incorrect and correct example* to solve it. As previously discussed, the student then had to answer two self-explanation questions, one to explain why the incorrect example is wrong and one to provide advice to a fellow student on how to correctly solve the problem;

It is worth noting that the interfaces in each condition were similar and the problem statements were the same. Only the manner in which the problems were presented to the students (e.g. correct solution, incorrect solution, or problem to solve) and the interaction pattern with the interfaces were a bit different, as described above. The PS condition had 91 subjects; the WE condition 79 subjects; and the ErrEx condition 85 subjects. Each condition had 30 treatment problems, 5 problems targeted at each of the 6 misconception types of interest (Table 1 – The *, **, *** are indications of an identical item across conditions). The goal of the first problem of each set of 5 was to familiarize students with the misconception type. In the PS condition this first problem was a partially-supported problem to solve, while in the WE and ErrEx condition it was a worked example with the same problem statement. The rationale behind using a worked example as the first problem in the ErrEx condition was that an ErrEx problem might be too overwhelming and confusing as the first problem. Problems 2 and 5 in each sequence of 5 problems were embedded test problems to check whether a student shows evidence of the particular misconception and whether the intervention helped him/her to overcome it. Finally the problems 3 and 4 were the treatment problems unique to each condition. The problem statements were the same across conditions, however, the way students interact and solve the problems was different across the conditions. For example, the fourth problem for the *megz* misconception cluster was: "*Jorge has 3 cups. The first cup holds 0.47 L, the second holds 0.5 L and the third holds 0.613 L. Jorge's friend asks him to pick the smallest cup. Which cup should Jorge choose?*"; in the PS condition students had a multiple choice menu to select one of the three cups as smallest; in the WE condition students were shown the correct answers (0.47 L cup) and had to build two self-explanation sentences; and in the ErrEx condition students were shown an incorrect answer, related to the misconception type *megz* (0.5 L, because this cup is the shortest among the three), then they could click on a button to see the correct answer, and finally, they had to build two self-explanation sentences.

Table 1. Study design with an overview of each condition. *n* is the number of students per condition.

Condition 1 (*n = 91*): Problem Solving	**Condition 2 (*n = 79*): Worked Examples**	**Condition 3 (*n = 85*): Erroneous Examples**
Megz: "Longer is larger" group of 5 problems		
1: Problem to solve 2: Embedded Test 1** 3: Problem to solve 1 4: Problem to solve 2 5: Embedded Test 2***	1: Worked example* 2: Embedded Test 1** 3: Worked example 1 4: Worked example 2 5: Embedded Test 2***	1: Worked example* 2: Embedded Test 1** 3: Erroneous example 1 4: Erroneous example 2 5: Embedded Test 2***
Segz: "Shorter is larger" group of 5 problems		
Negz: "Decimals seen as Negative Numbers" group of 5 problems		
Pegz: "Two sides of decimal viewed as separate numbers" group of 5 problems		
Mugz: "Multiplication always makes bigger" group of 5 problems		
Dugz: "Division always makes smaller" group of 5 problems		

Test and diagnostic problems were created to target the six misconceptions treated in the study. The pretest, posttest and delayed posttest were three isomorphic and counterbalanced tests (A, B, C) with 42 problems on each test. Thus, one third of the students used version A of the test as the pretest while other students used version B or C as pretest (similarly with the other tests). Each test took 25 to 50 minutes to complete and all of the materials took students between 160 and 220 minutes to complete. Students complete the pretest, intervention, immediate posttest and diagnostic problems in four consecutive periods (42 minutes each). The delayed posttest was applied in class a week after the immediate posttest. Students were also asked to fill a questionnaire that contained demographic questions (e.g. gender and age) and questions about their confidence on working with decimals and questions about their familiarity with computers. The diagnostic problems focused on checking metacognitive understanding which required students to verify whether a solution to a problem is correct or not and explain the concepts and/or processes used to answer the problem correctly/incorrectly.

All of the materials were developed based on problems used in earlier studies reported in the mathematics education literature. In particular, we selected problems that: (a) have been reported as difficult to solve in prior studies; (b) students' answers can lead to common misconceptions (one of the six discussed above); and (c) require more than procedural knowledge to answer correctly. All of the materials were made available through the MathTutor website (*https://mathtutor.web.cmu.edu/*). All participants received private user-IDs and passwords and worked alone at a computer at their own pace. Students were allowed to stop when desired to continue during the next period, except for tests in which all the problems had to be completed in a single period.

5 Results and Discussion

Table 2 shows the mean (and SD) for each of the three groups (and all of the groups combined) on the pretest, immediate posttest, and delayed posttest. To examine whether students' performance improved from undergoing the intervention, t-tests were conducted and revealed that there were significant differences between the scores on the pretest and the immediate posttest, $t(254) = -.7928$, $p < .001$, and between the pretest and the delayed posttest scores, $t(254) = -7.725$, $p < .001$, demonstrating the students' performance improved significantly overall in all conditions.

Table 2. Percentage correct on pretest, posttests, and diagnostic decimal problems.

	Pretest		Immediate Posttest		Delayed Posttest		Diagnostic Problems	
Condition	M	SD	M	SD	M	SD	M	SD
PS	0.5	0.15	0.56	0.16	0.56	0.17	0.57	0.17
WE	0.54	0.17	0.6	0.18	0.6	0.19	0.61	0.2
ERREX	0.53	0.16	0.58	0.18	0.58	0.19	0.59	0.2

To test the general hypothesis that the intervention condition would affect learning outcomes, an ANCOVA was conducted using pretest score as a covariate. No significant differences were found between the three conditions for either the immediate posttest, $F(2,251)=.253$, $MSE=.003$, $p=.777$, or the delayed posttest, $F(2,251)=.178$, $MSE=.002$, $p=.837$. To examine whether condition affected performance on the diagnostic problems, an ANCOVA with pretest as a covariate revealed no significant differences between the three conditions, $F(2,251)= .119$, $MSE = .002$, $p = .888$. Looking at performance on the embedded test problems within the intervention, an ANOVA found no significant differences on overall embedded test performance between the three conditions, $F(2,252)=1.023$, $MSE=138.37$, $p= .36$.

In terms of how long it took the students to complete the intervention, there was a significant difference in the total amount of time if took for the three conditions, $F(2,252)=117.623$, $MSE=72802881.01$, $p<.001$. Tukey post-hoc comparisons of the three condition revealed that the problem solving condition took significantly less time compared to the worked example condition, $p < .001$, and the erroneous examples condition, $p < .001$. However the worked examples condition and erroneous examples did not differ from each other significantly in terms of the amount of time taken to complete the intervention, $p=.805$.

Table 3 shows the mean number of misconceptions (and SD) for each group (and all groups combined) on the pretest, immediate posttest, and delayed posttest.

Table 3. Misconception percentages per condition for pretest, posttest, and delayed posttest

	Problem Solving			Worked Exampled			Erroneous Examples			Overall		
	Pre	Post	Delay	Pre	Post	Delay	Pre	Post	Delay	Pre	Post	Delay
Regz	21%	16%	21%	21%	17%	16%	19%	18%	18%	20%	17%	18%
Megz	36%	28%	28%	34%	25%	21%	33%	25%	24%	34%	26%	25%
Segz	29%	25%	23%	30%	25%	21%	29%	23%	19%	29%	24%	21%
Negz	21%	15%	18%	22%	16%	14%	21%	14%	17%	21%	15%	17%
Pegz	40%	34%	39%	36%	31%	35%	41%	35%	36%	39%	33%	36%
Mugz	37%	32%	31%	35%	32%	30%	34%	31%	30%	35%	32%	31%
Dugz	31%	30%	32%	32%	30%	27%	29%	29%	27%	31%	30%	29%
Overall	**31%**	**26%**	**28%**	**30%**	**26%**	**24%**	**30%**	**26%**	**25%**	**31%**	**26%**	**26%**

Similar to previous research, students displayed many of the misconceptions targeted by this intervention with misconceptions representing 31% of the overall answers given on the pretest, posttest, and delayed postest. For each problem, answers were scored according to whether the answer was correct, incorrect, or incorrect due to a misconception (e.g. On a Negz problem did the student treat the decimal like a negative number?) Examining the difference between the number of misconceptions displayed, there was a significant improvement between the pretest and posttest, $t(254)=8.921$, $p<.001$, and between the pretest and delayed posttest, $t(254)=9.956$, $p<.001$. For the pretest, there were not significant differences between the three conditions for number of misconceptions made, $F(2,252)=1.45$, $MSE=29.17$, $p= .24$ Examining how the three learning conditions affected the overall number of misconceptions displayed there were no significant differences between the three groups for posttest, $F(2,252)=.606$, $MSE=12.71$, $p= .546$. On the delayed posttest, however, there was a significant difference for number of overall misconceptions made, $F(2,252)=3.72$, $MSE=89.72$, $p= .026$. Post hoc Tukey analysis revealed that the worked examples group ($M= 9.62$, $SD= 4.7$) made significantly fewer misconceptions than the problem solving group, ($M= 11.57$, $SD= 5.12$) $p = .028$, but not the erroneous examples group, ($M= 10.09$, $SD= 4.87$) $p= .881$. There was also no significant difference between the misconceptions on the delayed posttest between the erroneous example group and the problem solving group, $p= .116$.

To examine whether the prior knowledge affected the use of the intervention, participants were first placed in high and low prior knowledge groups using a median split based on pretest performance. As predicted, students classified as having high prior knowledge scored higher on both the immediate, $F(1,249)=3.5$, $MSE=220.82$, $p<.001$, and the delayed posttest, $F(1,249)=3.53$, $MSE=178.87$, $p<.001$. There was no significant interaction between high/low prior knowledge and condition for the immediate, $F(2,249)=.845$, $MSE=.013$, $p=.431$, or delayed posttest, $F(2,249)=.915$, $MSE=.018$, $p=.402$. ANOVAs examining the effect of high/low prior knowledge and total misconceptions made revealed those individuals with higher prior knowledge were less likely to show misconceptions on the pretest, $F(1,249)=268.83$, $MSE=2641.256$, $p<.001$, immediate posttest, $F(1,249)=139.54$, $MSE=1888.75$, $p<.001$, and delayed posttest, $F(1,249)=165.17$, $MSE=2405.8$, $p<.001$. No significant interaction was found for condition and prior knowledge level on overall number of misconceptions made on any of the tests.

Our main hypothesis that erroneous examples would lead to positive learning effects, particularly for higher prior-knowledge learners, was partially confirmed. Students in the erroneous examples group learned as much as students in the two control groups. Similar to previous findings in the literature our results indicate that erroneous example is a good instructional tool to help students learn and the view that errors can be dangerous and potentially hurt learning is not supported. However, conversely to what has been reported in other studies, high prior knowledge students did not learned more than low prior knowledge ones and no significant learning difference was found between the three conditions. A possible explanation is that erroneous examples are not as effective as hypothesized. In fact, students in problem solving condition spent considerably less time to learn the same content compared to students in worked examples and erroneous examples conditions; thus, problem solving was a more efficient learning treatment.

Another explanation for the results may lie in how the three-step learning process discussed in section 3 was implemented in our materials. While in the problem solving condition whenever students make a mistake the "Show me the answer" button appears and they could check the correct answer immediately, in the erroneous examples and worked examples condition to self-explain an incorrect or a correct solution students have to *read* all alternatives and select the best one to build the correct self-explanation sentence. Furthermore, no hint is given to students if they fail to build the correct sentence. Thus, we hypothesize that students were overwhelmed and that our erroneous examples and worked examples materials were too wordy and difficult for middle school students to read. It is also possible that these materials did not draw enough attention of students to motivate them to truly reflect on the correct/incorrect solutions because the tests were not used as a class grade. Therefore, there is a chance that some of the benefits of erroneous examples were lost. A third hypothesis is that perhaps the erroneous examples effect, like other effects (e.g., politeness effect), is much harder to achieve in the lively, buzzing confusion of real-life classroom (vs. lab) settings where most learning actually occurs.

To answer some of these questions and hypotheses we are preparing new materials that will simplify the self-explanation sentences and animate the examples to draw attention to important steps. We will also ask teachers to use materials as a class grade, so the students have more motivation to solve the problems. It also part of our plan to conduct think-aloud pilot studies with the new materials to have a better understanding about how the treatment might work in classrooms.

6 Conclusions

This paper discusses a classroom study with 255 subjects from 6th, 7th and 8th grades to better understand the impact of erroneous examples on decimal learning. To evaluate the effects of erroneous examples we developed web-based materials that use decimal misconceptions to present errors to students and help them learn by reflecting upon errors using a three-step process. We also compared erroneous examples with more typical instructional materials namely, worked examples and problem solving.

Our study produced mixed results. First, similar to previous research, students displayed many misconceptions representing 31% of the overall answers given on the pretest, posttest, and delayed posttest. The most frequent ones were *pegz* - two sides of decimals viewed as separate numbers - and *mugz* – multiplication makes bigger. Second, unlike earlier empirical work that showed the benefits of erroneous examples for higher-prior knowledge learners, our results indicate that higher prior knowledge students did not benefit more from erroneous examples than from worked examples or problem solving. Furthermore, although the incidence of students' misconceptions declined significantly in all conditions, it was more accentuated for the worked examples group. These results suggest that erroneous examples are not as effective for learning as we – and other – researchers have hypothesized. It also could indicate that erroneous examples, like some other e-learning techniques, do not work as well in the rough and tumble classroom as they do in a laboratory setting. Nevertheless, there are some changes in the manner in which erroneous examples were presented that could improve learning from errors. Thus, we are currently developing new erroneous examples materials that will potentially provide better learning outcomes and planning a new study to provide more precise answers about the effectiveness of learning from erroneous examples.

Acknowledgments. The U.S. Department of Education (IES), Award No: R305A090460, provided support for this research. We also thank the Pittsburgh Science of Learning Center, NSF Grant # 0354420, for technical support of our work.

References

1. National Math Panel Report (2008), http://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf
2. Glasgow, R., Ragan, G.; Fields, W.M.; Reys, R. & Wasman, D. The decimal dilemma. Teaching Children Mathematics, vol. 7, pp. 89--93 (2000)
3. Rittle-Johnson, B., Siegler, R.S., & Alibali, M.W. Developing conceptual understanding and procedural skill in mathematics: An iterative process. Journal of Educational Psychology, 93(2), 346-362 (2001)
4. Stacey, K. & Steinle, V. Refining the classification of students' interpretations of decimal notation. Hiroshima Journal of Mathematics Education, vol. 6, pp. 49--69 (1998)
5. Putt, I. J. Preservice teachers ordering of decimal numbers: When more is smaller and less is larger! Focus on Learning Problems in Mathematics, 17(3), pp. 1--15 (1995)
6. Grosse, C.S. & Renkl, A. Finding and fixing errors in worked examples: Can this foster learning outcomes? Learning and Instruction, 17(6), pp. 612-634 (2007)

7. Tsovaltzi, D., Melis, E., McLaren, B.M., Meyer, A-K., Dietrich, M. & Goguadze, G. Learning from erroneous examples: When and how do students benefit from them? In Proc. of the European Conference on Technology Enhanced Learning, LNCS vol. 6383, pp. 357—373. Springer, Heidelberg (2010)
8. Borasi, R. Exploring mathematics through the analysis of errors. For the Learning of Mathematics: An Int. Journal of Mathematics Education, 7(3), pp 2--8 (1987)
9. OECD. Knowledge and Skills for Life. First Results from PISA 2000, Paris (2001)
10. Bransford, John, Brown, Ann L., & Cocking, Rodney R. How People Learn: Brain, Mind, Experience, and School. Washington, D.C.: National Academy Press (1999)
11. Skinner, B. F. The Technology of Teaching. New York: Appleton-Century-Crofts (1968)
12. Tsamir, P. & Tirosh, D. In-service mathematics teachers' views of errors in the classroom. In International Symposium: Elementary Mathematics Teaching (2003).
13. Siegler, R.S. Microgenetic studies of self-explanation. In N. Granott and J. Parziale (eds). Microdevelopment, Transition Processes in Development and Learning, pp. 31-58. Cambridge University Press. (2002)
14. Durkin, K. & Rittle-Johnson, B. (submitted). The effectiveness of comparing correct and incorrect examples for learning about decimal magnitude. The Journal of Experimental Child Psychology.
15. Kopp, V., Stark, R., & Fischer, M. R. Fostering diagnostic knowledge through computer-supported, case-based worked examples: Effects of erroneous examples and feedback. Medical Education, vol. 42, pp. 823--829 (2008).
16. Melis, E. Design of erroneous examples for ActiveMath. Proceedings of the Int. Conference on Artificial Intelligence in Education, pp. 451-458. IOS Press (2005)
17. Brueckner, L.J. Analysis of Difficulties in Decimals. Elementary School Journal, 29, 32-41 (1928)
18. Graeber, A. & Tirosh, D. Multiplication and division involving decimals: Preservice elementary teachers' performance and beliefs. Journal of Mathematics Behavior, 7, 263-280 (1988)
19. Hiebert, J. Mathematical, Cognitive, and Instructional Analyses of Decimal Fractions. Chapter 5 in Analysis of arithmetic for mathematics teaching, pp 283-322. Lawrence Erlbaum (1992).
20. Irwin, K.C. Using everyday knowledge of decimals to enhance understanding. Journal for Research in Mathematics Education, 32(4), pp. 399--420 (2001)
21. Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. Conceptual bases of arithmetic errors: The case of decimal fractions. Journal for Research in Mathematics Education, 20(1), pp. 8--27 (1989)
22. Isotani, S., McLaren, B., Altman, M. Towards Intelligent Tutoring with Erroneous Examples: A Taxonomy of Decimal Misconceptions. Proc. of the Int. Conference on Intelligent Tutoring Systems. LNCS 6095, pp. 346--348 Springer, Heidelberg (2010)
23. Bell, A. Swan, M. & Taylor, G. Choice of operations in verbal problems with decimal numbers. Educational Studies in Mathematics, 12, pp. 399-420 (1981)
24. Fischbein, E., M. Deri, M. Nello, & M. Marino The role of implicit models in solving verbal problems in multiplication and division. Journal of Research in Mathematics Education, 16, pp 3-17 (1985)
25. Ohlsson, S. (1996). Learning from performance errors. Psychological Review, Vol. 103, No. 2, 241-262
26. Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. The Cognitive Tutor Authoring Tools (CTAT): Preliminary Evaluation of Efficiency Gains. Proc. of the Int. Conf. on Intelligent Tutoring Systems. LNCS 4053, pp. 61-70. Springer (2006)
27. Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z. Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game The System". Proceedings of ACM CHI Computer-Human Interaction, pp. 383-390 (2004)