

McLaren, B.M., Adams, D., Durkin, K., Goguadze, G. Mayer, R.E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S. & Van Velsen, M. (2012). To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In: A. Ravenscroft, S. Lindstaedt, C. Delgado Kloos, & D. Hernández-Leo (Eds.), *Proceedings of EC-TEL 2012: Seventh European Conference on Technology Enhanced Learning*, LNCS 7563 (pp. 222-235). Springer, Berlin.

To Err is Human, To Explain and Correct is Divine: A Study of Interactive Erroneous Examples with Middle School Math Students

Bruce M. McLaren^{1,2}, Deanne Adams³, Kelley Durkin⁴, George Goguadze², Richard E. Mayer³, Bethany Rittle-Johnson⁴, Sergey Sosnovsky², Seiji Isotani⁵, and Martin van Velsen¹

¹ Carnegie Mellon University, U.S.A.

² The Center for e-Learning Technology (CeLTech), Saarland University, Germany

³ University of California, Santa Barbara, U.S.A.

⁴ Vanderbilt University, U.S.A.

⁵ The University of São Paulo, Brazil

bmclaren@cs.cmu.edu

Abstract. Erroneous examples are an instructional technique that hold promise to help children learn. In the study reported in this paper, sixth and seventh grade math students were presented with erroneous examples of decimal problems and were asked to explain and correct those examples. The problems were presented as interactive exercises on the Internet, with feedback provided on correctness of the student explanations and corrections. A second (control) group of students were given problems to solve, also with feedback on correctness. With over 100 students per condition, an erroneous example effect was found: students who worked with the interactive erroneous examples did significantly better than the problem solving students on a delayed posttest. While this finding is highly encouraging, our ultimate research question is this: how can erroneous examples be *adaptively* presented to students, targeted at their most deeply held misconceptions, to best leverage their effectiveness? This paper discusses how the results of the present study will lead us to an adaptive version of the erroneous examples material.

Keywords: erroneous examples, interactive problem solving, adaptation of problems, self-explanation, decimals, mathematics education

1 Introduction

An instructional technique that has recently drawn attention from learning science researchers is *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. Students can be challenged to find the error(s), explain the error(s), and/or fix the error(s) in order to more deeply learn the domain content and develop metacognitive skills. However, the use of erroneous examples for learning is controversial. On the one hand, some teachers fear that presenting errors to students will make them more inclined to make those errors [1], which is an idea supported by behaviorist theory

[2]. On the other hand, some educators have argued that presenting students with errors for review and discussion can be valuable for learning. For instance, Borasi [3] has argued that mathematics education might benefit from students working with errors, encouraging critical thinking about mathematical concepts and motivating reflection and inquiry.

Our view is that erroneous examples are likely to be helpful to students under three basic conditions. First, the errors should be fictitious examples of *other* students' errors, so the student reviewing the errors is freed from embarrassment – and possible demotivation – of having their own errors exposed. Furthermore, no other real student is put on the spot in front of classmates. Second, the erroneous examples should be *interactive* and *engaging*; in particular, they should be computer-based materials that prompt for explanations, ask students to find and correct errors, and provide feedback. Finally, the erroneous examples should be *adaptively* targeted to the particular needs of individual students. That is, the types of problems presented to students should be aimed at their most deeply held misconceptions and misunderstandings about the target domain.

In short, our hypothesis is that the erroneous examples, presented to students in an interactive and adaptive fashion (e.g., presenting examples when a student is ready, withholding when not), can provide the opportunity to find and reflect upon errors in a manner that will lead to deeper, more robust learning. In this paper we present the results of a study that shows that interactive erroneous examples of others *can* provide learning benefits. Furthermore, we present some data and ideas regarding the next step of our research; that is, making the erroneous examples adaptive to student needs.

2 The Potential of Erroneous Examples for Learning

Research on erroneous examples derives from work on *correct* worked examples, which has attracted much attention in learning science empirical research [4, 5, 6, 7, 8]. Much research has also shown the importance of prompted self-explanation of worked examples, particularly in multi-media learning environments [9]. The theory behind the worked examples effect is that human working memory, which has a limited capacity, is taxed by strictly solving problems, which requires focused thinking, such as setting subgoals. Problem solving consumes cognitive resources that could be better used for learning [10]. The rationale is that worked examples free cognitive resources for learning, in particular, for the induction of new knowledge.

Erroneous examples also appear to free working memory for learning, by providing much of what students need to understand and solve problems, but, at the same time, may engage students in a different form of active learning. It appears that erroneous examples may help students become better at evaluating and justifying solution procedures, which, in turn, may help them learn material at a deeper level. Learning with erroneous examples may also be related to the notion of “learning by teaching”, as students who find, correct, and explain errors assume a role akin to teaching or tutoring [11].

Some researchers have begun to investigate empirically the use of erroneous examples, attempting to better understand whether, how, and when they make a difference to learning. For instance, Siegler [12] investigated whether self-explaining correct and incorrect examples of mathematical equality were more beneficial than self-explaining correct examples only. He found that students who studied and self-explained both correct and incorrect examples led to the best learning outcomes. Grosse and Renkl [13] studied whether explaining both correct and incorrect examples made a difference to university students as they learned statistics. Their studies also showed learning benefits of erroneous examples but unlike the less ambiguous Siegler results, the benefit was only for learners with higher prior knowledge and for far transfer learning only. When errors were highlighted, on the other hand, low prior knowledge individuals did significantly better, while high prior knowledge students did not benefit, presumably because they were already able to identify errors on their own.

Recently, there has been increasing investigation of *interactive* erroneous examples, those that are computer-based, that allow students editing and correction, and for which feedback is provided. Unlike the Siegler and Grosse and Renkl studies, Tsovaltzi *et al* [14] presented erroneous examples of fractions to students using an interactive intelligent tutoring system with feedback. They found that 6th grade students improved their metacognitive skills when presented with erroneous examples with interactive help, as compared to a problem solving condition and an erroneous examples condition with no help. Older students – 9th and 10th graders – did not benefit metacognitively but did improve their problem solving skills and conceptual understanding by working with interactive erroneous examples that included help.

Encouraged that interactive erroneous examples are promising instructional materials, our project team ran a study of decimal learning, in which we compared an interactive erroneous examples condition to (1) a worked examples condition and (2) a supported (i.e., with correctness feedback) problem solving condition [15]. However, the interactive erroneous examples did not lead to better learning results than worked examples or problem solving, nor was there an interaction between high and low prior knowledge and condition. We attributed this finding to two things. First, the prompted self-explanation of erroneous examples in this study was (potentially) too cognitively taxing. Students were asked to complete explanations of incorrect steps by filling in two phrases of a sentence, using pull-down menus. We observed students struggling with this task, possibly undercutting their math learning. Second, while we presented erroneous examples to review and compare to correct examples, we did not prompt students to *find* and *correct* the errors.

The study presented in this paper was focused, first, on correcting the perceived problems with the prior study's materials and, second, on collecting data so we can learn how to adapt the presentation of erroneous examples to lead to the best possible learning outcomes.

3 The Domain: Learning Decimals

The domain we have focused on for this study is decimals. A variety of studies have shown that many students have difficulty mastering decimals and have common and persistent misconceptions [16, 17, 18], as well as problems that extend into adulthood [19, 20]. For instance, students often treat decimals as if they are whole numbers (e.g. they think 0.15 is greater than 0.8, since 15 is greater than 8, i.e., “longer decimals are larger”) or they think that all decimals are less than zero. Persistent misconceptions in students’ decimal knowledge must be overcome so students can handle everyday tasks (e.g., money calculations) and tackle more advanced mathematics.

Our general approach to addressing decimal learning with erroneous examples has been to develop problems that focus on single, key misconceptions. Based on an extensive literature review, we created short names for and developed a taxonomy of misconceptions that represents 17 misconceptions [15]. The present study focuses on four of these misconceptions, the ones that prior research has shown are most common and contributory to other misconceptions: Megz (“longer decimals are larger”, e.g., $0.23 > 0.7$), Segz (“shorter decimals are larger”, e.g., $0.3 > 0.57$), Negz (“decimals less than 1.0 are negative”), and Pegz (“the numbers on either side of a decimal are separate and independent numbers”, e.g., $11.9 + 2.3 = 13.12$).

4 The Study

For the current study we revised the materials from Isotani *et al* [15] by, in the interactive erroneous examples condition, simplifying the self-explanation step – asking students to complete sentences with one multiple-choice phrase instead of two – and by prompting the student to find and fix the errors in the erroneous examples. We also removed all problems, both on the tests and the intervention, related to two of the misconceptions explored in the earlier study (i.e., “multiplication always makes bigger” and “division always makes smaller”). This was done so we could focus on the most common misconceptions. Finally, we also simplified the experimental design, comparing only supported problem solving and interactive erroneous examples, while dropping the worked examples condition. We did this for two reasons. First, we wanted to compare the most common ecological control condition – that of students solving problems – to the much less typical learning experience of working with erroneous examples. Second, erroneous examples and problem solving are more comparable from a cognitive load perspective. As designed, they both require active problem solving – in the case of erroneous examples, the correction step; in the case of problem solving, generating the solution from the given problem – something worked examples do not (typically) require. Besides discovering whether erroneous examples could make a difference to learning, we had a goal of collecting data to help us determine how to implement automated material adaptation in a subsequent experiment.

4.1 Participants

247 students from a suburban middle school started the study, but 39 were eliminated from analysis (due to missing class or not trying on one or more of the tests, i.e., test score of 0, or close to 0). The remaining 208 participants consisted of 101 male and 107 female students. 105 of the students were in the 6th grade and 103 were in the 7th grade. Ages ranged from 11 to 13 ($M = 11.99$, $SD = .722$). All students worked on the materials exclusively during class time, under the supervision of the teachers and at least one experimenter, as further explained below.

4.2 Design and Materials

The materials consisted of six components, all presented to students on the Internet with all interactions and feedback implemented and logged using the CTAT intelligent tutoring authoring software [21]. The activities of the study are shown in Fig. 1: a pretest, a questionnaire on demographic/math experience (Q1), the intervention problems, a questionnaire asking about the student's experience with the materials (Q2), an immediate posttest, and a delayed posttest.

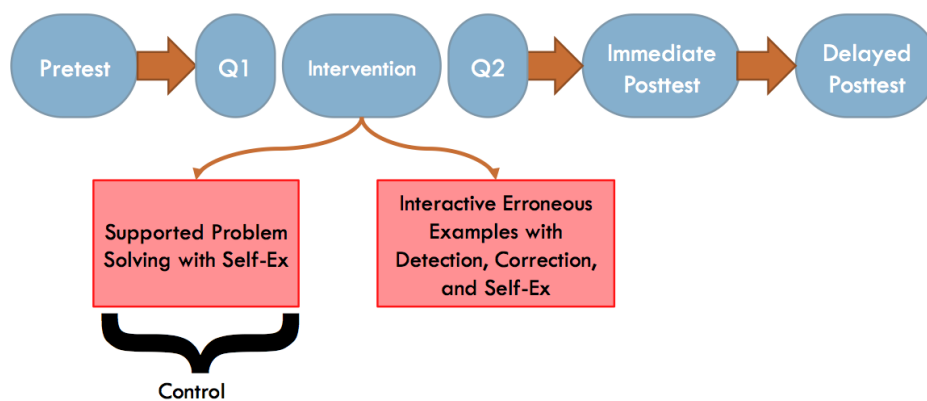


Fig. 1: The study design and sequence of activities

For the pretest and two posttests, three separate but isomorphic tests were constructed (Tests A, B, and C). Every problem on the test was designed to diagnose at least one of the misconceptions of interest – Megz, Segz, Negz, or Pegz. The problems on the tests included placing decimals on a number line, arranging decimal numbers in order by dragging and dropping numbers into a sequence, providing the next two numbers in a given sequence, and defining key decimal concepts by selecting from a multiple-choice list of possible definitions. All tests contained a total of 50 possible answers. For the demographic questionnaire (Q1), students were asked their gender, age, and grade level, as well as questions about their experience with decimals and computers. For the post-questionnaire (Q2), students were asked

questions such as, “I would like to do more lessons like this”, with answers provided on a 5-point Likert scale.

In the intervention, the two groups were presented with isomorphic decimal problems, but with different presentations and ways of interacting with the problems. As shown in Fig. 2, the erroneous examples subjects were (a) presented with an incorrect solution by a fictitious student (upper left panel), (b) prompted to explain what the student had done incorrectly (upper right panel), (c) asked to detect and correct the error (middle left panel) and (d) prompted to explain and reflect on the correct answer (middle right panel and bottom left panel). They received feedback on their responses (i.e., green=correct; red=incorrect; with supportive feedback such as “You’ve got it. Well done.” displayed in the lower right panel).

Ryan has a chocolate bar that weighs 0.9 pounds and a chocolate bar that weighs 0.37 pounds. If Ryan wants to choose the largest chocolate bar, which one should he choose?

Ryan said this incorrect answer: I want the 0.37 pound chocolate bar because it is the largest. I used a number line to solve this problem.

What did Ryan do wrong?
Ryan thinks that decimals are like ____.

- fractions, and shorter numbers are less than zero
- whole numbers, and shorter numbers are larger
- whole numbers, and shorter numbers are smaller
- fractions, and shorter numbers are smaller

Click on the line where the incorrect point should go to fix Ryan's error.

Using the corrected number line, which chocolate bar is larger?

- 0.9 pounds because it is closer to 1
- 0.37 pounds because it is closer to 0

What advice would you give to Ryan to solve the problem right next time?
Ryan, to find the largest decimal, you should see which decimal is ____.

- longest
- farthest to the left on the number line
- shortest
- farthest to the right on the number line

Decimals are not the same as whole numbers because a decimal being longer ____ mean it is larger.

- does not
- does

Message Window
You've got it. Well done.

← Previous Next →

Done

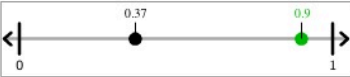
Fig. 2: Sample interactive erroneous example, targeted at the Megz misconception

As shown in Fig. 3, the supported problem solving subjects were (a) asked to solve problems (upper panel) and (b) prompted to explain and reflect on the correct answers. These students also received feedback on their solutions (i.e., green=correct; red=incorrect; with supportive statements, as in the erroneous examples condition).

As shown in Table 1, the intervention comprised a total of 36 problems, 24 of which had interactions such as that illustrated in Figures 2 and 3 (according to condition), and 12 of which were problems to solve (the same across conditions), presented to the students to encourage active processing of the concepts and skills just presented. The problems were arranged in groups of three, each group targeting one

of the misconceptions of interest (Megz, Segz, Pegz, and Negz – Highlighting in the table shows the grouping by threes).

You have a chocolate bar that weighs 0.9 pounds and a chocolate bar that weighs 0.37 pounds. If you want to choose the largest chocolate bar, which one should you choose? Figure out how big 0.9 is by placing it on the number line. 0.37 is already marked.



How would you explain your solution to another student?

To find the largest decimal, I chose the one that was _____.

- farthest to the right on the number line
- longer
- shorter
- farthest to the left on the number line

When comparing two decimals, you can compare the numbers in each place value _____.

- and then choose the longer decimal as the larger
- from left to right
- and then choose the shorter decimal as the larger
- from right to left

Message Window

You've got it. Well done.

← Previous Next →

Done

Fig. 3: Sample supported problem to solve, (isomorphic to the erroneous example problem in Fig. 2 and also targeted at the Megz misconception)

4.3 Procedure

The students were randomly assigned to either the supported problem solving or the erroneous example condition and to one of the six possible pretest / immediate posttest / delayed posttest orderings (e.g., ABC, ACB, BCA, etc.). The study took place exclusively in computer rooms in the school, replacing regular class time. The students were given a total of five 43-minute periods to complete the entire set of materials shown in Figure 1. They started the pretest on Day 1 and were allowed to continue immediately to the questionnaires and intervention, using as much of the first three days to work on these materials as needed. If students finished the materials early, they were asked to work on other, non-decimal materials and not disturb the students still working. On the 4th day all of the students were given the immediate posttest and 6 days later, on Day 5, they were all given the delayed posttest. Between the time the students took the immediate posttest and the time they took the delayed posttest, they received no classroom exposure to decimals and were blocked from working with the web-based decimal materials.

Table 1: Design of the Intervention

Supported Problem Solving (PS)	Int. Erroneous Examples (ErrEx)
1. Megz Supported PS 1	1. Megz ErrEx 1
2. Megz Supported PS 2	2. Megz ErrEx 2
3. Megz PS 1	3. Megz PS 1
4. Segz Supported PS 1	4. Segz ErrEx 1
5. Segz Supported PS 2	5. Segz ErrEx 2
6. Segz PS 1	6. Segz PS 1
7. Pegz Supported PS 1	7. Pegz ErrEx 1
8. Pegz Supported PS 2	8. Pegz ErrEx 2
9. Pegz PS 1	9. Pegz PS 1
10. Negz Supported PS 1	10. Negz ErrEx 1
11. Negz Supported PS 2	11. Negz ErrEx 2
12. Negz PS 1	12. Negz PS 1
13. Megz Supported PS 3	13. Megz ErrEx 3
14. Megz Supported PS 4	14. Megz ErrEx 4
15. Megz PS 2	15. Megz PS 2
16. Segz Supported PS 3	16. Segz ErrEx 3
17. Segz Supported PS 4	17. Segz ErrEx 4
18. Segz PS 2	18. Segz PS 2
19. Pegz Supported PS 3	19. Pegz ErrEx 3
20. Pegz Supported PS 4	20. Pegz ErrEx 4
21. Pegz PS 2	21. Pegz PS 2
22. Negz Supported PS 3	22. Negz ErrEx 3
23. Negz Supported PS 4	23. Negz ErrEx 4
24. Negz PS 2	24. Negz PS 2
25. Megz Supported PS 5	25. Megz ErrEx 5
26. Megz Supported PS 6	26. Megz ErrEx 6
27. Megz PS 3	27. Megz PS 3
28. Segz Supported PS 5	28. Segz ErrEx 5
29. Segz Supported PS 6	29. Segz ErrEx 6
30. Segz PS 3	30. Segz PS 3
31. Pegz Supported PS 5	31. Pegz ErrEx 5
32. Pegz Supported PS 6	32. Pegz ErrEx 6
33. Pegz PS 3	33. Pegz PS 3
34. Negz Supported PS 5	34. Negz ErrEx 5
35. Negz Supported PS 6	35. Negz ErrEx 6
36. Negz PS 3	36. Negz PS 3

5 Results

Due to bugs in four of the problems in the tests, the data for those problems was removed from all students, leaving a total possible score of 46 for each test for all students. The mean score for the pretest, immediate posttest, and delayed posttest, per

condition, is shown in Figure 4. Standard deviations were 9.4, 9.6, 9.5, 9.0, 9.2, and 7.5 for the left-to-right scores of Figure 4.

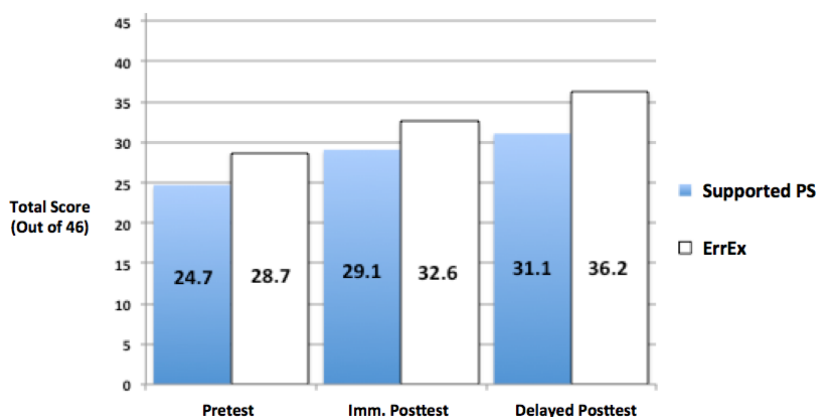


Fig. 4: Mean scores of pretest, immediate posttest, and delayed posttest

All students significantly improved their test performance from pretest to immediate posttest ($t(207) = -8.058, p < .001$, mean increase of 9%) and from pretest to delayed posttest ($t(207) = 14.496, p < .001$, mean increase of 15%, a medium-to-large effect, $d = .75$). An independent sample t-test revealed that the ErrEx group performed significantly better on the pretest than the Supported PS group, $t(206) = 3.045, p = .003$. An ANOVA revealed that there was no significant difference between the order of the three tests (A, B, C), $F(5,202) = 1.293, MSE = .057, p = .268$; thus, we assume the three tests were truly isomorphic to one another.

To determine whether the conditions differed significantly on the immediate and delayed posttest, an ANCOVA with pretest as a covariate was used. Results showed:

1. There were no significant differences in performance on the immediate posttest between the two groups, $F(1,205) = .768, MSE = 34.97, p = .382, d = .38$.
2. On the delayed posttest there was a significant difference, in favor of the ErrEx condition, $F(1,205) = 9.896, MSE = 349.08, p = .002, d = .62$.

In other words, although students in the ErrEx condition did not do significantly better than the students in the Supported PS condition on the immediate posttest, they *did* do significantly better on the delayed posttest, after a six-day delay. This occurred even though none of the students received further decimal training or practice between the immediate and delayed posttests.

To see whether higher prior knowledge students benefitted more from erroneous examples than low prior knowledge students, as they had in prior studies (e.g., [13]), we divided the participants using a median split into a high prior knowledge group (score of 26-46, total of 101 students) and a low prior knowledge group (score of 8-25, total of 107 students). ANCOVAs on posttest scores, with pretest covariates, showed the same outcome as 1-2 above, regardless of prior knowledge. That is, for *both* low and high prior knowledge learners there were no significant differences between conditions on the immediate-posttest, while the differences on the delayed-posttest were significantly better for the erroneous examples condition.

6 Discussion

Our study provides clear evidence that erroneous examples can help students learn. The delayed posttest results show that the students who were presented with interactive erroneous examples learned better than those who were presented with supported problems to solve. This suggests that erroneous examples may provide a deeper learning experience, one that can help students build upon their initial understanding of decimals, leading to a deeper understanding over time. More specifically, it appears that erroneous examples might encourage generative processing (i.e., deeper cognitive processing that organizes the material and relates it to prior knowledge). This generative processing may be due to an effect referred to as “desirable difficulty” [22], in which problems of a more challenging form have been shown to lead to delayed learning benefits. The interactive erroneous examples of this study were very likely the harder of the two intervention types – and almost certainly less familiar to students – so they may have helped long-term retention.

On the other hand, unlike the results of Grosse and Renkl’s study [13], we did not find that higher prior knowledge students benefited more from erroneous examples. We also expected that higher prior knowledge students would get more benefit than lower prior knowledge students in this study. Perhaps our materials, unlike those of the Grosse and Renkl study, were designed straightforwardly enough so that even lower prior knowledge students could easily follow, interact with, and learn from the examples without incurring excessive cognitive load. Indeed, one of our goals in this study was to simplify and streamline the prior year’s study materials, in which no erroneous examples effect was found [15]. We did the streamlining by, for instance, making the self-explanation statements pure multiple choice, rather than sentence construction items with multiple components. The Grosse and Renkl work was also different in that it focused on errors related to confusing problem types instead of deeply entrenched misconceptions, which is what the current study focused on. In other words, erroneous examples may be more helpful for students with low prior knowledge when they involve common misconceptions.

One caveat to our results is that more than a single variable differs between the erroneous examples and the supported problem solving materials. For instance, while students in both condition had to self-explain their work, the erroneous examples condition had the extra self-explanation step, prompting students to explain why the fictitious student may have made the given error. Yet, our goal in this study, which we view as one step in an exploration of how erroneous examples might benefit learning, was not to demonstrate the full generality of erroneous examples. Rather, our objective was to see if we could find an advantage to erroneous examples – which our results have clearly shown – and, in future studies, isolate the instructional features of the materials that might account for the benefits. It is also worth repeating that our erroneous examples intervention, while differing in multiple ways from supported problem solving, has still demonstrated advantages to the most obvious and common control condition, that of students solving problems in conventional fashion.

7 Future Work

Our ultimate goal is to determine *when* and *with what students* we should present erroneous examples, and also to determine *what types of erroneous examples* to present to students. We have been pursuing this objective by modeling students with a Bayes Net of decimal misconceptions [23], which is updated when students take the tests described above. The Bayes Net represents the misconceptions that a student might have – Megz, Segz, Negz, and Pegz – and is updated based on carefully crafted test questions that probe for each of the misconceptions. Our isomorphic tests, A, B, and C, contain 9 Megz problems, 10 Segz problems, 10 Pegz problems, and 9 Negz problems, after the 4 buggy problems are eliminated (there are also 8 problems that are targeted at a more general misconception called Regz, which contributes to all of the other misconceptions). Students can either get these problems correct, in which case the probability of the targeted misconception drops, they can get them incorrect in an unexpected way, in which case the misconception is only partially increased, or they can get them incorrect in a way that provides direct evidence for the misconception in the Bayes Net. The tests were designed so that the misconception problems are relatively evenly distributed across the tests. Some of the misconception problems have possible answers that can indicate more than one misconception. The details of the Bayes Net are discussed in [23]. Our approach was inspired by the similar implementation of Stacey *et al* [24].

Given how the Bayes Net of each of the 208 students in the present study were updated, we calculated mean probabilities over all misconceptions: Segz=0.37; Megz=0.31; Pegz=0.15; Negz=0.15. Furthermore, we created *misconception profiles* for all of the students, based on the order of probability of each of the misconceptions for each student. For instance, a student with a Megz probability of 0.92, Segz probability of 0.75, Pegz probability of 0.32 and Negz probability of 0.2 would have a misconception profile of Megz>Segz>Pegz>Negz. Table 2 summarizes the misconception profiles of all the students by most prominent misconception, i.e., the misconception that has the highest probability.

Table 2: Summary of the misconception profiles of all 208 students

Description	#	Pre	Megz	Segz	Pegz	Negz	General Misconception Profile
Megz is the most prominent misc.	42	18.1	0.97	0.55	0.28	0.19	Megz>Segz>Pegz>Negz
Segz is the most prominent misc.	60	19.3	0.34	0.89	0.10	0.21	Segz>Megz>Negz>Pegz
Pegz is the most prominent misc.	58	33.9	0.03	0.02	0.22	0.03	Pegz>Megz>Negz>Segz
Negz is the most prominent misc.	48	34.3	0.00	0.00	0.00	0.19	Negz>Pegz>Megz>Segz

Key: The “#” column is the number of students with this misconception profile. The “Pre” column is the number of items, on average out of 46, that students in this row got correct on the pretest. The values under the “Megz”, “Segz”, “Pegz”, and “Negz” columns are the average probabilities, according to the Bayes Net, that students in this row have each of these misconceptions.

As can be seen, the students were reasonably well distributed across the most prominent misconception categories, but there are stark differences in the mean values. Note that students who displayed the Megz (“longer decimals are larger”) and Segz (“shorter decimals are larger”) misconceptions as their most likely misconception, show a very high probability for actually having those misconceptions (see bold items in rows 1 and 2), while the students who displayed the Pegz (“each side of the decimal is separate and independent”) and Negz (“decimals less than 1.0 are negative”) misconceptions as most likely, show a much lower probability for actually having those misconceptions (see bold items in rows 3 and 4). Furthermore, the pretest scores of the Megz and Segz students are dramatically lower than the Pegz and Negz students. Finally, the *other* possible misconceptions of the Megz and Segz students have a much high probability than those of the Pegz and Negz students.

What does this tell us? First, these results are in line with the math education literature, which clearly indicates that Megz and Segz are the most likely decimal misconceptions of middle school math students. Having recently learned whole numbers and fractions, middle school students are very susceptible to the mistake of thinking longer decimals are larger (as is so with whole numbers) or that shorter decimals are larger (as is so with shorter denominators in fractions). Students who struggle with either (or both) of these misconceptions are much more likely to do poorly on decimal tests. Second, these results give us some clues about how to adapt our materials to particular students. Clearly, our system is more likely to be successful in helping students by emphasizing the Megz and Segz problems. Our initial plan is, not surprisingly, to provide more intervention problems aimed at the misconceptions for which students have shown they may have, according to the Bayes Net and the resulting misconception profiles. We will retrieve the misconception profile for each student and then provide that student with an intervention curriculum catered to that profile. The curriculum associated with each misconception profile will be weighted toward providing more problems aimed at that student’s highest-probability misconception, less problems at the next highest probability misconception, and so forth. Considering the three-problem “problem groups” of Table 1, given a student with misconceptions in the order A, B, C, D, we might present 4 problem groups aimed at misconception A; 3 problem groups aimed at misconception B; 2 problem groups aimed at misconception C; and 1 problem group aimed at misconception D. Curricula will be variable, though, dependent on how different the probabilities are within a profile, e.g., given the clear need to ameliorate the Negz misconception in the misconception profile of the last row of Table 2, we might present a student that has such a profile with many more Negz problem groups than any other problem groups. We are well positioned to identify the most likely curricula needed; we have mined data for all of the 208 students in the study from the Bayes Net, created their misconception profiles, and have quantitative data to guide our approach. For instance, we have discovered that, of the 24 possible profiles (all of the permutations of the 4 misconceptions), 6 profiles never occur. Within specific profiles we have also discovered that some students have high probability values, very close to 1, while others have very low probability values, very close to 0, suggesting that even within each misconception profile, we will want to adjust curricula per student.

Besides this relatively straightforward adaptation proposal, we will investigate more sophisticated strategies. For instance, it is likely that misconceptions are causally interrelated, to a certain extent, so we will investigate ways to identify causality and use it to make adaptation decisions. The misconception profiles could also be adjusted based on either (or both) the Likert or questionnaire data that we collect as part of our study. For instance, a student who says he or she is “very sure” of an incorrect (and misconception) answer would lead to a more weighted update of the Bayes Net than a student who says he or she is “unsure” of an incorrect answer.

8 Conclusion

This paper has presented a study that provides evidence that interactive erroneous examples may be helpful to learning, especially over time, when a student has had an opportunity to reflect. Our next step is to investigate how we can adapt our erroneous examples material according to user models represented as Bayes Nets of decimal misconceptions. We will investigate a relatively straightforward adaptation strategy to see if it can be helpful to learning and then explore more complex strategies.

Acknowledgements: 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.

9 References

1. Tsamir, P., Tirosh, D.: In-service mathematics teachers' views of errors in the classroom. In *International Symposium: Elementary Mathematics Teaching*, Prague, (2003)
2. Skinner, B.F.: *The behavior of organisms: An experimental analysis*. Appleton-Century, New York (1938)
3. Borasi, R.: *Reconceiving Mathematics Instruction: A Focus on Errors*. Ablex Publishing Corporation (1996)
4. Catrambone, R.: The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General* 1998, Vol. 127, No. 4, 355-376 (1998)
5. McLaren, B.M., Lim, S., Koedinger, K.R.: When and how often should worked examples be given to students? New results and a summary of the current state of research. In: *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, (pp. 2176-2181). Austin, TX: Cog. Sci. Society (2008)
6. Renkl, A., Atkinson, R.K.: Learning from worked-out examples and problem solving. In: J.L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory*, Cambridge: Cambridge University Press (2010)
7. Sweller, J., Cooper, G.A.: The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2, 59-89 (1985)
8. Zhu, X., Simon, H.A.: Learning mathematics from examples and by doing. *Cognition and Instruction* 4 (3), 137-66 (1987)

9. Roy, M., Chi, M.T.H.: The self-explanation principle in multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 271-286). New York: Cambridge University Press (2005)
10. Sweller, J., Van Merriënboer, J.J.G., Paas, F.G.W.C.: Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296 (1998)
11. Leelawong, K., Biswas, G.: Designing learning by teaching agents: The Betty's Brain system. *Int'l Journal of Artificial Intelligence in Education*, 18(3), 181-208 (2008)
12. 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)
13. Grosse, C.S., Renkl, A.: Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction*, 17(6), 612-634 (2007)
14. Tsovaltzi, D., Melis, E., McLaren, B.M., Meyer, A-K., Dietrich, M., Gogvadze, G.: Learning from erroneous examples: When and how do students benefit from them? In: *Proceedings of the Fifth European Conference on Technology Enhanced Learning (EC-TEL 2010)*, LNCS vol. 6383, (pp. 420-422) Springer, Heidelberg (2010).
15. Isotani, S., Adams, D., Mayer, R.E., Durkin, K., Rittle-Johnson, B., McLaren, B.M.: Can erroneous examples help middle-school students learn decimals? In: *Proceedings of the Sixth European Conference on Technology Enhanced Learning: Towards Ubiquitous Learning (EC-TEL 2011)*. (pp. 181-195) (2011)
16. Irwin, K.C.: Using everyday knowledge of decimals to enhance understanding. *Journal for Research in Mathematics Education*, 32(4), 399-420 (2001)
17. 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)
18. Sackur-Grisvard, C., Léonard, F.: Intermediate cognitive organizations in the process of learning a mathematical concept: The order of positive decimal numbers. *Cognition and Instruction*, 2, 157-174 (1985)
19. 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), 1-15 (1995)
20. Stacey, K., Helme, S., Steinle, V., Baturo, A., Irwin, K., Bana, J.: Preservice teachers' knowledge of difficulties in decimal numeration. *Journal of Mathematics Teacher Education*, 4, 205-225 (2001)
21. Aleven, V., McLaren, B.M., Sewall, J., Koedinger, K.R.: A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, 19(2), 105-154 (2009)
22. Schmidt, R.A., Bjork, R.A.: New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psych. Sci*, 3(4), 207-217 (1992)
23. Gogvadze, G., Sosnovsky, S., Isotani, S., McLaren, B.M.: Evaluating a bayesian student model of decimal misconceptions. In: *Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011)* (pp. 301-306) (2011)
24. Stacey, K., Sonenberg, E., Nicholson, A., Boneh, T., Steinle, V.: A teacher model exploiting cognitive conflict driven by a Bayesian network. In: *User Modeling 2003: Proceedings of the Ninth International Conference*. (pp. 352-362) (2003)