Erroneous Examples Versus Problem Solving: Can We Improve How Middle School Students Learn Decimals?

Deanne Adams[1] (adams@psych.ucsb.edu)
Bruce M. McLaren[2] (bmclaren@cs.cmu.edu)
Kelley Durkin[3] (kelley.l.durkin@vanderbilt.edu)
Richard E. Mayer[1] (mayer@psych.ucsb.edu)
Bethany Rittle-Johnson[3] (bethany.rittle-johnson@vanderbilt.edu)
Seiji Isotani[4] (sisotani@icmc.usp.br)
Martin van Velsen[2] (vvelsen@cs.cmu.edu)

[1]Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106-9660
[2]Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA 15213
[3]Department of Psychological Sciences, Vanderbilt University, Nashville, TN 37240-7817
[4]Department of Computer Systems, University of Sao Paulo (ICMC-USP)

Abstract

Worked examples have been found to be effective tools in reducing cognitive load and supporting learning. Erroneous examples are worked examples that include incorrect steps and are intended to help students learn how to identify important principles and errors to avoid. The current study examines whether using erroneous examples in an online intelligent tutoring system can help middle-school children learn decimals beyond simple problem solving with feedback. Results showed that although students did not differ between the two conditions on an immediate posttest, students in the erroneous examples group performed better on a delayed posttest. This suggests that working with errors, and thus processing the decimal problems at a deeper level, helped students retain more about decimals and build upon that understanding over time.

Keywords: erroneous examples, math learning, computer-based tutors

Worked Examples and Math Learning

One effective method that has been applied to mathematics education to increase learning is worked-out-examples (also called *worked examples*). Worked examples consist of a problem formation, the steps taken to reach the solution, and the final solution (Cooper and Sweller, 1987; McLaren, Lim, and Koedinger, 2008; Renkl, 2005, 2010; Renkl and Atkinson, 2010; Zhu and Simon, 1987). Worked examples may be effective because they facilitate learning by helping to manage intrinsic processing levels (i.e. cognitive processing required to learn the material presented in a lesson); decreasing extraneous processing (i.e., cognitive processing that does not support the instructional goal); and by encouraging generative processing (i.e., cognitive processing that enables deeper learning). According to the cognitive theory of multimedia learning (Mayer, 2009) and cognitive load theory from which it is derived (Moreno and Park, 2010) learners have a limited processing capacity in working memory and every learning task has an intrinsic level of processing required to understand and learn the task. During problem solving such as mathematics, students use strategies such as means-ends analyses to solve problems, comparing the state of the problem to the goal state and trying to reduce the differences (Renkl and Atkinson, 2010). Over time they develop procedural and schematic knowledge that facilitates problem solving. Worked examples can decrease both intrinsic and extraneous cognitive processing during learning by showing the students the solution procedures to follow. The freed up cognitive resources can then be applied to understanding and eventually to automatizing the different steps in the problem's procedure.

A study by Cooper and Sweller (1987) compared learning by doing/traditional problem solving and learning from worked examples. The results showed that participants in the learning by examples group could answer transfer problems much faster than students who learned by doing although the later group actually had more practice in solving problems.

An important issue with worked examples is that although students may have freed up cognitive resources, this does not mean that the freed cognitive capacity will be used for generative processing (also called germane processing) which requires deeper processing of the material (Renkl and Atkinson, 2010). Students may need further assistance in fully absorbing and learning solution methods or principles. Self-explanation is one way to achieve this. Chi et al. (1989) found that good problem solvers are more likely to generate self-explanation statements while thinking out loud when reading a lesson on physics. In addition, other research has shown the importance of explicitly prompting for self-explanation

(Hausman and Chi, 2002). Explanations can therefore be used to encourage further processing of the material and increase learning.

Erroneous Examples and Learning

One other proposed way to encourage deeper processing while using worked examples is to present students with incorrect (or erroneous) examples. Erroneous examples may encourage students to use more explanations since they must identify and explain to themselves why the solution is incorrect and how it can be corrected. Erroneous examples may also help students focus on each step of a solution method separately to identify where the error occurred. However erroneous examples could also place additional processing demands on learners, overloading working memory. The student may have to simultaneously represent both the correct and incorrect solution steps while searching for what is wrong in the worked example (Grosse and Renkl, 2007). Therefore, learners with low prior knowledge may be more likely to be adversely affected by incorrect examples because they would be unable to hold large chunks of new information in memory while also looking for an error. Grosse and Renkl (2007) suggest relieving this processing demand by highlighting the error. Reiss, Hellmich, and Thomas (2002) found that learners only had a .35 probability of identifying a math false argument as being false while correct arguments had a .67 probability of being identified as correct.

Yet research has shown that erroneous examples can facilitate learning of mathematics. In a study by Kawasaki (2010), 170 5th grade students were presented with either a correct or incorrect solution to a math problem by one of the participants. The teacher then explained the correct solution either contrasting the two procedures for the incorrect or displaying the correct. Students who had used similar incorrect solutions benefitted the most from the instruction with the incorrect example. Tsovaltzi et al. (2010) found mixed results for whether erroneous examples facilitated learning of fractions. For 6th graders they found that including erroneous example, especially with help, increased metacognitive skills such as answering conceptual questions. With 9th and 10 graders, on standard problem solving tests, students in the erroneous examples with help condition outperformed students in the erroneous examples without help and the no erroneous examples groups. They propose that this was due to the low prior knowledge level of the students.

Grosse and Renkl (2007) also found an effect of prior knowledge on the effectiveness of erroneous examples. College level students were taught a lesson on probability. In their first experiment half of the conditions were presented with correct solutions only while the other half were presented with both correct and incorrect solutions. For groups with both incorrect and correct solutions, half of the participants had the error highlighted while the other half did not. The study found an interaction between the prior knowledge of the individual and the inclusion of incorrect solutions. High prior knowledge students benefitted from having both correct and incorrect solutions and scored higher on far transfer problems that did not have solution structures similar to the problems presented during the lesson. In contrast, low prior knowledge students did worse on a far transfer test when given both correct and incorrect solutions. For highlighting the error, high prior knowledge students did not benefit from having errors highlighted (presumably because they were already able to identify the error on their own). Low prior knowledge individuals did significantly better when the errors were highlighted than when they were not. Grosse and Renkl's (2007) second study replicated the prior knowledge incorrect solution interaction but also found that including errors changes the sort of self-explanation statements students made. Students made more elaborations that were error related such as identifying the error or the reasons for the error, however, students in this group also made less principle-based self-explanation. Principle-based explanations have been proven to foster learning outcomes (Renkl, 1997).

In a recent study by Isotani et al. (2011) an online tutoring system with erroneous examples was used to teach decimals to middle school students. Six commonly held misconceptions dealing with decimals were identified, such as decimals being treated as negative numbers or students treating the two sides of a decimal as separate numbers. Participants were separated into three conditions: problem solving, worked examples, and erroneous examples. During the problem solving condition students had to at least attempt to answer a problem once and were given feedback in the form of green or red lettering as to whether their answer was correct. If the student supplied an incorrect answer they could choose to have the correct answer displayed. In the worked example condition students were given a word problem in which the correct answer was given. The students were then asked to complete two sentences that described how the problem was solved and what knowledge about decimals was needed to answer the problem. Students would select responses for the two blanks in the sentence and then receive feedback from the tutor as to whether their created explanation was correct. The erroneous examples problems were similar to the worked examples except that an incorrect solution was presented. It was the job of the students to fill in the blank to generate two sentences: the first identifies the particular decimal misconception while the second sentence prompts the student to explain how the individual in the problem could correctly solve the problem.

The results uncovered no significant differences among the three groups for either immediate posttest or the delayed posttest and unlike Grosse and Renkl (2007) there was no interaction between high and low prior knowledge and condition. One possible reason for no significant differences among the three groups is the amount of cognitive load that the sentence completion task required of the participants. Instead of focusing on the math, the students may have been devoting their cognitive processing to selecting the correct sentence portions and reading their completed sentence.

Present Study

For the current study we have streamlined the materials from Isotani et al. (2011) to increase the focus on finding and fixing errors in erroneous examples. In particular, we simplified that design to compare problem solving to erroneous examples. This study focused on whether erroneous examples could encourage more generative processing than problem solving, even though both conditions encourage at least some problem solving (for erroneous examples: finding and fixing errors). The two groups were presented with isomorphic problems, but with different ways of interacting with those problems. The erroneous examples subjects were presented with an incorrect solution, were prompted to explain and correct the error and reflect on the correct answer, and received feedback on their responses. The problem solving subjects were asked to solve the same problems, reflect on the correct answers, and received feedback on their work. The additional steps in the erroneous examples condition of explaining and correcting the error/misconception made in each problem was intended to improve learning outcomes by encouraging learners to engage in generative processing concerning decimal principles. The problems were also simplified from Isotani et al. (2011) by providing more complete explanations for the students to choose from. Previous research on self-explanation prompts by Johnson and Mayer (2010), demonstrated that providing the explanation statements, rather than having learners generate their own, facilitated learning from an educational game. By providing the students with possible complete explanations to choose from rather than parts of sentences, processing demands should decrease.

Participants. Participants consisted of 208 (Male = 101, Female = 107) middle-school students from Pittsburgh, PA. Of those students, 105 were in the 6th grade while 103 were in 7th grade. Ages ranged from 11 to 13 (M= 11.99, SD = .722).

Materials. The computer-based materials consisted of 6 components, three tests (pretest, posttest, and delayed posttest), two surveys (demographic/math experience and evaluation), and the intervention problems. For the pretest and two posttests, three separate but isomorphic tests were constructed. Question types including placing decimals on a number line, putting a group of three or four decimal numbers in order, providing the next two numbers in a sequence, and answering true/false statements. All three tests contained 46 problems with a total of 50 points possible. For the demographic survey, along with basic information about age and grade level, students were asked about their experience with decimals and computers. They were also asked a few self-efficacy questions such as, “I am good in math at school”, with 5-point Likert answers, ranging from “Strongly Agree” to “Strongly Disagree.” For the evaluation survey, students were

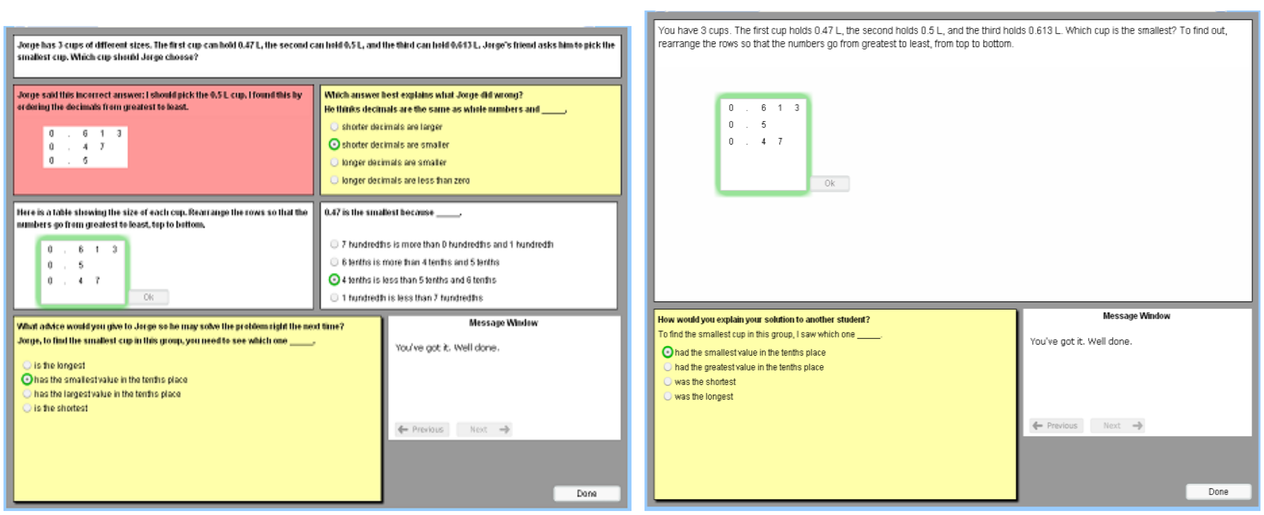

Figure 1: Side-by-side comparison of the isomorphic questions from the two intervention conditions. An erroneous example problem is on the left and the equivalent problem to solve is on the right.

asked how they felt about the intervention using a 5-point Likert scale ranging form. Questions included items such as, “I would like to do more lessons like this.”

During the intervention students completed a total of 36 problems, with interaction and feedback implemented by intelligent tutoring software (Aleven et al, 2009). The problems were arranged in four groups of three (with each group targeted at one of four misconception types) making a total of 12 groups. For the erroneous condition, students would first receive two problems dealing with a misconception such as “shorter decimals are smaller.” The third problem was then a problem to solve (with feedback) related to the misconception (i.e. putting decimals of different lengths in order from largest to smallest). The erroneous problems contained up to 5 components (not including the problem statement) for the students to interact with (see Figure 1 for a comparison between the two interventions). In the top left box students read the error made by the individual in the word problem. After pressing a “Next” button students were asked to identify what the subject had done wrong from a list of 3-4 options, one of which was the misconception exhibited by that student. In the left middle panel students were then asked to correct the mistake. This involved either placing the decimal correctly on a number line, changing a decimal addition, correctly ordering a list of decimals (largest to smallest or smallest to largest), or correctly completing a sequence of decimals. In the right middle panel participants explained why the new answer was correct. Finally, in the bottom left panel the students were asked to give advice to the fictional student that had gotten the answer incorrect. For every panel that required the student to make a selection feedback was provided (green = correct; red = incorrect). Students also received text feedback from a message window that was placed at the bottom right corner of the intervention. Messages include encouragement for students to try incorrect steps again or feedback for students to continue on to the next step or problem.

In the problem-solving version of the intervention, students were given the same problems as in the erroneous examples condition except they were asked to provide the solutions themselves. These problems were also arranged in groups of three with a simple correct / incorrect feedback for the third problem in each sequence. On the first two problems of the problem solving condition, after solving the problem students were asked how they would explain their solution to another student. These options included the correct procedure along with misconception distracters. Students in this group also received feedback from a message window in the bottom right panel as well as green / red feedback on their solution and multiple-choice selections.

Procedure. Students were randomly assigned to one of the two conditions **(**PS = 108; ErrEx= 100). Students in both conditions were given a total of five 43-minute sessions to complete the entire intervention. The students were randomly assigned to either the problem solving or the erroneous worked example condition. Students were also randomly assigned to receive one of the six possible pretest / posttest / delay-posttest orderings (ABC, ACB, BAC, BCA, CAB, CBA). On the second day the students answered the demographic and math/computer experience questionnaire before starting the intervention. The students were given two days to complete the problem solving/worked example problems. Upon completion they were given the intervention assessment questionnaire. The next day students were given the immediate posttest. Finally, during the following week, students were given the delayed posttest.

Results

Due to an error in data recording for four of the problems, the data for those problems was removed from the pretest, posttest, and delayed posttest scores making the total possible score out of 46. To first examine whether the problem solving (PS) and erroneous examples (ErrEx) condition performed similarly on the pretest an independent sample t-test was conducted. It was found that the ErrEx group performed significantly better on the pretest than the PS group, t (206) = 3.045, p = .003 (See Table 1 for means and standard deviations). An ANOVA revealed that there was no significant difference between the test orders, F (5,202) = 1.293, MSE = .057, p = .268.

In general students significantly improved their test performance after the intervention, regardless of condition, t(207) = -8.058, p < .001, with a mean increase of 9%. Students continued to significantly improve between the immediate and delayed posttest, t(207) = -8.230, p < .001, with a mean increase of 6%. Overall students increased their performance an average of 15% between the pretest and the delayed posttest, yielding a medium-to-large effect, d = .75.

To examine whether one condition increased learning more than the other gain scores were calculated between the pretest and posttest, pretest and delayed posttest, and posttest and delayed posttest. An ANCOVA with pretest as a covariate revealed that for the pretest-to-immediate-posttest gain did not differ significantly for the two groups, F (1,205) = .768, MSE = 34.97, p = .382. There were significant differences between the two conditions for pretest-to-delayed-posttest gains, F (1,205) = 9.896, MSE = 349.08, p = .002, and between immediate posttest and delayed posttest, F (1,205) = 7.027, MSE = 163.07, p = .009, with participants in the ErrEx condition having higher gain scores. That is, although participants

Table 1: Test performance for the two conditions

	Groups	
	Problem Solving	Erroneous Examples
Test score	*M* *SD*	*M* *SD*
Pretest	24.68 (9.42)	28.69 (9.58)
Posttest	29.07 (9.48)	32.58 (8.95)
Delayed Posttest	31.06 (9.20)	36.23 (7.47)

in the ErrEx condition may not have scored higher on the immediate posttest, they showed superior gains when tested after the week delay.

To determine whether the intervention had a different effect for students with high prior knowledge versus those with low prior knowledge, similar to Grosse and Renkl's (2007),we conducted an additional analysis. Participants were first classified as high verses low by using a median split on the pretest scores (8-25 points for low and 26-45 points for high). This divided the groups so that there were 107 students classified as low prior knowledge and 101 as high prior knowledge. For high prior knowledge individuals an ANCOVA with pretest as a covariate revealed the participants did not differ for pretest to immediate posttest gains, F (1,98) = .122, MSE = 3.76, p = .728, or posttest test to delayed posttest gains, F (1,98) = 2.01, MSE = 46.15, p = .160 (see Table 2 for means and standard deviations). There was a significant difference with ErrEx showing greater gains between the pretest and delayed posttest, F (1,98) = 4.75, MSE = 76.27, p = .032. For low prior knowledge individuals there was still not significant difference between pretest and posttest gains, F (1,104) = .489, MSE = 28.49, p = .486. However there the ErrEx condition did have significantly higher gains between the pretest and delayed posttest, F (1,104) = 5.21, MSE = 265.73 p = .025, and the posttest and delayed posttest, F (1,104) = 5.02, MSE = 120.21, p = .027. Thus, the pretest-to-delayed posttest gain was greater for the ErrEx condition for both low and high prior knowledge learners.

Discussion

The results of this study show that although using erroneous examples did not facilitate learning gains for an immediate pretest, students in the erroneous group had significantly higher gains on the delayed posttest. These results suggest that students taught with erroneous examples may have had a deeper learning experience, one that helped them build upon their initial understanding of decimals to gain a deeper understanding by the time they took the delayed posttest.

Previous research by Grosse and Renkl's (2007) found that prior knowledge interacted with incorrect examples; higher prior knowledge students performed better when presented with incorrect solutions. For our study, however, no significant interaction was found between prior knowledge and condition. The data showed that both low and high prior knowledge individuals did better in the erroneous examples condition than the problem solving condition. This might have occurred because the erroneous example students, both low and high prior knowledge, were enticed to engage in more generative processing than the problem solving students, through the prompted explanation and correction of errors.

One limitation of our study is that we did not include a correct worked examples condition. The reasons for this were straightforward. First, in the present study 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, as we revised the instructional materials from the Isotani et al (2011) study, we realized that erroneous examples and problem solving were 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 does not require. Renkl and Atkinson (2010) mention a reversal of the worked examples effect when students already have sufficient knowledge. Studying just the examples without any sort of active problem solving may become redundant for the students therefore decreasing the amount of mental effort they put into the lesson. Nevertheless, to compare other possible instructional approaches, in a future study we intend to include a worked examples condition.

Table 2: Test performance for low/high prior knowledge individuals for the two conditions

	Low Prior Knowledge		High Prior Knowledge	
	PS	ErrEx	PS	ErrEx
Test Score	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Pretest	17.73 (3.89)	19.25 (4.01)	34.40 (5.36)	36.11 (5.03)
Posttest	24.24 (8.82)	26.80 (8.06)	35.84 (5.33)	37.13 (6.75)
Delayed Posttest	26.30 (8.55)	31.07 (7.47)	37.71 (5.96)	40.29 (4.34)

In summary, our study provides evidence that presenting students with errors that they are prompted to analyze, explain, and correct can facilitate learning decimals from a computer-based tutor.

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

Aleven, V., McLaren, B.M., Sewall, J., & Koedinger, K.R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, *19*(2), 105-154.

Chi, M.T.H., Bassok, M., Lewis, M.W., & Reimann, R. & Glaser, R. (1989). Self explanations: How students study and used examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79,* 347-362.

Grosse, C.S. & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17*(6), 612-634.

Hausmann, R.G.H. & Chi, M.T.H. (2002). Can a computer interface support self-explaining? *Cognitive Technology*, *7*(1),4–14.

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? In: Proceedings of the *Sixth European Conference on Technology Enhanced Learning: Towards Ubiquitous Learning (EC-TEL-2011)*.

Johnson, C.I. & Mayer, R.E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior*, *26*, 1246-1252.

Kawasaki, M. (2010. Learning to solve mathematics problems: The impact of incorrect solutions in fifth grade peers' presentations. *Japanese Journal of Developmental Psychology, 21* (1), 12-22.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed). New York: Cambridge University Press.

McLaren, B.M., Lim, S., & Koedinger, K.R. (2008). When and how often should worked examples be given to students? New results and a summary of the current state of research. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society,*. Austin, TX: Cognitive Science Society.

Moreno, R., & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In. J.L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory*, Cambridge: Cambridge University Press.

Reiss, K., Hellmich, F., & Thomas, J. (2002). Individual and scholastic factors for argumentations and proofs in mathematics instruction]. In M. Prenzel, & J. Doll (Eds.), *Bildungsqualitä t von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. Zeitschrift für Pädagogik (45. Beiheft*)

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21,* 1-29.

Renkl, A. (2005). The worked-out examples principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Renkl, A. (2010). Instruction based on examples. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction*. New York: Routledge.

Renkl, A., & Atkinson. R.K. (2010). Learning from worked-out examples and problem solving. In. J.L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory,*, Cambridge: Cambridge University Press.

Tsovaltzi, D., Melis, E., McLaren, B.M., Meyer, A-K., Dietrich, M. & Goguadze, G. (2010). G. Learning from erroneous examples: When and how do students benefit from them? In Proceedings of the *European Conference on Technology Enhanced Learning,* LNCS vol. 6383,. Springer, Heidelberg.

Zhu, X., and H.A. Simon. 1987. Learning mathematics from examples and by doing. *Cognition and Instruction 4* (3),137–66.