THE EFFECTIVENESS OF COMPARING CORRECT AND INCORRECT EXAMPLES FOR LEARNING ABOUT DECIMAL MAGNITUDE

By

Kelley Durkin

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

December, 2009

Nashville, Tennessee

Approved:

Professor Bethany Rittle-Johnson

Professor Georgene Troseth

ACKNOWLEDGEMENTS

This research was supported with funding from the Institute of Education Sciences, U.S. Department of Education, grants R305H050179 and R305B040110. The opinions expressed are mine and do not represent views of the U.S. Department of Education. A special thanks to the students and teachers at Overbrook School, Saint Bernard Academy, and St. Ann School for participating in this research. I am especially grateful to my advisor, Bethany Rittle-Johnson, and my committee member, Georgene Troseth, who offered endless guidance and support. Thanks to Jon Star for suggestions on the study design and to Holly Harris, Kristin Tremblay, and Calie Traver for help in collecting and coding the data. Finally, thanks to my family and friends who have been with me every step of the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Students face misconceptions when learning in a variety of domains. For example, students often think that the seasons are caused by earth's changing distance from the sun (Mestre, 1994) and elementary-school children often think that 0.25 is greater than 0.8 because 25 is greater than 8 (Resnick et al., 1989). Misconceptions such as these can persist over a long period of time, and in order to master a domain, these misconceptions need to be overcome (Eryilmaz, 2002). In the current study, we examined whether comparison of correct and incorrect examples could help alleviate misconceptions and aid students learning about decimal fractions. In the introduction, we outline the benefits of using incorrect examples, the potential role of comparison in learning from incorrect examples, and common misconceptions in the target domain of decimal fractions.

Benefits of Incorrect Examples

People's intuition and original behaviorist principles suggest that exposure to incorrect examples may reinforce incorrect responses, and therefore should not be used. However, research indicates that presenting students with examples of their misconceptions can be beneficial for correcting them and for improving knowledge of concepts (Eryilmaz, 2002; Huang, Liu, & Shiu, 2008; Van den Broek & Kendeou, 2008) and procedures (Große & Renkl, 2007; Siegler, 2002).

Presenting students with incorrect examples may allow them to recognize misconceptions and consequently improve their knowledge of correct concepts. This has been best studied in the physics learning literature. Directly exposing students to physics misconceptions was more effective for helping students overcome their misconceptions than focusing exclusively on correct concepts (Eryilmaz, 2002; Mestre, 1994). Science texts in many domains have sometimes capitalized on this idea with refutation texts (Alvermann & Hague, 1989; Diakidoy, Kendeou, & Ioannides, 2003; Van den Broek & Kendeou, 2008). These refutation texts place incorrect examples in textbooks along with correct examples to teach scientific concepts (although the examples are not directly compared). Students who used such refutation texts were more likely to revise incorrect knowledge than students who did not see any incorrect examples in their text (Van den Broek & Kendeou, 2008). By engaging students' conflicting correct and incorrect concepts, students may create a new mental representation of the material that labels incorrect concepts as wrong (Van den Broek & Kendeou, 2008). In addition, using incorrect examples may motivate students to think more deeply about the correct concepts (VanLehn, 1999). Thus, presenting incorrect examples may help students recognize their misconceptions and focus their attention on understanding the correct concepts.

A vast majority of the literature on incorrect examples has focused on scientific misconception, but there is limited evidence that exposing students to incorrect mathematics examples can help students overcome their misconceptions and increase their knowledge of math concepts. In one study, sixth-grade students were exposed to incorrect examples when learning about the meaning of decimals (e.g. in 5.4, saying the

.4 represents 4 ones instead of 4 tenths). These students retained correct concepts after a 4 week delay better than students who were not presented with incorrect examples (Huang et al., 2008).

Exposure to incorrect examples may also benefit students by reconciling competing procedures and improving the likelihood that correct procedures will be used. People use multiple procedures and ways of thinking at any given time, and often people continue to use incorrect procedures after correct ones have been learned (see Siegler, 2002 for a review). For example, the misconception that the equal sign means "add up the numbers" is pervasive in elementary school, and it leads many children to use incorrect procedures, such as adding all the numbers present in the equation, when solving math equivalence problems such as $3 + 4 + 5 = \_\_ + 5$ (Carpenter, Franke, & Levi, 2003; McNeil, 2008; Perry, 1991). Even after children learn correct procedures for solving math equivalence problems, they continue to use their old, incorrect procedures (e.g., Alibali, 1999; Rittle-Johnson & Alibali, 1999). Having students explain why incorrect procedures are wrong is one way to reduce their use of incorrect procedures. Children who explained both correct and incorrect solutions were able to solve more difficult problems than those who only explained correct solutions because they were more likely to learn and use correct procedures that were applicable to a range of problem types (Siegler, 2002). Similar results have been found for college students learning algebra and probability (Curry, 2004; Große & Renkl, 2007). Siegler (2002) proposed that explaining incorrect examples helps decrease the strength of incorrect procedures, reducing the probability that the procedure will be selected in the future. This may occur in part because studying incorrect procedures can lead people to verbalize more thoughts, including explanations of why the procedure is wrong and elaborations of correct procedures (Große & Renkl, 2007). However, the benefits of studying incorrect examples may only arise for learners with sufficient prior knowledge to generate reasonable explanations (Große & Renkl, 2007). Although not always effective, presenting people with examples of misconceptions and incorrect procedures often helps them recognize their incorrect ways of thinking and learn correct concepts and procedures.

Comparison of Incorrect and Correct Examples

Currently, it is unclear what role comparison plays in learning from incorrect examples. It seems likely that comparison of incorrect examples to correct ones is a powerful learning process. For example, if incorrect examples promote deeper reflection on correct concepts (VanLehn, 1999) and change the relative probability of selecting correct over incorrect procedures (Siegler, 2002), incorrect examples are likely being compared to correct ones. However, the role of comparison has never been tested in studies on incorrect examples. We briefly review the design of past studies on incorrect examples and relate these design features to what is known about how to support learning from comparison of correct examples.

In most past research, incorrect examples were presented alone and were not directly compared to correct examples (Große & Renkl, 2007; Huang et al., 2008). Students may have spontaneously compared the incorrect examples to correct examples, but doing so would require good metacognitive skills (recognizing the benefits of doing so) and impose high-working memory load (the need to recall a correct example and keep it in mind during the comparison process) (Richland, Morrison, & Holyoak, 2006).

Indeed, when studying multiple *correct* examples, sequential presentation of examples does not support learning as well as simultaneous presentation of the examples, in part because simultaneous presentation greatly increases the frequency of comparison (Gentner, Loewenstein, & Thompson, 2003; Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2009). We could only identify one prior study on mathematics learning in which an incorrect solution was presented at the same time as a correct solution (Siegler, 2002), as well as a few studies on refutation science texts in which incorrect examples were presented on the same page as correct examples (Alvermann & Hague, 1989; Diakidoy et al., 2003; Van den Broek & Kendeou, 2008). However, in none of these studies were students prompted to compare the two – rather, they were prompted to think about each example individually. Explicit and focused prompts to compare greatly enhance the benefits of studying multiple correct examples (Catrambone & Holyoak, 1989; Gentner et al., 2003). We suspect the same is true for incorrect examples.

Given the potentially important role of comparison in learning from incorrect examples, we were interested in the benefits of learners comparing correct and incorrect examples, over and above the general benefits of comparison. Comparison is often lauded as an effective and important learning process in both cognitive science (Gentner et al., 2003; Gick & Holyoak, 1983) and in mathematics education (NCTM, 2000). Comparing two correct examples can help people create correct knowledge categories, increase knowledge of concepts, and increase flexible use of procedures (e.g., Kotovsky & Gentner, 1996; Rittle-Johnson & Star, 2009). For example, comparing two correct solution procedures improved students' learning of equation-solving and estimation procedures more than examining the same examples one at a time (Rittle-Johnson & Star, 2007; Rittle-Johnson, Star, & Durkin, in press; Star & Rittle-Johnson, 2009). In the domain of estimation, comparison also improved students' conceptual knowledge more than viewing examples sequentially (Star & Rittle-Johnson, 2009). Comparing correct examples may also reduce misconceptions, but this possibility has not been tested. Because comparison of two correct examples is an effective route to establishing correct concepts and procedures, it should also reduce misconceptions in conflict with them.

The Target Domain and Outcomes

To evaluate the effectiveness of comparing incorrect examples to correct ones, we chose an important mathematical domain with many documented misconceptions. The National Mathematics Advisory Panel Report (2008) emphasized the importance of students mastering decimal fractions, commonly referred to as decimals (Resnick et al., 1989). However, the Panel also reported that in general, students receive very poor preparation in decimals. Algebra teachers surveyed by the Panel identified their students' lack of knowledge about fractions and decimals as the second largest barrier to learning algebra, while also listing fraction and decimal knowledge as the third most important skill students need to successfully learn algebra (National Mathematics Advisory Panel Subcommittee, 2008). Indeed, it is well documented that students often have difficulty understanding decimals (Glasgow, Ragan, Fields, Reys, & Wasman, 2000; Kouba, Carpenter, & Swafford, 1989; Rittle-Johnson, Siegler, & Alibali, 2001), and in fact, adults frequently have trouble with decimals as well (Putt, 1995; Stacey et al., 2001).

One reason for these difficulties is that students have common and persistent misconceptions involving decimal magnitude (Glasgow et al., 2000; Irwin, 2001; Resnick et al., 1989; Sackur-Grisvard & Leonard, 1985). 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). A related misconception is specific to the role of zero, which is different for whole and decimal numbers. When a zero is in the tenths place, students will often ignore it and treat the following digit as if it is in the tenths place (e.g. students will think 0.08 is the same as 0.8). Students will also frequently assume that adding a zero on the end of a number increases its magnitude (e.g. 0.320 is greater than 0.32). Again, students are incorrectly applying knowledge about whole numbers to decimals (e.g., 08 is the same as 8 and 320 is greater than 32). Finally, some students think that all decimals less than one are less than zero because they start with "0" (Glasgow et al., 2000; Irwin, 2001; Resnick et al., 1989).

These misconceptions interfere with conceptual understanding of decimals and manifest themselves in various ways. Without an early knowledge of decimal concepts, students have difficulty doing later mathematical tasks involving decimals (Hiebert & Wearne, 1985). For example, when asked to add or subtract two decimals, students often do not know how to align the numbers properly. This seems to be a result of students relying on a set of learned procedures without a solid conceptual foundation (Hiebert & Wearne, 1985). The persistent misconceptions evident in students' decimal knowledge must be overcome so that students can become proficient with decimals, as well as go on to treat decimal numbers appropriately in more advanced mathematics.

Number line tasks can serve as a useful way to teach students about decimal magnitude (National Mathematics Advisory Panel, 2008; Rittle-Johnson et al., 2001). It is a relevant but novel task for most students. Placing fractions and decimals on a number line can help students better understand magnitude relations and link knowledge of procedures and concepts (National Mathematics Advisory Panel, 2008). Therefore, we focused on children learning to place decimals on number lines, a task that should engage their concepts of decimal magnitude as well as support generation and use of procedures for placing values on number lines.

Current Study

In the current study, students either compared two correct procedures for solving a problem or compared an incorrect procedure based on a common misconception to a correct one. All students studied worked examples of number line problems and were prompted to self-explain because the combination of worked examples and self-explanation improves learning in a large variety of domains (see Atkinson, Renkl, & Merrill, 2003 for a review). We also included practice problems on which students received immediate feedback, since such practice is important when learning from worked examples (Atkinson, Derry, Renkl, & Wortham, 2000). Before and after the intervention, we assessed students' knowledge of related concepts and procedures, as well as their misconceptions of decimal magnitude. In addition, we explored students' explanation quality to gain insights into how different types of comparison impacted learning.

We hypothesized that although both groups would show learning gains, students who compared correct and incorrect examples would reduce their misconceptions more, would perform better on conceptual and procedural measures and would retain knowledge better over a delay than students who were prompted to compare two correct examples. We made this prediction based on evidence for the general benefits of exposing students to incorrect examples (e.g., Eryilmaz, 2002; Huang et al., 2008) and the expectation that incorrect examples would help flag misconceptions as errors (Van den Broek & Kendeou, 2008), would increase attention to correct concepts (VanLehn, 1999), and would increase the likelihood that correct procedures would be selected over incorrect ones (Siegler, 2002).

CHAPTER II

METHOD

Participants

Consent was attained from 103 fourth- and fifth-grade students from three urban parochial schools (59 female, 44 male). Students solving 75% or more of the learning problems correctly at pretest were excluded from this study because they had already mastered the skills that would be taught in the intervention.

The final sample consisted of 74 students (48 female, 26 male): 43 from fourth-grade classes and 31 from fifth-grade classes. The average age was 10.5 years (range 7.0 to 12.7 years). At one school, four of the students were actually third graders working at a fourth-grade level. The fourth-grade classes used the Saxon Math or Sadlier-Oxford Math textbook, and the fifth grade classes used either the Saxon Math or Harcourt Math textbook. Most students had not yet received much formal instruction in decimals, although some fifth-grade students had had some instruction.

Design

Students participated in a pretest-intervention-posttest design, including a two-week retention test. Students were randomly assigned to one of two intervention conditions: the *incorrect* condition (n = 37) or the *correct* condition (n = 37). In both conditions, students received the same introductory lesson on decimals and then completed a packet with 12 pairs of worked examples. Each pair illustrated two different procedures for placing a decimal value on a number line from 0 to 1, followed by questions prompting them to reflect on the two examples (see Figure 1 for examples). The students in the *incorrect* condition compared one correct and one incorrect example in each pair, while students in the *correct* condition compared two different correct examples in each pair. During the intervention, students also solved four practice problems and received accuracy feedback. After completing the intervention, students completed an immediate posttest. Approximately two weeks later, students completed a retention test.

Materials

Introductory Lesson

All students received a five minute lesson on decimals and place value adapted from a lesson in the *Everyday Mathematics* curriculum (Bell et al., 2004). The purpose of this lesson was to help students become acquainted with terminology (e.g. tenths, hundredths, and thousandths) and understand what different place values meant (e.g. that the value in the hundredths place is ten times greater than the value in the thousandths place).

Intervention Packet

Each packet contained 24 worked examples of decimal number line problems and 24 corresponding questions, presented in pairs. Each worked example showed where a hypothetical student placed a decimal on a number line from 0 to 1 and an explanation of his or her procedure (See Figure 1). The decimals varied in number of digits (tenths, hundredths or thousandths), magnitude (greater than or less than 0.5), and whether a zero was included. On the first 12 worked examples, the tenths were marked on the number line, since tenths marks help students learn about decimal magnitude (Rittle-Johnson et al., 2001). On the remaining 12 worked examples, the tenths were not marked since fading of instructional supports has been shown to improve the robustness of student learning (Atkinson et al., 2003).

Three different correct procedures were illustrated across the worked examples, based on procedures students have reported using in past research (Irwin, 2001; Rittle-Johnson et al., 2001). The first focused on the number of tenths in a decimal, placing the value close to where that number of tenths would go (e.g., For 0.312, estimating where 3 tenths would go, and then moving over a little more). A second was imaging the line divided into the number of pieces specified by the smallest place value and placing the decimals based on this (e.g. For 0.312, imagine dividing the line into 1000 pieces and estimating where 312 would be placed). The third was a benchmark procedure in which the location was estimated based on knowledge of its magnitude in relation to 0, 0.5, and 1.

Three different incorrect procedures were illustrated in the incorrect condition. Each corresponded to common decimal misconceptions (Irwin, 2001; Resnick et al., 1989). One procedure was to treat decimals like whole numbers (e.g. thinking that 0.9 is like 9, a small number, and thus placing 0.9 close to 0). A second procedure was to think of all decimals as less than zero and place them before zero. The third procedure incorporated a misconception about the role of zero. Specifically, this procedure was to ignore zeros in the tenths place and to think that zeros placed at the end of a decimal made its value bigger.

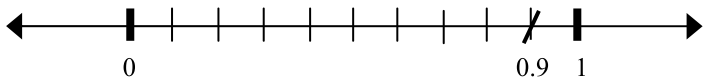
In the packet for the *incorrect* condition, one correct and one incorrect procedure was illustrated on each page, with the three correct and three incorrect procedures evenly distributed throughout the packet. Each page of the packet also contained two questions prompting students to explain why an example was correct or incorrect, explain why a decimal should be placed closer to 0 or 1, and/or describe an additional correct procedure for solving the problem.

In the packet for the *correct* condition, two different correct procedures were illustrated on each page, with the three correct procedures evenly distributed throughout the packet. The two questions on each page prompted students to explain why each example was correct, explain why a decimal should be placed closer to 0 or 1, describe an additional correct procedure for solving the problem, and./or describe one of the correct procedures to a new student.

After every three pairs of worked examples, a practice problem was presented. Students were asked to place a slash on the number line where a decimal would be located (i.e., 0.164, 0.89, 0.8, and 0.310) and to justify their answer. Following their justification, students were given feedback as to whether their placement was correct and to where the slash should have been on the number line. On the first two problems, the tenths were marked, and on the second two problems, they were not.

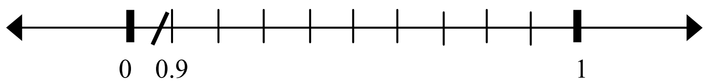
Incorrect Condition

Correct



John said, “9 tenths is 9 out of 10 tenths. Because the line is divided into 10 tenths, I counted over to 9 tenths.”

Incorrect



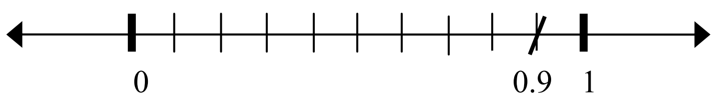
Tyler said, "9 is a small number. So I’m going to put 0.9 close to 0.”

1. Why is Tyler’s way of thinking incorrect?

2. Explain why John’s way of thinking is correct.

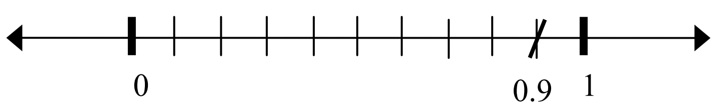
Correct Condition

Correct



John said, “9 tenths is 9 out of 10 tenths. Because the line is divided into 10 tenths, I counted over to 9 tenths.”

Correct



Kim said, "I know 9 tenths is only one tenth smaller than 1. So I marked 0.9 a little before 1."

1. How are Kim’s and John’s ways different?

2. Explain why both ways of thinking are correct.

Figure 1: Sample Intervention Packet Page for Each Condition

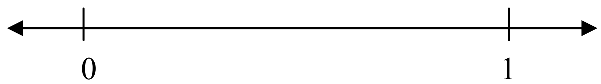
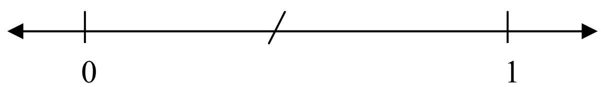
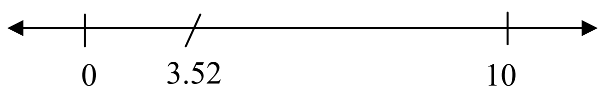
Assessment

The same assessment was used as a pretest, posttest, and retention test. This assessment was designed to measure knowledge of procedures and concepts and was adapted from the one used in Rittle-Johnson et al. (2001). Sample items of each knowledge type are shown in Table 1. The decimals varied in number of digits (tenths, hundredths, and thousandths), magnitude (greater than or less than 0.5), and whether a zero was included. The procedural items were four familiar, *learning problems* and eight novel, *transfer problems*. Target values were specifically chosen to make it easy to recognize misconception errors. On the *learning problems*, students needed to place a decimal on a number line from 0 to 1, similar to the problems they completed during the intervention. On one kind of *transfer problem*, students needed to identify from a list which decimal was already marked on a number line. Thus, students were asked to do the reverse of what they had done during the intervention. The second kind of *transfer problem* involved number lines from zero to ten, and students needed to place a decimal greater than 1 in relation to another number marked on the number line (e.g. place 3.8 in relation to 3.52). Internal consistency on this measure was good ($\alpha = .79$).

The conceptual items were designed to measure students' understanding of fundamental decimal concepts and the four types of items were based on past assessments (see Table 1; Irwin, 2001; Rittle-Johnson et al., 2001). *Magnitude comparison* items assessed students' understanding of the size of various decimals (Irwin, 2001; Resnick et al., 1989). The *continuous nature of decimals* items evaluated students' understanding that there are an infinite number of decimals that can come between any two numbers (e.g. between 0.76 and 0.77) (Irwin, 2001; Resnick et al., 1989; Rittle-

Johnson et al., 2001). The *role of zero* items evaluated students' understanding of when a zero made a difference in a decimal's magnitude (Irwin, 2001; Rittle-Johnson et al., 2001). Finally, the *greater than zero* items assessed students' knowledge that decimals presented were greater than zero (Irwin, 2001). Each item was designed to contain an answer choice that fit a misconception error. Internal consistency on this measure was very high ($\alpha = .91$).

Table 1: Sample Assessment Items

<table>
<tr><th></th><th>Example Item</th><th>Scoring</th></tr>
<tr><td>Procedural Items
(learning)
(n = 4)</td><td>Mark about where 0.9 goes on the number line.
</td><td>1 point for each if within one tenth of the correct placement in either direction on the number line</td></tr>
<tr><td>Procedural Items
(transfer)
(n = 4)</td><td>What number tells about where the slash is on the number line?
a) 0.214
b) 0.84
c) 0.489
d) 0.05
</td><td>1 point for each correct answer circled</td></tr>
<tr><td>Procedural Items
(transfer)
(n = 4)</td><td>The number line now goes from 0 to 10. 3.52 is marked. Mark where 3.8 goes.
</td><td>1 point for each if properly placed as greater or less than the marked number and within one tenth of the correct placement</td></tr>
<tr><td>Conceptual Items
(n = 20)</td><td>1. (Magnitude, n = 9) Circle the decimal that is greater: 0.87 0.835
2. (Continuous Nature, n = 5) Write a decimal that comes between 0.5 and 0.6.
3. (Role of zero, n = 4) Circle all the numbers that are worth the same amount as 0.51:
0.5100 0.051 0.510 51
4. (Greater than zero, n = 2) 0.8 is _______ 0
a) greater than
b) less than
c) the same as</td><td>1 point for each correct answer

1 Magnitude item and 1 Role of zero item had 2 correct answers, each worth 1 point.</td></tr>
</table>

Procedure

Students completed the pretest as a group in a 20 minute session in their classroom. Then, each student was pulled out of class individually for the intervention. The intervention began with the five minute scripted lesson. Following the lesson, students completed the intervention packet, which took about 25 minutes. Students began each page by reading the example procedures aloud. Next, the researcher asked the student two questions that the student answered verbally. The four practice problems were completed when appropriate. Students were audio taped and videotaped so that their explanations could be recorded and later transcribed. Immediately after the intervention, students were given the posttest. Approximately two weeks later, all students completed the retention test together in their classrooms.

Coding

Assessment

Items were scored for accuracy using the criteria specified in Table 1. When possible, students' answers on the assessment were also coded for the three misconception errors described in the introduction. For the learning number line problems, an incorrect answer was considered a whole number misconception if a number in the tenths was incorrectly placed on the first third of the number line (e.g., 0.9 placed at 0.1), a number in the hundredths was placed on the second third (e.g., 0.83 placed at 0.5), or a number in the thousandths was placed in the final third (e.g., 0.256 was placed near 1). Recall that target values on the assessment were chosen so that misconceptions could not lead to a correct answer (e.g., 0.2 was not included). For the whole-number transfer problems, an incorrect answer was considered a whole number misconception if a number was placed within one unit on the wrong side of the already marked decimal (e.g. when 3.52 is marked, placing 3.8 between 2.52 and 3.52). An error was considered a role of zero misconception error if a decimal with a zero in the tenths was placed as if the zero was not there (e.g. placing 0.07 at 0.7). Students were marked as having a less than zero misconception if they placed a decimal anywhere before zero on the number line. For the conceptual continuous nature of decimal items, an incorrect answer was considered a whole number misconception if it fit between the given numbers viewed as whole numbers (e.g. 0.7 comes between 0.5 and 0.52 since 7 comes between 5 and 52). The other items were multiple choice, and distracter choices had been designed to reflect misconceptions when possible; students' errors on these items were scored for misconceptions accordingly. In all, there were 24 items on which a whole number misconception could be detected, 11 items on which a role of zero misconception could be detected, and 6 items on which a less than zero misconception error could be detected. The total number of misconception errors of each type was tallied for each student.

Intervention

Students' 24 verbal explanations given during the intervention were coded for discussing concepts and endorsing misconceptions. The details of the coding scheme will be discussed in the Results section. Inter-rater reliability for 20% of explanations ranged from 87% to 95% (exact agreement). Students' accuracy on the practice problems was coded in the same manner as for the procedural learning problems on the assessment.

Missing Data

Thirteen students were absent for the retention test. We imputed this missing data because imputation has been shown to result in the same conclusions as if there were no missing data when the data are missing at random and less than 20% of the data are missing (Barzi & Woodward, 2004; Schafer & Graham, 2002). To impute the missing data, we used the expectation-maximization algorithm for maximum-likelihood estimation via the missing value analysis module of SPSS (Schafer & Graham, 2002). The missing data were estimated from all nonmissing values on continuous variables in our models.

CHAPTER III

RESULTS

We first discuss students' performance on the pretest. We follow this with a report of the effects of condition on students' posttest performance and then on retention test performance. Finally, we explore how condition affected performance on intervention activities, including the quality of students' explanations, and how this related to performance on the assessments.

Pretest Knowledge

All students completed a pretest to assess their knowledge of procedures and concepts. Students had some knowledge of decimal magnitude at pretest (see Table 2), but it was not extensive. Students in both conditions were best at solving conceptual *greater than zero* items and were worst at solving procedural *transfer* items. There were no significant differences between conditions on conceptual or procedural items.

Misconception Errors

At pretest, misconception errors were prevalent for students in both conditions, and there were not significant differences between conditions (see Table 3). Students in both conditions made *whole number misconception* errors on about half of relevant items. *Role of zero misconception* errors were also prevalent. In contrast, *less than zero misconception* errors were made infrequently. For both conditions, 55% of the errors made on the pretest were misconception errors, with the remaining errors being random errors.

Table 2: Performance on Outcome Measures by Condition (Proportion Correct)

	Pretest		Posttest		Retention Test	
Outcome	*M*	*SD*	*M*	*SD*	*M*	*SD*
Conceptual Items						
Correct	.33	.20	.41	.26	.39	.27
Incorrect	.29	.20	.42	.28	.45	.29
Magnitude						
Correct	.28	.29	.34	.33	.36	.33
Incorrect	.21	.25	.36	.36	.39	.34
Continuous Nature						
Correct	.21	.21	.35	.30	.28	.32
Incorrect	.21	.19	.31	.25	.34	.31
Role of Zero						
Correct	.39	.37	.43	.40	.43	.40
Incorrect	.32	.35	.42	.39	.53	.40
Greater than Zero						
Correct	.70	.45	.89	.30	.71	.43
Incorrect	.80	.40	.96	.18	.79	.40
Procedural Items						
Correct	.21	.18	.43	.23	.38	.27
Incorrect	.20	.15	.51	.25	.41	.26
Learning						
Correct	.30	.18	.61	.22	.53	.29
Incorrect	.28	.16	.69	.22	.54	.29
Transfer						
Correct	.16	.22	.34	.28	.31	.31
Incorrect	.16	.18	.42	.30	.34	.30

Table 3: Proportion of Misconception Errors by Condition

	Pretest		Posttest		Retention Test	
Outcome	*M*	*SD*	*M*	*SD*	*M*	*SD*
Whole Number						
Correct	.45	.19	.37	.21	.38*	.24
Incorrect	.50	.20	.37	.21	.33	.18
Role of Zero						
Correct	.42	.23	.42	.22	.39	.23
Incorrect	.38	.17	.40	.27	.34	.24
Less than Zero						
Correct	.11	.17	.04	.10	.09	.14
Incorrect	.10	.19	.01	.06	.08	.13
Overall Misconception Errors						
Correct	.39	.15	.33	.17	.34*	.18
Incorrect	.41	.14	.32	.19	.30	.16

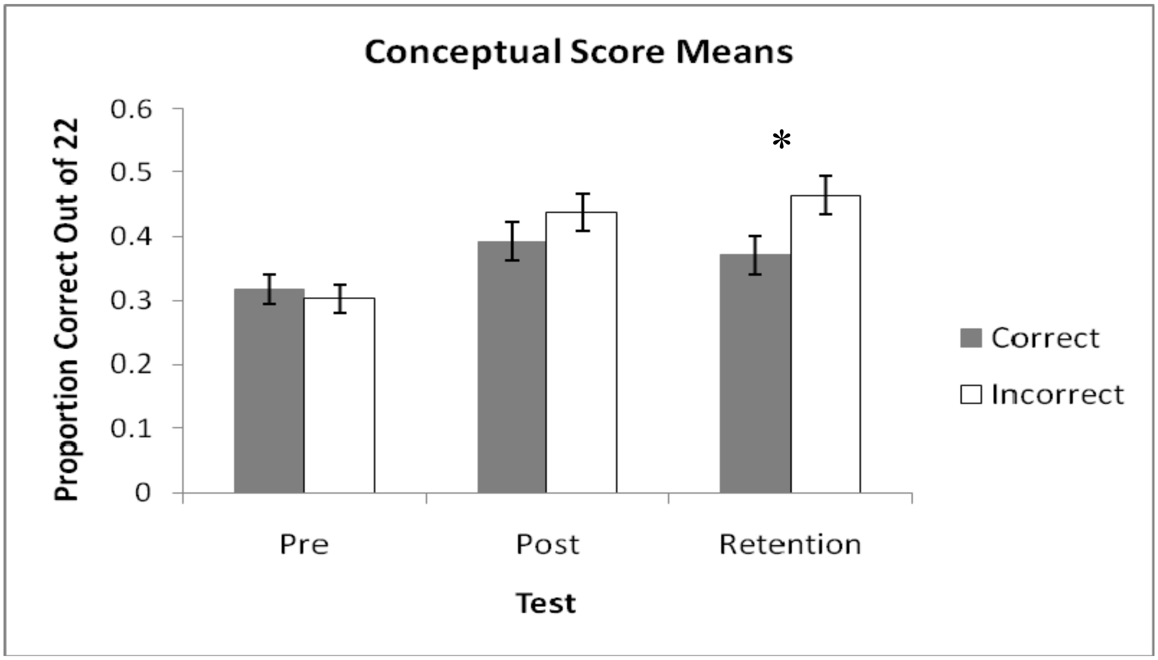
* Conditions differ from each other at $p < .05$

Effect of Condition on Posttest Knowledge

We analyzed differences between conditions on the posttest using analysis of covariance, with accuracy on the conceptual items and procedural items at pretest, grade level, and school as covariates. We ran separate models for each outcome.

Conceptual Items

There was no significant effect for condition on conceptual scores at posttest, $F(1, 67) = 1.21$, $p = .276$. As shown in Figure 2, students in both conditions had similar conceptual scores following the intervention, getting slightly fewer than 50% of the problems correct. Exploratory analyses indicated that there were not significant differences between the conditions on any of the different types of conceptual items.

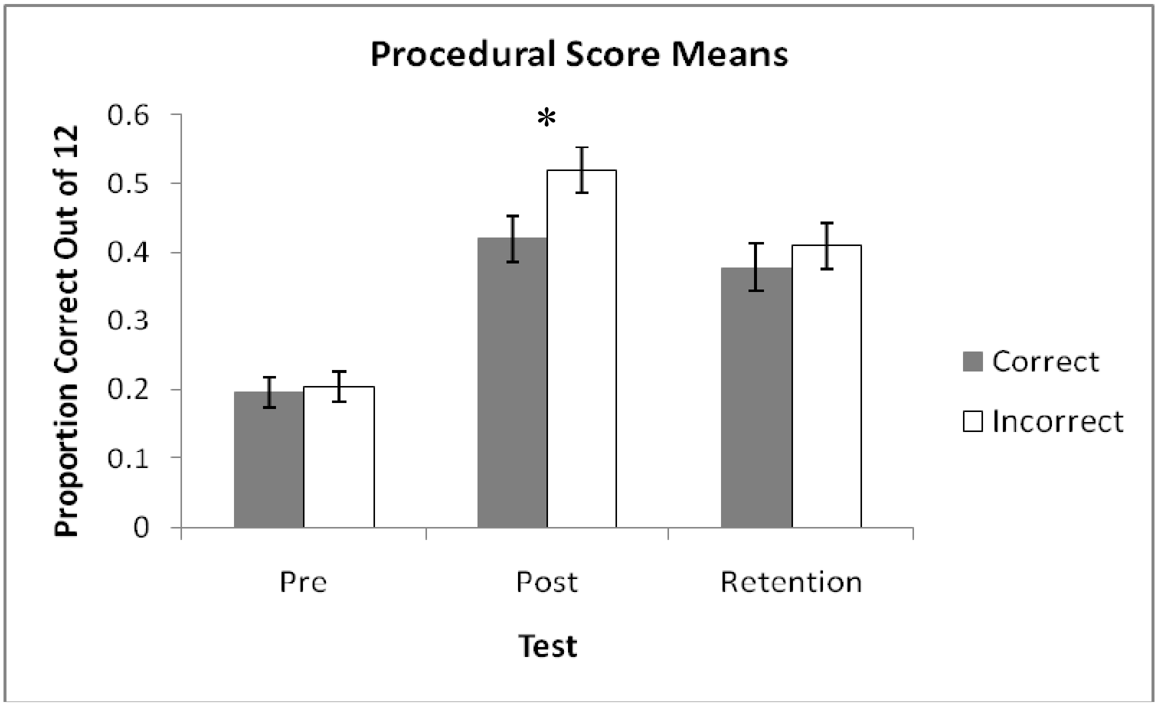


* Conditions differ from each other at $p < .05$

Figure 2: Effect of Condition on Conceptual Scores

Procedural Items

There was a significant effect for condition on procedural item scores at posttest, $F(1, 67) = 4.65$, $p = .035$. As shown in Figure 3, comparing incorrect and correct examples led students to have significantly higher procedural scores than those students who only compared correct examples. This difference was present for each type of procedural item.



* Conditions differ from each other at p < .05

Figure 3: Effect of Condition on Procedural Scores

Misconception Errors

As previously mentioned, we coded students' answers on the posttest for whether or not they fell into three common misconception errors: whole number (e.g. thinking 0.7 is less than 0.25 because 7 is less than 25), role of zero (e.g. treating 0.08 the same as 0.8), and less than zero misconceptions (e.g. thinking all decimals are less than 1). From pretest to posttest, overall misconception errors decreased significantly across conditions, $t(73) = 4.48$, $p < .001$ (see Table 3). This was mainly due to a decrease in whole number misconception errors from pretest to posttest, $t(73) = 4.77$, $p < .001$. Less than zero misconception errors also decreased from pretest to posttest, $t(73) = 3.66$, $p < .001$. However, role of zero misconception errors did not decrease, $t(73) = -.19$, $p = .850$. There were no significant differences between conditions in misconception errors at posttest. Rather, accuracy in the *incorrect* condition was higher because they made fewer random errors (22% vs. 29% of answers).

On the posttest, students in the *incorrect* condition had higher procedural scores than students in the *correct* condition. However, they did not have higher conceptual scores, and they made similar numbers of misconception errors.

Effect of Condition on Retention Test Knowledge

Conceptual Items

On the two week retention test, there was a significant effect for condition on conceptual item scores, $F(1, 67) = 5.27$, $p = .025$. Students who compared incorrect and correct examples had significantly higher conceptual scores than those students who only compared correct examples. As shown in Figure 2, from posttest to retention test, students who compared correct and incorrect examples maintained their conceptual scores, but students who compared correct examples had forgotten some information. Exploratory analyses indicated that the difference between conditions at retention test was greatest for items examining the role of zero, $F(1, 67) = 7.21$, $p = .009$ (see Table 2). Comparing incorrect examples to correct examples helped students remember the concepts they learned.

Procedural Items

At retention test, there was not a significant effect for condition on procedural item scores, F (1, 67) = 0.44, p = 0.511. As shown in Figure 3, students in the *incorrect* condition did not maintain their greater gains in procedural scores.

Misconception Errors

Across conditions, students made significantly fewer misconception errors from pretest to retention test, $t(73) = 5.54$, $p < .001$ (see Table 3). This was mostly due to a decrease in whole number misconception errors from pretest to retention test, $t(73) = 5.24$, $p < .001$. Furthermore, students in the *incorrect* condition made whole number misconception errors less frequently than students in the *correct* condition, F (1, 70) = 4.40, p = 0.040. The decrease in role of zero misconception errors was only marginally significant, $t(73) = 1.77$, $p = .081$, and there was not a significant decrease in less than zero misconception errors, $t(73) = .96$, $p = .340$. Although students in the *incorrect* condition made these errors less frequently than those in the *correct* condition at retention test, these differences were not significant for these two error types.

A greater reduction in misconception errors helped to explain why students in the *incorrect* condition had higher conceptual scores on the retention test. Examining misconception errors on only the conceptual items indicated that students in the *incorrect* condition made somewhat fewer role of zero misconception errors (38% vs. 43% of possible items), and somewhat fewer whole number misconception errors (46% vs. 51% of possible items) than students in the *correct* condition, F (1, 69) = 3.17, $p = .079$ and F (1, 69) = 3.03, $p = .086$, respectively.

In summary, on the retention test, students in the *incorrect* condition had higher conceptual scores and made fewer misconception errors than students in the *correct* condition. However, they no longer had significantly higher procedural scores.

Effect of Condition on Intervention Activities

To help understand how learning was impacted by our two intervention conditions, we examined students' responses during the intervention. During the intervention, students completed 4 practice problems and answered 24 explanation prompts.

Practice Problems

Accuracy on the four practice problems was higher in the *incorrect* than in the *correct* condition ($M = 0.70$, $SD = 0.30$ vs. $M = 0.59$, $SD = 0.31$, *respectively*), but only marginally so, $F(1, 67) = 3.02$, $p = .087$.

We also coded what procedures students used to solve the practice problems based on their verbal reports. These procedures included the 3 correct and 3 incorrect procedures that were demonstrated in the intervention packets. There were no significant differences between conditions in frequency of using the different procedures. Not surprisingly, students in both conditions were most likely to use the correct procedures modeled in the intervention packets.

Explanations on Worked Examples

The explanation prompts varied between conditions because they were designed to facilitate the appropriate processes in each condition. Therefore, the coding of these explanations was used as a manipulation check. We used ANCOVA models, and adopted the more conservative alpha value of .005 to determine significance due to the exploratory nature of these analyses that required multiple tests.

We coded students' explanations for their correct use of decimal concepts, such as correctly explaining a misconception, discussing relative magnitude, or using decimal and fraction terminology (see Table 4 for details). Students in the *incorrect* condition were more likely to correctly explain misconceptions and to correctly discuss relative magnitude. There were not significant differences between conditions in their use of decimal and fraction terminology. We also coded when students made statements that endorsed misconceptions, including the less than zero, role of zero, whole number, and other misconceptions. Students endorsed misconceptions relatively infrequently, as we hoped. However, students in the *incorrect* condition were more likely to endorse misconceptions than students who only saw correct examples.

To explore the impact of explanation quality on outcomes, we evaluated which explanation features predicted performance on the posttest and retention test. In the ANCOVA models, frequency of each of the explanation types was used as a predictor of procedural and conceptual scores and whole number misconceptions. Pretest knowledge scores were included as covariates.

Correctly discussing overall decimal concepts positively predicted procedural scores at posttest and conceptual scores at retention, F (1, 70) = 5.02, p = .028, η^2 = .067

and F (1, 70) = 4.26, $p = .043$, $\eta^2 = .057$, respectively. The benefit of discussing decimal concepts was due in large part to correctly discussing decimal misconceptions, as this type of concept explanation positively predicted procedural scores at posttest and conceptual scores at retention as well, F (1, 70) = 4.30, $p = .042$, $\eta^2 = .058$ and F (1, 70) = 6.07, $p = .016$, $\eta^2 = .080$, respectively. Recall that students in the *incorrect* condition discussed decimal concepts more often and had greater procedural scores at posttest and conceptual scores at retention test, suggesting that their focus on concepts may be one reason that comparing incorrect examples to correct ones aided learning. The other explanation characteristics were not predictive of outcomes.

Table 4: Percentage of Intervention Explanations Containing Each Feature, By Condition

Explanation Characteristic	Sample Explanations	Correct	Incorrect
Correct concepts			
Correctly discuss misconceptions** [a]	"0.08 is not the same as 0.8."	.01	.24
Discuss relative magnitude**	"0.15 is a little bigger than 0.1."	.15	.26
Include decimal or fraction terms	"It's like the fraction 15 over 100." "It's the same as 15 out of 100."	.16	.15
Any correct concept** [a]		.26	.52
Endorse misconceptions**	"0.6 is smaller than 0.365."	.03	.11

**Conditions differ from each other at p < .005; [a] Frequency of this explanation predicted outcomes

Summary

Comparing incorrect examples to correct examples led to better performance on procedural items on the posttest and on conceptual items on the retention test, in part because the incorrect condition reduced misconception errors on the retention test, particularly whole number misconceptions. One advantage of the incorrect condition was that it focused students' attention on decimal concepts—both on why misconceptions were incorrect and on relative magnitude—which in turn, predicted learning outcomes.

CHAPTER IV

DISCUSSION

Presenting students with common misconceptions and contrasting them with correct solutions helped focus students' attention on decimal concepts, aided their short-term knowledge of procedures and their longer-term retention of decimal concepts, and diminished misconceptions. In the discussion, we integrate these findings with findings from past research, discuss instructional implications and suggest future directions for research in this area.

Integrating With Past Research on Incorrect Examples

As past research has suggested, studying correct and incorrect examples improves learning (e.g., Eryilmaz, 2002; Große & Renkl, 2007; Huang et al., 2008; Siegler, 2002). Importantly, the current findings indicate that this benefit is above and beyond the general benefits of comparison. In prior research on incorrect examples, comparison of incorrect examples to correct ones was not directly supported, and there were few opportunities for comparison in the control condition. In the current study, comparison of solution procedures was supported in both conditions, and including incorrect solutions promoted greater learning. We consider two important processes underlying the benefits of comparing incorrect examples to correct ones: improving formation of correct category concepts and decreasing use of incorrect procedures.

Comparing correct and incorrect examples may have enhanced students' learning by leading them to create advanced category concepts. Studying conflicting examples prompted students to think about concepts more deeply, and this may have helped students to create categories for correct concepts as well as misconceptions. Past research has suggested that activating opposing ideas simultaneously may get students to think about more complex concepts (VanLehn, 1999) and to create more developed categories for concepts (Kotovsky & Gentner, 1996). Similarly, studies have suggested that comparing correct and incorrect examples can improve students' knowledge of concepts (Eryilmaz, 2002; Van den Broek & Kendeou, 2008), especially after a delay (Huang et al., 2008). The current study provides further evidence that studying correct and incorrect examples may have led students to form more advanced category concepts. Students in the *incorrect* condition were much more likely to discuss correct concepts during the intervention than students in the *correct* condition. By comparing correct and incorrect examples, what makes the correct example right and what makes the incorrect example wrong are emphasized, and students may have created a new mental representation of concepts that labels incorrect concepts as incorrect (Van den Broek & Kendeou, 2008). Though this did not produce immediate differences in conceptual scores, it did help students retain the concepts better over a delay, including a reduction in misconceptions.

Including incorrect and correct examples can also improve procedure selection. Learning a correct procedure does not eliminate the use of incorrect ones (Siegler, 2002); directly contrasting correct and incorrect examples helped decrease the frequency of using incorrect procedures. Comparing correct and incorrect examples may have encouraged students to differentiate correct and incorrect procedures, which in turn supported greater selection of correct procedures over incorrect ones immediately after the intervention. However, after a two-week delay, use of correct procedures did not remain higher for these students. We suspect that additional practice solving problems with feedback would be needed to maintain the strength of correct procedures over a delay.

One reason students in the *incorrect* condition learned concepts and procedures better than those in the *correct* condition may be that they were provided with more noticeable contrast in their example pairs. Indeed, contrast may be a particularly important part of comparison (Rittle-Johnson & Star, 2009; Schwartz & Bransford, 1998). In the past, researchers have often focused on similarities when discussing comparison and have not focused on contrast (e.g., Catrambone & Holyoak, 1989; Gentner et al., 2003; Gick & Holyoak, 1983; Richland, Holyoak, & Stigler, 2004). However, recent evidence suggests that comparing contrasting examples improves learning (Schwartz & Bransford, 1998; Waxman & Klibanoff, 2000). For example, for young children completing a categorization task, those children who saw examples of category members and non-category members (i.e., contrasted correct and incorrect examples) learned more than those who saw examples of only category members (i.e., no contrast; correct examples only) (Namy & Clepper, under review). In addition, students learning algebra benefited most from the type of comparison that had the most contrast—comparing correct solution methods (Rittle-Johnson & Star, 2009). Providing learners with examples that have salient similarities and differences due to contrast appears to improve learning more than just providing examples with similarities. In the current study, though students in both conditions were asked to explain similarities and differences, the differences between the examples in the *incorrect* condition were much more salient. Perhaps this more salient contrast led to the more frequent discussion of misconceptions in the *incorrect* condition, which in turn may have supported less frequent use of incorrect procedures, fewer misconceptions and greater retention of concepts.

The current study provides clues to how comparing correct and incorrect examples may improve learning outcomes in mathematics. Comparison with incorrect examples improved concept categories and decreased use of incorrect procedures, possibly because of increased differentiation between examples due to contrast.

Instructional Implications

Having students compare correct and incorrect examples can be beneficial for learning, and there are several instructional implications from the current study. Textbook publishers and classroom teachers should include more incorrect examples in their lessons, along with materials to support comparison to correct examples. As previously mentioned, science refutation texts have begun to capitalize on this idea (Alvermann & Hague, 1989; Diakidoy et al., 2003; Van den Broek & Kendeou, 2008). However, incorrect examples are rarely incorporated in mathematics texts, and only recently have researchers begun developing a curriculum that encourages teachers to incorporate incorrect examples into their math lessons (Curtis, Heller, Clarke, Rabe-Hesketh, & Ramirez, 2009).

While further use of incorrect examples in math instruction needs to be encouraged, simply showing students incorrect examples without proper support seems unlikely to be effective for learning. Several features of the intervention materials in the current study may have been especially important. First, a correct and an incorrect example were presented simultaneously, side-by-side. Numerous studies on comparing correct examples indicate that side-by-side presentation of examples supports learning much better than sequential presentation (Gentner et al., 2003; Oakes & Ribar, 2005; Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2009). Second, students were directly prompted to make comparisons between correct and incorrect examples. Asking students to identify similarities and differences between examples is often encouraged by expert mathematics teachers (Fraivillig, Murphy, & Fuson, 1999; Huffred-Ackles, Fuson, & Sherin Gamoran, 2004; Lampert, 1990; Silver, Ghousseini, Gosen, Charalambous, & Strawhun, 2005), and has been an important aspect of other mathematics interventions using worked examples (e.g., Rittle-Johnson & Star, 2007, 2009). Third, the incorrect examples used in this study were illustrating a commonly held misconception. Previous research on effectively using incorrect examples for learning science has used incorrect examples that directly address students' misconceptions (Eryilmaz, 2002; Mestre, 1994). Simply using incorrect examples that illustrate random errors may not be as effective. Students may get the maximal instructional benefit by being presented with a common misconception next to a correct example and being asked to identify the similarities and differences in the examples.

Future Directions

Several next steps are necessary to understand the role incorrect examples can play in learning. First, it is unknown when in the mastery process incorrect examples may be best able to aid learning. Students in the current study had relatively high prior knowledge; students with low prior knowledge might not benefit from studying incorrect examples. Past research indicates that inexperienced learners may be best served by learning from only correct examples because they do not have enough knowledge to use incorrect examples effectively (Große & Renkl, 2007). In this study, college students with low prior knowledge provided fewer principle-based explanations and learned less when studying correct and incorrect examples than those who studied only correct ones (Große & Renkl, 2007). Other studies have also found that learners' prior knowledge can impact the effectiveness of an instructional method (Kalyuga, 2007; Kalyuga & Sweller, 2004; Rittle-Johnson & Kmicikewycz, 2008; Rittle-Johnson et al., in press; Snow, 1992). Consequently, it may be best to initially present students with only correct examples until they have some basic understanding of the domain, and then the instructor can move onto comparisons between correct and incorrect examples. Students may then be prepared to discuss correct concepts and notice meaningful differences between the correct and incorrect examples. Further research is needed with students with differing prior knowledge levels to determine if prior knowledge influences the effectiveness of instruction with correct and incorrect examples.

In addition, most students in the current study did not achieve mastery at posttest or retention test. There may be instructional strategies to further augment the general benefits of comparison, particularly with incorrect examples. Such instructional strategies may include additional time spent on practice problems. Practice problems with feedback can be very useful for learning (Atkinson et al., 2000); however, in the current study, students only completed four practice problems. More practice problems should increase students' opportunities to use correct procedures and to inhibit incorrect procedures (Siegler, 1996). Increased explanations from the instructor about decimal concepts might also improve students' abilities to benefit from comparison and incorrect examples. Instructional explanations can emphasize important conceptual structures that may help eliminate competing incorrect concepts (Renkl, 2002). These conceptual explanations may also help increase students' discussion of concepts in their own explanations and increase the benefits of using incorrect examples. Also, further studies should explore the effects of comparing incorrect and correct examples on learning in a classroom setting over multiple sessions.

Comparing correct and incorrect examples helped students reduce their misconceptions, learn correct procedures, and remember correct concepts above and beyond the general benefits of comparison. Further research should evaluate when in the learning process such comparison may be most beneficial for students and how to maximize the effectiveness of these comparisons.

REFERENCES

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*(1), 127-145.

Alvermann, D. E., & Hague, S. A. (1989). Comprehension of counterintuitive science text: Effects of prior knowledge and text structure. *Journal of Educational Research, 82*(4), 197-202.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from Examples: Instructional Principles from the Worked Examples Research. *Review of Educational Research, 70*(2), 181-214.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps. *Journal of Educational Psychology, 95*(4), 774-783.

Barzi, F., & Woodward, M. (2004). Imputations of Missing Values in Practice: Results from Imputations of Serum Cholesterol in 28 Cohort Studies. *American Journal of Epidemiology, 160*, 34-45.

Bell, M., Bretzlauf, J., Dillard, A., Hartfied, R., Isaacs, A., McBride, J., et al. (2004). Unit 4: Decimals and Their Uses. In *Grade 4 Everyday Mathematics: Teacher's Lesson Guide* (pp. 214-265). New York, NY: Wright Group/McGraw-Hill.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking Mathematically: Integrating Arithmetic and Algebra in Elementary School*. Portsmouth, NH: Heinemann.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1147-1156.

Curry, L. (2004). The Effects of Self-Explanations of Correct and Incorrect Solutions on Algebra Problem-Solving Performance. In K. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 1548). Mahwah, NJ: Erlbaum.

Curtis, D. A., Heller, J. I., Clarke, C., Rabe-Hesketh, S., & Ramirez, A. (2009). *The Impact of Math Pathways and Pitfalls on Students' Mathematics Achievement*. Paper presented at the Annual Meeting of the American Educational Research Association.

Diakidoy, I.-A. N., Kendeou, P., & Ioannides, C. (2003). Reading about energy: The effects of text structure in science learning and conceptual change. *Contemporary Educational Psychology, 28*, 335-356.

Eryilmaz, A. (2002). Effects of Conceptual Assignments and Conceptual Change Discussions on Students' Misconceptions and Achievement Regarding Force and Motion. *Journal of Research in Science Teaching, 39*(10), 1001-1015.

Fraivillig, J. L., Murphy, L. A., & Fuson, K. (1999). Advancing Children's Mathematical Thinking in Everyday Mathematics Classrooms. *Journal for Research in Mathematics Education, 30*, 148-170.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*(2), 393-405.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1-38.

Glasgow, R., Ragan, G., Fields, W. M., Reys, R., & Wasman, D. (2000). The Decimal Dilemma. *Teaching Children Mathematics, 7*(2), 89-93.

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.

Hiebert, J., & Wearne, D. (1985). A model of students' decimal computation procedures. *Cognition and Instruction, 2*, 175-205.

Huang, T.-H., Liu, Y.-C., & Shiu, C.-Y. (2008). Construction of an online learning system for decimal numbers through the use of cognitive conflict strategy. *Computers & Education, 50*, 61-76.

Huffred-Ackles, K., Fuson, K., & Sherin Gamoran, M. (2004). Describing Levels and Components of a Math-Talk Learning Community. *Journal for Research in Mathematics Education, 35*(2), 81-116.

Irwin, K. C. (2001). Using everyday knowledge of decimals to enhance understanding. *Journal for Research in Mathematics Education, 32*(4), 399-420.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509-539.

Kalyuga, S., & Sweller, J. (2004). Measuring Knowledge to Optimize Cognitive Load Factors During Instruction. *Journal of Educational Psychology, 96*(3), 558-568.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*(6), 2797-2822.

Kouba, V. L., Carpenter, T. P., & Swafford, J. O. (1989). Number and operations. In M. M. Lindquist (Ed.), *Results from the Fourth Mathematics Assessment of the*

National Assessment of Educational Progress (pp. 64-93). Reston, VA: National Council of Teachers of Mathematics, Inc.

Lampert, M. (1990). When the Problem is Not the Question and the Solution is Not the Answer Mathematical Knowing and Teaching. *American Educational Research Journal, 27*, 29-63.

McNeil, N. M. (2008). Limitations to teaching children 2+2=4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development, 79*(5), 1524-1537.

Mestre, J. P. (1994). Cognitive Aspects of Learning and Teaching Science. In S. J. F. L. C. Kerplelman (Ed.), *Teacher Enhancement for Elementary and Secondary Science and Mathematics: Status, Issues, and Problems*. Washington, D. C.: National Science Foundation (NSF 94-80).

Namy, L. L., & Clepper, L. (under review). The Differing Roles of Comparison and Contrast in Children's Categorization.

National Mathematics Advisory Panel. (2008). *Foundations of Success: The Final Report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

National Mathematics Advisory Panel Subcommittee. (2008). *Chapter 9: Report of the Subcommittee on the National Survey of Algebra I Teachers*. Washington, DC: U.S. Department of Education.

NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Oakes, L. M., & Ribar, R. J. (2005). A Comparison of Infants' Categorization in Paired and Successive Presentation Familiarization Tasks. *Infancy, 7*(1), 85-98.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6*(4), 449-468.

Putt, I. J. (1995). Preservice Teachers Ordering of Decimal Numbers: When More Is Smaller and Less Is Larger! *Focus on Learning Problems in Mathematics, 17*(3), 1-15.

Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction, 12*(5), 529-556.

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. (1989). Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education, 20*, 8-27.

Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2004). Analogy Use In Eighth-Grade Mathematics Classrooms. *Cognition and Instruction, 22*, 37-60.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*(3), 249-273.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*(1), 175-189.

Rittle-Johnson, B., & Kmicikewycz, A. O. (2008). When generating answers benefits arithmetic skill: The importance of prior knowledge. *Journal of Experimental Child Psychology, 101*, 75-81.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*(3), 561-574.

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*(3), 529-544.

Rittle-Johnson, B., Star, J. R., & Durkin, K. (in press). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology*.

Sackur-Grisvard, C., & Leonard, F. (1985). Intermediate cognitive organizations in the process of learning a mathematical concept: The order of positive decimal numbers. *Cognition and Instruction, 2*, 157-174.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147-177.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475-522.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Garnott & J. Parziale (Eds.), *Microdevelopment: A process-oriented perspective for studying* *development and learning* (pp. 31-58). Cambridge, MA: Cambridge University Press.

Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Strawhun, B. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *Journal of Mathematical Behavior, 24*, 287-301.

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist, 27*(1), 5-32.

Stacey, K., Helme, S., Steinle, V., Baturo, A., Irwin, K., & Bana, J. (2001). Preservice teachers' knowledge of difficulties in decimal numeration. *Journal of Mathematics Teacher Education, 4*, 205-225.

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 101*, 408 - 426.

Van den Broek, P., & Kendeou, P. (2008). Cognitive Processes in Comprehension of Science Texts: The Role of Co-Activation in Confronting Misconceptions. *Applied Cognitive Psychology, 22*, 335-351.

VanLehn, K. (1999). Rule-Learning Events in the Acquisition of a Complex Skill: An Evaluation of Cascade. *The Journal of the Learning Sciences, 8*(1), 71-125.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571-581.