THE EFFECTIVENESS OF INCORRECT EXAMPLES AND COMPARISON WHEN LEARNING ABOUT DECIMAL MAGNITUDE

By

Kelley Durkin

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSPHY

in

Psychology

May, 2012

Nashville, Tennessee

Approved:

Professor Lynn Fuchs

Professor Bruce McCandliss

Professor Bethany Rittle-Johnson

Professor Georgene Troseth

ACKNOWLEDGEMENTS

First, I would like to thank Bethany Rittle-Johnson for being the best advisor and mentor imaginable. This work would not have been possible without Bethany's valuable guidance, feedback, and encouragement. No matter how busy she was, Bethany always managed to make time for me, and I would certainly not be where I am today if it were not for her. She is my role model on how to be a wonderful mentor, researcher, and person.

In addition, thank you to all of my committee members for their priceless comments and suggestions throughout the research process: Lynn Fuchs, Bruce McCandliss, Jon Star, and Georgene Troseth. I will be forever grateful for their time and attention, and each of them has provided me with important lessons.

I am also grateful to all of my lab members throughout the years who have helped me at each step along the way. Thank you to Emily Fyfe, Percival Matthews, and Katie McEldoon for reading countless drafts, listening to numerous practice talks, and providing me with endless encouragement. These last five years would not have been as meaningful or fun without you. Thank you to Marci DeCaro and Michael Miller for providing me with valuable feedback and for being exemplary models of researchers. Also, I am extremely grateful to all of the research assistants, our lab manager, and my fellow graduate students who have helped me collect, enter, and code all of this data. So thank you to Matthew Faulkner, Dylan Hillman, Abbey Loehr, Jennifer Samson, Michael Weinstein, and Alanne Wheeler. This research would not have happened without their countless hours of work and support.

Next, I would like to thank all of my friends who have stuck with me throughout this process. They have put up with me talking their ears off about incorrect examples and comparison, listened to me obsess about graphs, and dealt with me disappearing from the social world for periods of time. In particular, thank you to Waylon Envik for never complaining about the fact that he now knows more about psychology and education than he ever wished.

I am also thankful for the funding and help I have received throughout the years. The research reported here was funded by the Institute of Education Sciences, U.S. Department of Education, Grant #R305B040110. The views expressed in this paper are mine and not theirs. In addition, thank you to all of the schools, teachers, and students who participated in this research.

Finally, thank you to my family for everything. I would not be the person I am today without you. I feel like I won the sister lottery with Erin and Kathleen Durkin. I cannot imagine more supportive, intelligent, and fun sisters. I also do not know how I got so lucky to have such amazing parents. James and Lynn Durkin are always encouraging and optimistic, and they can always manage to make me smile. And so I say to my family, YTBO!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Misconceptions can persist over a long period of time and must be overcome (Eryilmaz, 2002). For example, students have common and persistent misconceptions involving decimal magnitude, such as treating decimals the same as whole numbers (e.g., Irwin, 2001; Resnick et al., 1989). This is often due to the fact that students who are learning a new topic relate it to their prior knowledge in other domains. While this can be helpful for learning, it can also provide barriers for learning if prior knowledge is misapplied to new domains (e.g., Stafylidou & Vosniadou, 2004; Vamvakoussi & Vosniadou, 2004). Unfortunately, these misconceptions can become entrenched and adverse to change (McNeil & Alibali, 2005). One potential way to correct misconceptions and improve student learning is through the use of incorrect examples (e.g., Huang, Liu, & Shiu, 2008; Siegler, 2002). Past research has shown that presenting students with incorrect examples can help students correct their misconceptions and improve their knowledge of correct concepts (Eryilmaz, 2002; Huang et al., 2008; Van den Broek & Kendeou, 2008) and procedures (Durkin & Rittle-Johnson, 2012; Große & Renkl, 2007; Siegler, 2002).

However, it remains unclear *how and when* exposure to incorrect examples improves learning. One promising reason for why incorrect examples are beneficial is because students are comparing them to correct examples. Past research has illustrated that comparison of two correct examples can improve learning (e.g., Gentner, Loewenstein, & Thompson, 2003), and the specific mechanisms involved in comparison may be especially helpful when using incorrect examples. However, prior knowledge may play a central role in determining when incorrect examples and comparison are helpful (Große & Renkl, 2007; Rittle-Johnson, Star, & Durkin, 2009). In the current study, I experimentally evaluated whether comparison of correct and incorrect examples benefited students' knowledge beyond exposure to incorrect and correct examples or exposure to only correct examples. Participants varied in their prior knowledge, so the moderating effects of prior knowledge on students' learning from incorrect examples and comparison could be examined.

First, I discuss past research misconceptions and on the effectiveness of exposing learners to incorrect examples. Next, I describe the importance of comparison and how comparing incorrect and correct examples may improve learning. Then, I describe the target domain of the current study, decimal magnitude. Finally, I describe the current study and my hypotheses.

Misconceptions in Academic Domains

Misconceptions can be defined as incorrect or erroneous ideas that contradict accepted ideas, and they can form as students attempt to "assimilate…new information into their existing conceptual structures" (Stafylidou & Vosniadou, 2004, p. 505). While the term "misconception" is used frequently in research on science learning (e.g., Eryilmaz, 2002), it is not always used in research on mathematics learning. When discussing erroneous ideas, math education researchers will often label them as buggy errors or rules (e.g., Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Resnick et al., 1989). However, the term "misconception" has been used by some math education researchers (e.g., Irwin, 2001), and it is used frequently in conceptual change research on mathematics learning (e.g., Vamvakoussi & Vosniadou, 2004). Consequently, I will use the term "misconception" in this paper because it more accurately reflects the persistent misunderstandings students have in the current study that relate to incorrectly understanding concepts as opposed to simply misunderstanding operations or notation. Such misconceptions can become deeply entrenched and difficult to change over time (e.g., McNeil & Alibali, 2005).

Effectiveness of Incorrect Examples

One possible way to help correct misconceptions is by using incorrect examples. The use of incorrect examples has been studied primarily in science domains, and this research can provide insight into the possible benefits of incorrect examples. For example, numerous classroom experiments on science learning indicated that exposing students to incorrect examples in physics was more effective for overcoming misconceptions than focusing on correct concepts alone (Eryilmaz, 2002; Mestre, 1994). Also, students were more likely to revise incorrect concepts when using refutation science texts that included both incorrect and correct examples than when using texts that only included correct examples (Alvermann & Hague, 1989; Diakidoy, Kendeou, & Ioannides, 2003; Van den Broek & Kendeou, 2008).

Similarly, incorrect examples have been shown to improve students' mathematical knowledge, including both conceptual and procedural knowledge. Conceptual knowledge is defined as the ability to recognize and understand key domain concepts. Procedural knowledge is defined as the ability to execute action sequences to solve problems. These two distinct knowledge types have been used frequently in past research (e.g., Canobi, Reeve, & Pattison, 2003; Große & Renkl, 2006; Hiebert & Wearne, 1996), and while the two types of knowledge are related, they are distinct constructs that can be measured separately (Schneider, Rittle-Johnson, & Star, 2011). Many tasks may involve both conceptual and procedural knowledge, but procedural demands dominate some tasks and conceptual demands dominate others.

Generally, procedural knowledge is involved in more familiar tasks, on which students likely know a solution procedure. On the other hand, conceptual knowledge is involved in unfamiliar tasks, on which students must transfer knowledge to new problem types that assess underlying domain principles. Incorrect examples can help improve both of these types of knowledge.

In two studies, students showed improvements in conceptual knowledge for mathematics when using incorrect examples (Durkin & Rittle-Johnson, 2012; Huang et al., 2008). For instance, 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). Similarly, fourth- and fifth-grade students who compared incorrect procedures for placing decimals on number lines to correct procedures remembered decimal concepts better than students who only compared correct procedures (Durkin & Rittle-Johnson, 2012).

Exposure to incorrect examples can also improve procedural knowledge in mathematics. In a given domain, learners often have and use multiple procedures and ways of thinking at any given time. Because of the coexistence of multiple procedures, they often continue to use incorrect procedures after correct procedures have been learned (see Siegler, 2002 for a review). Having students explain why incorrect procedures are wrong is one way to reduce their use of incorrect procedures. Third- and fourth-graders who explained correct and incorrect procedures to mathematical equivalence problems (e.g., $3 + 4 + 8 = \_ + 8$) were able to solve more difficult problems than those who only explained correct procedures. Primarily, this was because they generated more generalizable correct procedures that were applicable to a wider range of problem types (Siegler, 2002). Similar results have been found for college students learning algebra and probability (Curry, 2004; Große & Renkl, 2007) and fourth- and fifth-grade students learning decimals (Durkin & Rittle-Johnson, 2012).

However, students' prior knowledge may influence the effectiveness of using incorrect examples. Past research suggests that learners with low prior knowledge may learn best from correct examples alone, but past findings are not consistent. Learners with low prior knowledge often do not have sufficient knowledge to use incorrect examples effectively, nor do they always understand why an error is wrong (Große & Renkl, 2007; Stark, Kopp, & Fischer, 2011). For example, college students with low prior knowledge of probability and general mathematics learned more from studying correct examples alone than from studying both correct and incorrect examples. However, students with sufficient prior knowledge benefited from studying the incorrect examples (Große & Renkl, 2007). In contrast, I found that prior domain knowledge did not influence the effectiveness of incorrect examples, and students benefited from incorrect examples regardless of prior knowledge (Durkin & Rittle-Johnson, 2012). In this study, fourth- and fifth-grade students learned about decimals by comparing incorrect and correct examples or by comparing correct examples only. Thus, there is limited evidence on the importance of prior knowledge when using incorrect examples, and the findings are inconsistent. This could be due to a variety of factors, including the use of different methods. The first study included college students studying probability, a domain that involves commonly confused problem types rather than strongly held misconceptions (Große & Renkl, 2007). The second study included younger students studying decimals, a domain with many prevalent misconceptions. Furthermore, all students in this second study received the additional scaffold of comparison. Also, past studies have not measured students' prior prevalence of misconceptions as a measure of prior knowledge. In domains with persistent misconceptions, the prevalence and strength of students' prior misconceptions may play an important role in the effectiveness of incorrect examples. Thus, it is important to attend to students' prior knowledge and misconceptions when using incorrect examples, and further evidence is needed on this topic.

In addition to investigating *who* will benefit from incorrect examples, it is also necessary to examine *how* incorrect examples can benefit learning. Exposure to incorrect examples can improve learning, at least under some circumstances, but less is known about how this exposure supports knowledge growth. There is little direct evidence, but researchers have offered at least three reasons. First, exposure to correct and incorrect examples may improve conceptual knowledge because it engages students' conflicting correct and incorrect concepts. In studies involving incorrect examples, participants are often told when an example is incorrect and asked to explain why it is incorrect or to correct the error. Consequently, students may create a new mental representation, or schema, of the material that labels incorrect concepts as wrong. Students' responses during a think-aloud procedure when studying correct and incorrect examples supported this idea (Van den Broek & Kendeou, 2008). For example, students with misconceptions who saw correct and incorrect examples were more likely to notice and respond to conflicting ideas than if they saw correct examples alone. Second, students may be motivated to think more deeply about correct concepts (Durkin & Rittle-Johnson, 2012; Van den Broek & Kendeou, 2008; VanLehn, 1999). For example, students who compared correct and incorrect examples were twice as likely to correctly identify misconceptions and mention correct concepts as students who only compared correct examples (Durkin & Rittle-Johnson, 2012). Third, explaining examples of incorrect procedures may help decrease the strength of incorrect procedures, reducing the probability that the procedure will be selected in the future and leading students to use more correct procedures (Durkin & Rittle-Johnson, 2012; Siegler, 2002). 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 (Durkin & Rittle-Johnson, 2012; 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).

In addition to these three potential mechanisms, comparison may be an essential mechanism supporting learning from incorrect examples. For instance, direct comparison of incorrect examples to correct examples may allow students to more easily identify critical attributes of both correct examples and misconceptions. However, most past research has not directly supported comparison nor has it evaluated the potential role of spontaneous comparison when correct and incorrect examples are presented. In fact, in most past research, incorrect examples were presented alone, and students did not directly compare them to correct examples (Curry, 2004; Große & Renkl, 2007; Huang et al., 2008). There are a few exceptions. In one study, an incorrect solution was presented at the same time as a correct solution (Siegler, 2002), and in a few studies on refutation science texts, incorrect examples were presented on the same page as correct examples and discussed together (Alvermann & Hague, 1989; Diakidoy et al., 2003; Van den Broek & Kendeou, 2008). In these cases, direct comparison of the two was not explicitly encouraged, and the frequency of spontaneous comparison was not assessed. While students may have spontaneously compared incorrect examples to correct examples, doing so would require good metacognitive skills (Richland, Morrison, & Holyoak, 2006).

In two additional studies, all students compared examples, with some students comparing correct and incorrect examples while others compared only correct examples (Durkin & Rittle-Johnson, 2012, in preparation). However, due to the fact that all students in these studies compared examples, it is unclear whether comparison was necessary for learning. Perhaps mere exposure to incorrect examples is sufficient for helping students to identify and reduce their misconceptions and incorrect procedures. Yet, it is also possible that *comparing* incorrect and correct examples was an important part of why incorrect examples were beneficial to students in these studies. Past research on comparing correct examples suggests that it is.

The Role of Comparison

Given the potentially significant role of comparison in learning from incorrect examples, it is important to consider the benefits of comparison in learning. However, past research on comparison has primarily focused on comparing *correct* examples. Based on this work, comparison is often lauded as an effective and important learning process in cognitive science (Gentner et al., 2003; Gick & Holyoak, 1983) and in mathematics education (NCTM, 2000; Silver, Ghousseini, Gosen, Charalambous, & Strawhun, 2005). Comparing two correct examples can help people recognize more abstract, high-level commonalities between them (e.g., Kotovsky & Gentner, 1996) and increase conceptual and procedural knowledge (e.g., Rittle-Johnson & Star, 2009). For example, comparing two correct solution procedures improved students' knowledge of equation-solving and estimation procedures more than studying the same examples one at a time (Rittle-Johnson & Star, 2007; Rittle-Johnson et al., 2009; Star & Rittle-Johnson, 2009). In the domain of estimation, comparison also improved students' conceptual knowledge more than viewing examples sequentially (Star & Rittle-Johnson, 2009). Further, explicit prompts to compare greatly improve the benefits of studying multiple correct examples (Catrambone & Holyoak, 1989; Gentner et al., 2003). For example, college students who were explicitly told to compare two problems that differed in surface features but could be solved using the same procedure were more likely to notice the convergent procedure than students who were not prompted to compare (Catrambone & Holyoak, 1989). Together, these studies suggest that students benefit from comparison, especially when explicitly prompted to compare.

One of the primary benefits of comparison is that it allows learners to see the underlying structure of both examples (e.g., Loewenstein, Thompson, & Gentner, 1999). Students often focus on unimportant, surface features that are not relevant to the target concepts and procedures (Gentner, 1989). For example, when solving word problems, students often focus on the cover story as opposed to the underlying structure of the problems (Catrambone & Holyoak, 1989). Making direct comparisons between two examples can lead to structural alignment that highlights the shared relational structure between two examples for students (Gentner, 1983). In turn, this can facilitate students' transfer of knowledge to novel problems with the same underlying structure, but different surface features. Thus, comparisons can lead to students noticing important, deep structural aspects of examples instead of merely noticing surface features, and students can then more easily identify meaningful similarities and differences (Loewenstein et al., 1999).

Comparison and Incorrect Examples

For the same reasons comparing correct examples can improve learning, comparison may improve learning when comparing incorrect and correct examples. When comparing incorrect and correct examples, alignment of examples could help students recognize important structural features of incorrect and correct examples. In addition, detecting the differences between correct and incorrect examples may help students accurately update schemas of correct concepts and create schemas for misconceptions (VanLehn, 1999). For example, comparing incorrect and correct examples of the placement of decimals on a number line may allow students to update their schema of decimal magnitude. By aligning the incorrect and correct examples, students can go beyond noticing surface features, such as the total number of digits in the decimal, and focus on relevant structural features, such as the digit in the highest place value. While exposure to incorrect examples can help students identify misconceptions, it may not help them understand why these misconceptions are incorrect. *Comparing* incorrect examples to correct ones may push students to understand misconceptions more deeply and how they differ from correct concepts, and may encourage students to discuss correct concepts more frequently (Durkin & Rittle-Johnson, 2012). This should lead to deeper conceptual knowledge.

In addition to helping students update their concepts, comparing incorrect and correct examples can help students reduce competition from incorrect procedures. By comparing incorrect and correct examples, students may be more likely to notice important, meaningful differences between the procedures illustrated in the two examples. This may help students update their knowledge of what characteristics of a procedure are important for future success on problems. For example, after comparing incorrect and correct examples of how to place a decimal less than 1 on a number line, students may notice that procedures that focus on the tenths place are more successful than procedures that focus on how many digits a decimal has. This comparison may prompt them to encode the value of the digit in the tenths place and select a correct procedure instead of an incorrect one.

While comparison can be beneficial for learning, prior knowledge is important when considering the effectiveness of comparison. Sufficient domain knowledge may be required in order to benefit from comparison (Rittle-Johnson et al., 2009). For example, middle school students learning algebra benefited from comparing multiple solution methods if they were familiar with one of the target solution methods, but not if they were not (i.e., had low prior knowledge). Students with low prior knowledge benefited more from viewing these methods sequentially (Rittle-Johnson et al., 2009). This is possibly due to the fact that if students have low prior knowledge, the concepts and procedures illustrated in examples they compare are often unfamiliar. It is difficult for students to align two unfamiliar examples because this unfamiliarity makes it hard for students to recognize the aspects to which they should attend (Gentner et al., 2003; Rittle-Johnson et al., 2009; Schwartz & Bransford, 1998). This can make it hard for students to understand the importance of similarities and differences between examples. Consequently, comparison may be too overwhelming to significantly improve learning.

Overall, comparison is an important learning mechanism that can improve students' conceptual and procedural knowledge. Comparison may be particularly useful for students when using incorrect examples because it may help students recognize the underlying structure of examples, correctly update their schemas, and reduce competition between correct and incorrect procedures. However, students may need sufficient prior knowledge to benefit from comparison. Consequently, more evidence is needed on the role of prior knowledge and comparison, particularly when learning from incorrect examples.

Target Domain - Decimals

I evaluated the effects of incorrect examples and comparison in the domain of decimal fractions, commonly referred to as decimals (Resnick et al., 1989). It is important for students to master decimals to improve their learning in more advanced mathematics. For instance, mastering decimals is important for later algebra proficiency (National Mathematics Advisory

Panel, 2008) and for more advanced mathematical tasks involving decimals (Hiebert & Wearne, 1985).

Unfortunately, students and adults often have difficulty understanding decimals. For instance, 67% of fourth-graders could not solve a problem correctly that involved placing 1.7 on a number line from 0 to 3 (National Center for Education Statistics, 2011). Not only are decimals difficult for students to master, they can also be difficult for adults. For example, half of preservice teachers failed to get 75% of a test on decimals correct on their first try (Putt, 1995). Thus, students and adults struggle with the domain of decimals.

Such difficulties with decimals often stem from common and persistent misconceptions involving decimal magnitude (e.g., Desmet, Gregoire, & Mussolin, 2010; Glasgow et al., 2000; Irwin, 2001; Kouba, Carpenter, & Swafford, 1989; Putt, 1995; Resnick et al., 1989; Rittle-Johnson, Siegler, & Alibali, 2001; Sackur-Grisvard & Leonard, 1985; Stacey et al., 2001). This makes decimals an ideal domain for studying incorrect examples. Three common misconceptions are: 1) the *whole number* misconception, 2) the *role of zero* misconception, and 3) the *fraction* misconception. First, the *whole number* misconception involves thinking of decimals as if they are whole numbers (e.g. thinking 0.25 is greater than 0.7 because 25 is greater than 7). Students incorrectly apply their generally strong knowledge about whole numbers to decimals. Second, a related misconception specifically involves the *role of zero*, which is different for whole and decimal numbers. When a zero is in the tenths place, students often ignore it and treat the following digit as if it is in the tenths place (e.g. students will think 0.07 is the same as 0.7). In addition, students will assume that adding a zero on the end of a decimal 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. 320 is bigger than 32). A third common misconception is to think of decimals like common *fractions* and assume that longer decimals are smaller because they contain smaller parts, just as a fraction with the same numerator and larger denominator than another fraction is smaller because it contains smaller parts (e.g., 0.784 must be less than 0.3 because 1/784 has smaller parts than 1/3). These misconceptions are the ones most frequently held by students in the United States (Resnick et al., 1989; Sackur-Grisvard & Leonard, 1985), so they were the focus of the current study.

To teach students about decimals, I used a number line task. Number line tasks are a useful way to teach students about decimal magnitude (e.g., National Mathematics Advisory Panel, 2008; Rittle-Johnson et al., 2001). Placing decimals on a number line can help students better understand magnitude relations and create links between decimal concepts and procedures (National Mathematics Advisory Panel, 2008). In addition, this was a novel task for many students in the current study because most students had not been exposed to placing decimals on number lines yet. Therefore, I focused on children learning to place decimals on number lines during the intervention, a task that should engage their concepts of decimal magnitude and allow them to generate procedures for completing the task.

In summary, decimals are important for students to master, but students often have trouble understanding decimals. This is frequently due to commonly held misconceptions, which need to be overcome. The persistent nature of misconceptions in this domain makes it ideal for studying the effectiveness of incorrect examples and comparison.

Current Study

In the current study, students compared correct and incorrect examples (*Incorrect-Compare* condition), studied correct and incorrect examples sequentially (*Incorrect-Sequential* condition), or studied only correct examples sequentially (*Correct-Sequential* condition). Examining the differences between these groups, I was able to test the effects of studying incorrect examples and the effects of comparing incorrect examples to correct ones. I was also able to explore the role of prior knowledge, including misconceptions, when learning from incorrect examples and comparison.

Aspects of the current study's design merit discussion. First, this study was conducted in classrooms. This allowed me to examine the effects of comparing incorrect and correct examples in a classroom setting. All past empirical work on mathematics learning from studying correct and incorrect examples has been done using brief, one-on-one tutoring (e.g., Durkin & Rittle-Johnson, 2012; Große & Renkl, 2007; Siegler, 2002). Second, students were randomly assigned to one of the three conditions within their classrooms. This resulted in each classroom containing students in all conditions to minimize classroom affecting the impact of condition. Third, students studied examples and generated explanations with a partner. Students worked with a partner rather than by themselves because students learn more working with a partner than working alone, especially when encouraged to explain (Fuchs et al., 1997; Johnson & Johnson, 1994). Also, students were prompted for explanations throughout the intervention. Generating self-explanations has been shown to improve learning compared to students who do not spend time generating explanations, so all students in the current study generated explanations (e.g., Chi, de Leeuw, Chiu, & LaVancher, 1994). Fourth, practice problems were also completed at regular intervals throughout the packet. The combination of worked examples and practice problems was used during the intervention because this combination can improve learning in a variety of domains beyond solving practice problems alone (see Atkinson, Renkl, & Merrill, 2003 for a review). These basic design features have been used successfully in past research on comparison of correct examples in mathematics classes (e.g., Rittle-Johnson et al., 2009).

Hypotheses

In the current study, I hypothesized that students in the *Incorrect-Compare* condition would have higher procedural and conceptual knowledge than students in the *Incorrect-Sequential* condition based on past research illustrating the benefits of comparison (e.g., Rittle-Johnson & Star, 2007). I hypothesized that students in the *Incorrect-Sequential* condition would have higher conceptual and procedural knowledge than students in the *Correct-Sequential* condition based on past research emphasizing the benefits of studying incorrect examples (e.g., Durkin & Rittle-Johnson, 2012). I expected this to be true immediately after the intervention and after a delay, in part due to a reduction in misconception errors.

I also explored the potential moderating role of prior knowledge. Recall that some past work has suggested that students with low prior knowledge may have trouble with incorrect examples (Große & Renkl, 2007) and with comparison (Rittle-Johnson et al., 2009), making the combination of incorrect examples and comparison difficult. In the current study, I used several measures of prior knowledge, including accuracy at pretest and the prevalence of students' prior misconceptions. I did not have strong hypotheses on prior knowledge due to mixed results in past research.

CHAPTER II

METHOD

This was a classroom-based study that examined whether comparison of correct and incorrect examples benefited students' learning beyond simply being exposed to incorrect examples or being exposed to only correct examples. Students were prompted to compare incorrect and correct examples, to study incorrect and correct examples sequentially, or to study only correct examples sequentially. Students studied and discussed packets of worked examples with a partner in their math class for approximately 2 hours.

Participants

All 378 students from 9 fourth-grade and 9 fifth-grade classrooms at three public, suburban schools participated. The fourth-grade classrooms were from the two elementary schools that fed into a middle school that contained the fifth-grade classrooms. Thirty-six students were dropped from the analysis: 31 students because they were absent for at least 1 day of the intervention, 4 students because they were unable to use our materials due to very significant learning disabilities, and 1 student because he was absent for every assessment. Of the remaining 342 students, 185 were in fourth-grade, 157 were in fifth-grade, 53% were female, 95% were Caucasian (2% African-American, 2% Hispanic, and 1% Asian), and the mean age was 10.17 years (range 8.83 -12.83 years). At one elementary school, 31% of students were eligible for free or reduced lunch, at the other elementary school, 79% of students were eligible, and at the middle school, 47% of students were eligible. Out of a total of 13 teachers, the mean number of years teaching was 16 years, with a mean of 13 years of experience teaching elementary or middle school math in particular.

Teacher reports provided more details about the teachers and their classrooms. All teachers of the participating classes used their respective grade's version of Pearson Education's enVision Math textbook. In addition, previous exposure to decimals was limited, although varied; 7 teachers reported spending a few days on decimals, 4 had spent a few weeks, and 2 had not previously covered decimals. Students' experience with placing decimals on number lines also varied. Six teachers reported that none of their students had experience placing decimals on number lines, 2 teachers reported a few of their students could do so, and 4 teachers reported that some of their students could do so.

Design

A pretest-intervention-posttest design was used followed by a retention test administered 3 weeks after the posttest. For the intervention, students were randomly assigned a partner in their class, and then each pair of students was randomly assigned to one of three intervention conditions: the *Incorrect-Compare* condition (n = 116, 58 groups), the *Incorrect-Sequential* condition (n = 117, 57 groups), or the *Correct-Sequential* condition (n = 109, 55 groups). When there was an odd number of students in a classroom, students formed groups of three. There were 5 such groups in the current study. Students in the *Incorrect-Compare* condition studied sets of two worked examples of the same problem (one incorrect and one correct) on each page followed by questions that encouraged comparison of these two different solution methods for the same problem. Each example was labeled as correct or incorrect. Prompts encouraged students to discuss why one example was correct while the other was incorrect and to find similarities and differences between the examples. Students in the *Incorrect-Sequential* condition studied the same worked examples, but each page only contained one worked example and one question that encouraged reflection on that example. Each example was labeled as correct or incorrect. Students in the *Correct-Sequential* condition studied only correct worked examples, and each page only contained one worked example and one question that encouraged reflection on that example (see Figure 1). Each example was labeled as correct. All students completed the intervention over the course of two class periods, which were each about one hour long.

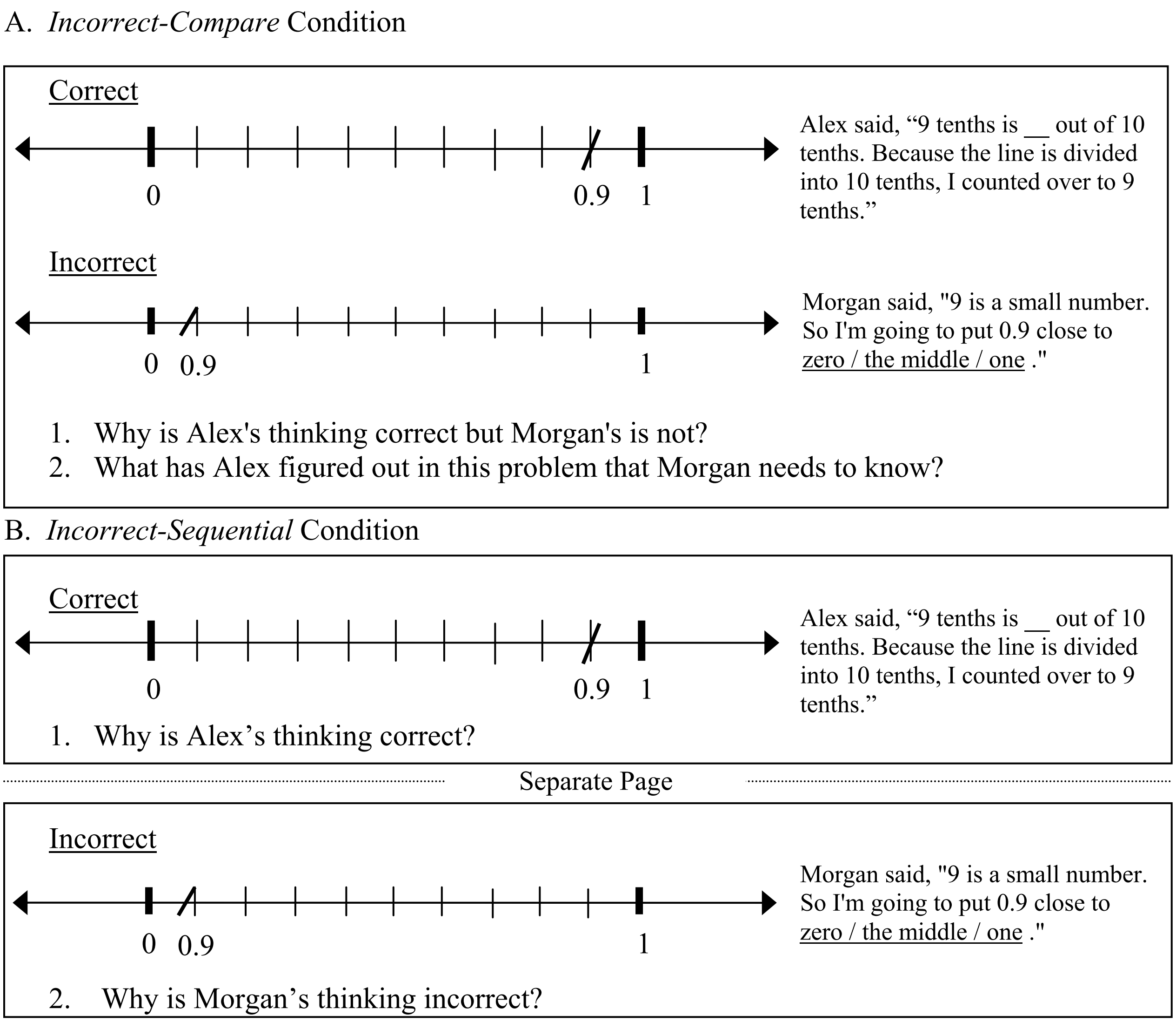
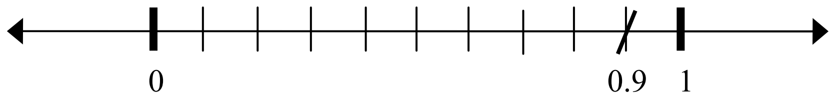


Figure 1: Sample intervention packet pages for each condition

C. *Correct-Sequential* Condition

<u>Correct</u>

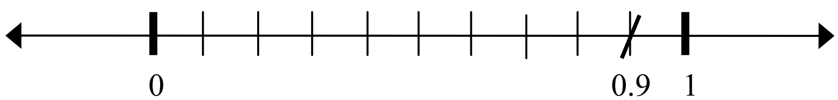


Alex said, "9 tenths is __ out of 10 tenths. Because the line is divided into 10 tenths, I counted over to 9 tenths."

1. Why is Alex's thinking correct?

Separate Page

<u>Correct</u>



Morgan said, "I know 9 tenths is only one tenth smaller than 1. So I marked 0.9 a little <u>before / after</u> 1."

2. Why is Morgan's thinking correct?

Figure 1 continued: Sample intervention packet pages for each condition

Materials

Introductory Lessons

All students received brief lessons after the pretest and at the beginning of each intervention day (see Appendix A). After students completed the pretest, a researcher and the classroom teacher modeled how students should work with their partners based on a previous lesson used in past studies that involved partner work (e.g., Rittle-Johnson et al., 2009). On the first intervention day, students were reminded about the meaning of different decimal place values with instruction adapted from a lesson in the *Everyday Mathematics* curriculum (Bell et al., 2004). On the second intervention day, students were told to attend to the labels on each number line, as the number lines might go from 0 to 1 or from 0 to 10. Most students had some previous experience with basic decimal concepts, and the purpose of these brief lessons was to activate students' prior knowledge of terminology and of what different place values mean.

Intervention Packet

Each intervention packet contained 36 worked examples of decimal number line problems presented with 36 corresponding questions (18 on each day of the intervention). Each worked example illustrated where a hypothetical student, Alex or Morgan, placed a decimal on a number line and an explanation of his or her procedure (see Figure 1). On 30 of the worked examples the number line went from 0 to 1, and on 6 of the worked examples the number line when from 0 to 10. I added these 0 to 10 number lines to aid in the ability to transfer knowledge learned from using the 0 to 1 number line because my past study illustrated that students had great difficulty transferring this knowledge (Durkin & Rittle-Johnson, 2012). The numbers used in the worked examples were the same in all conditions. 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 24 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). To encourage students to carefully read and process the examples, the worked examples also required students to fill in blanks and circle correct words to make the statements true (see Figure 1). It was not expected that all students would complete all examples on each day, and the packets were designed so that later examples repeated ideas that were illustrated in earlier examples. Consequently, students who did not complete the entire packet were still exposed to all the different problem features and ideas. In addition, the explanation prompts paired with each worked example were designed to elicit discussion about the worked example. In the *Incorrect-Compare* condition these prompts focused on students' comprehension of the examples and comparison between the correct and incorrect ideas. In the *Incorrect-Sequential* condition, the prompts focused on students' comprehension of the examples and discussion of correct and incorrect ideas separately. In the *Correct-Sequential* condition, the prompts focused on students' comprehension of the examples and discussion of correct ideas.

The three demonstrated correct procedures were based on procedures students have reported using in past research (Irwin, 2001; Rittle-Johnson et al., 2001). The first procedure targeted the number of tenths in a decimal by putting the value near where that number of tenths would go (e.g., For 0.280, estimating where 2 tenths would go, and then moving the mark over a little more). The second procedure involved imagining that the number line was divided into the number of pieces specified by the smallest place value and positioning the decimals based on this (e.g., For 0.280, imagine dividing the number line into 1000 pieces and estimating where 280 would be placed). The third correct procedure involved using benchmarks; the location was approximated based on knowledge of the decimal's magnitude in relation to 0, 0.5, and 1. The three demonstrated incorrect procedures corresponded to the three most common decimal misconceptions in the U.S. - the *whole number* misconception, the *role of zero* misconception, and the *fraction* misconception (e.g., Desmet et al., 2010; Irwin, 2001; Resnick et al., 1989). The packets also contained 25 practice problems spread throughout that asked students to place a slash on a number line where the decimal would be located (13 on the first day, 12 on the second day). On 8 practice problems, students were asked to give written explanations of how they got their answers so that I could code their procedure use. After completing the first four practice problems each day, students were told to stop and wait for an experimenter or teacher to check their answers. Students were given feedback as to whether their placement was correct or not. If their placement was incorrect, students were told to try again. If their placement was still incorrect after the second attempt, students were told to move on in the packet. Feedback was provided because feedback has been shown to improve learning with worked examples and practice questions (Krause, Stark, & Mandl, 2009). Students were given a second opportunity to try problems they solved incorrectly to improve learning from the practice problems. If students did not solve the problems correctly after the second attempt, they were told to move on and were not shown the correct answer.

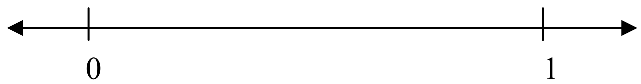
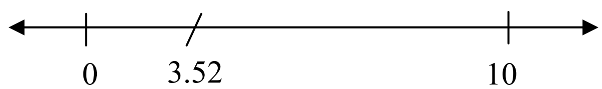
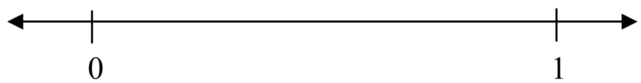
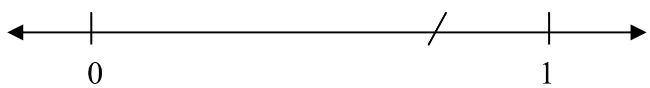
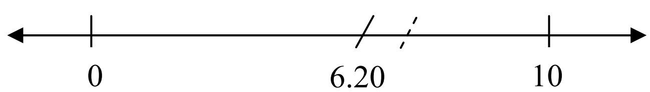
Assessment

An assessment was administered to students as a pretest, posttest, and retention test. This assessment measured procedural and conceptual knowledge of decimals and was modified and expanded from a past assessment that was shown to be reliable (Durkin & Rittle-Johnson, 2012). Sample items of each knowledge type are shown in Table 1 (see Appendix B for the complete assessment). The decimals varied in number of digits (tenths, hundredths, thousandths, and ten thousandths), magnitude (greater than or less than 0.5), and whether a zero was included.

Procedural knowledge items could be solved through the use or adaptation of a step-by-step solution procedure illustrated during the intervention. While students had not yet seen these solution procedures at pretest, I categorized these items as procedural knowledge at all assessment points for simplicity. The procedural items included four *familiar problem type* items, three *recognizable problem type* items, and eleven *novel problem type* items. Target values were specifically chosen to make it easy to recognize misconception errors. On the *familiar problem type*, students needed to place a decimal on a number line from 0 to 1, similar to the problems they completed during the intervention, with new values. On the *recognizable problems,* students needed to place a decimal on a number line from 0 to 10 in relation to another number marked on the number line (e.g., place 3.8 in relation to 3.52). Students had limited exposure to this type of problem during the intervention as only 6 worked examples contained this 0 to 10 number line. Most students were exposed to at least a couple of these worked examples during the intervention, but a few students did not get to these problems in the packet and therefore had no exposure to them. There were three novel problem types. One kind of *novel problem type* involved placing a number in the ten thousandths place on a number line from 0 to 1. On another kind of *novel problem type*, students needed to identify from a list which decimal was already marked on a number line from 0 to 1. Thus, students were asked to do the reverse of what they had done during the intervention. On the third kind of *novel problem type,* students needed to identify from a list which decimal was already marked on a number from 0 to 10 in relation to another marked number. Internal consistency on this measure across item types was good (α = .82, .85, and .87 at pretest, posttest, and retention test, respectively).

The conceptual items were designed to measure students' understanding of fundamental decimal concepts independent of the number line and the three 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 *density* items evaluated students' understanding that there are an infinite number of decimals that can come between any two numbers (e.g. between 0.5 and 0.6) (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). Each item was designed to contain an answer choice that fit a misconception error. Internal consistency on this measure across item types was also good (α = .88, .90, and .91 at pretest, posttest, and retention test, respectively).

Table 1: Sample Assessment Items

	Example Item	Scoring
Procedural Items *(familiar)* *(n = 4)*	Mark about where 0.9 goes on the number line. 	1 point for each if within one tenth of the correct placement in either direction on the number line
Procedural Items *(recognizable)* *(n = 3)*	The number line now goes from 0 to 10. 3.52 is marked. Mark where 3.8 goes. 	1 point for each if properly placed as greater or less than the marked number and within one tenth of the correct placement
Procedural Items *(novel 1)* *(n = 4)*	Mark about where 0.3481 goes on the number line. 	1 point for each if within one tenth of the correct placement in either direction on the number line
Procedural Items *(novel 2)* *(n = 4)*	What number tells about where the slash is on the number line? a) 0.76 b) 0.3 c) 0.08 d) 0.401 	1 point for each correct answer
Procedural Items *(novel 3)* *(n = 3)*	One number is already marked on the number line. What number tells about where the unmarked dashed slash is on the number line? a) 6.173 b) 6.8 c) 6.05 d) 0.45 	1 point for each correct answer
Conceptual Items *(magnitude)* *(n = 10)*	Circle the number that is greater: 0.87 0.835	1 point for each correct answer
Conceptual Items *(density)* *(n = 5)*	Write a decimal that comes between 0.5 and 0.6.	1 point for each correct answer
Conceptual Items *(role of zero)* *(n = 5)*	Circle all the numbers that are worth the same amount as 0.51: 0.5100 0.051 0.510 51	1 point for each correct answer

The assessment was created so that incorrect response patterns indicated different misconceptions. On multiple choice questions, each incorrect answer choice was designed to fit a particular misconception. On number line problems, a variety of decimals were selected to include instances where students could make errors based on each particular misconception type. This allowed incorrect answers to be easily categorized as fitting the *whole number* misconception, *role of zero* misconception, or *fraction* misconception.

Three additional measures were included to assess the strength of students' misconceptions. Confidence ratings were added to five items spread throughout the assessment: 1 conceptual item and 5 procedural items. These confidence rating items asked students how sure they were that they answered the item correctly on a scale from 1 (Not at all sure) to 5 (Very sure). These ratings helped determine whether students truly had misconceptions or were just randomly guessing, and were used to categorize students' answers. These kinds of ratings have not often been used in past research on misconceptions in math, but they have been used to evaluate misconceptions in physics (e.g., Murray, Schultz, Brown, & Clement, 1990) and in psychology (e.g., Taylor & Kowalski, 2004) with college students. In one study, confidence ratings were used to assess children's adherence to incorrect strategies when solving mathematical equivalence problems (McNeil & Alibali, 2005). Second, a "hidden decimal task" was used to determine how students generally thought of magnitude problems (Irwin, 2001; Resnick et al., 1989). This item asked students to identify whether 0.□ or 0.□□□□ was greater (the boxes represent numbers covered by pieces of paper) and why their answer was correct. The students could select 0.□, 0.□□□□, or "Can't tell" for their answer. The correct answer was "Can't tell" but knowing which way students answered illustrated the misconceptions they had when solving such problems. Third, students were asked to describe a general rule for how they can tell which decimal is bigger when making magnitude comparisons. However, the responses on this item were generally very low quality (e.g., "I looked at the numbers."), so this task was not included in any further analyses.

Procedure

On Day 1, students were given 35 minutes to complete the pretest within their math classrooms. On Days 2 and 3, all students received brief introductory lessons at the beginning of each class period. Upon completion of the lesson, students worked with their partner on their intervention packets. The class periods on Day 2 and 3 ranged from 55 minutes to 75 minutes, depending on the length of the students' regular math instruction. On Day 3, students started on the second half of the packet, regardless of where they had finished working on Day 2. This exposed most students to all the different problem types featured in the intervention. On Day 4, students completed a posttest that was isomorphic to the pretest. Finally, students completed the retention test 21 or 22 days after the posttest.

Coding

Assessment

Answers to items were scored for accuracy according to the criteria outlined in Table 1. When possible, assessment items were also coded for the three misconception errors previously described, and this is outlined in Table 2. There were 23 items on which a whole number misconception could be detected, 17 items on which a role of zero misconception could be detected, and 16 items on which a fraction misconception could be detected. The proportion of misconception errors was calculated by dividing the number of misconception errors made by the total possible number of misconception errors of that type (see Table 2).

Table 2: Assessment Misconception Coding

Item Type	Example Item(s)	Whole Number	Role of Zero	Fraction
Procedural Mark on number line from 0-1	Mark about where 0.9 goes on the number line. Mark about where 0.3481 goes on the number line.	Incorrectly placed tenths near 0, hundredths in the middle, or thousandths near 1 (e.g., placing 0.9 on the first third of the line)	Decimal with a zero in the tenths placed as if the zero was not there (e.g., placing 0.07 at 0.7)	Incorrectly placed tenths near 1, hundredths in the middle, or thousandths near 0 (e.g., placing 0.3 on the last third of the line)
Procedural Mark on number line from 0-10	The number line now goes from 0 to 10. 3.52 is marked. Mark where 3.8 goes. 4.35 is marked. Mark where 4.842 goes.	Placed within one whole of the already marked decimal on the wrong side viewed as whole numbers (e.g., placing 3.8 between 2.52 and 3.52)	A decimal with a zero in the tenths placed as if the zero was not there (e.g., placing 9.05 on the wrong side of 9.2 as if it was 9.5)	Placed within one whole of the already marked decimal on the wrong side viewed as fractions (e.g., placing 4.842 between 4.25 and 4.35)
Procedural Choose correct decimal	What number tells about where the slash is on the number line? 0.76 0.3 0.08 0.401 with slash at 0.76	One answer was designed to be chosen if viewed as whole numbers (e.g., 0.401)	One answer was designed to be chosen if ignored zero in tenths place (e.g., 0.08)	One answer was designed to be chosen if viewed as fractions
Conceptual Magnitude	Circle the decimal that is greater: 0.87 0.835 0.3 0.92	Incorrectly circling an answer that is greater if viewed as whole numbers (e.g., circling 0.835)	Incorrectly circling an answer that is greater if the zero in the tenths is ignored	Incorrectly circling an answer that is greater if viewed as fractions (e.g., circling 0.3)
Conceptual Density	Write a decimal that comes between 0.14 and 0.148.	An incorrect answer between the given numbers viewed as whole numbers (e.g., 0.16)	Not applicable	Not applicable
Conceptual Role of Zero	Circle all the numbers that are worth the same amount as 0.51: 0.5100 0.051 0.510 51	Ignoring the "0." and treating the decimal like a whole number (e.g., circling 51)	Circling one or more of the incorrect answer choices	Not applicable

In addition to scoring accuracy, students' confidence ratings were scored from five Likert scale items on the assessment (on a scale from 1 to 5). These confidence ratings were then categorized as high confidence (3-"Somewhat sure" or above) or low confidence (2-"A little sure" or below). Students' mean confidence level on each assessment was calculated. Also, items with a confidence scale were categorized as a high confidence error, a low confidence error, a high confidence correct response, or a low confidence correct response. Then the mean proportion of each of these types of responses occurring was calculated for each student. For example, if a student always chose 4 as their confidence level and got 3 of these items right and 2 of these items wrong, they would have mean proportions of .60 for high confidence correct responses, .40 for high confidence errors, .00 for low confidence correct responses, and .00 for low confidence errors.

Intervention

Several pieces of information were calculated and coded from the intervention including students' written explanations, students' completion of explanations and practice problems, and students' accuracy when filling in blanks within the worked examples.

Data Analysis

Missing Data

There was a small amount of missing data at each of the assessment time points. At pretest 5% of the data was missing, at posttest 3% of the data was missing, and at retention test 6% of the data was missing. This missing data was all due to students being absent from school on the day of the assessment. I did not impute this missing data as simulation studies have shown that it does not matter how you handle missing data when the proportion missing is this low (Schafer & Graham, 2002).

Multilevel Models

Because students worked with a partner, their performance may not be independent of each other, and multilevel modeling must be used. I calculated intra-class correlations, controlling for the predictor variables, to test for non-independence in posttest and retention test scores between partners (Kenny, Kashy, & Cook, 2006). Intra-class correlations ranged from 0.01 to 0.24, and there were significant intra-class correlations on many measures. Consequently, I used multilevel linear models to account for this non-independence in the data and specified the use of restricted maximum likelihood estimation (REML) (Kenny et al., 2006).

I used a two-level model. The first level of the model, the individual level, measured the effect of the prevalence of the individual's pretest misconceptions. Pretest accuracy was not included as a predictor in the model because its high correlation with prior misconceptions led to multicollinearity problems. The prevalence of pretest misconceptions was mean centered in the model. The second level of the model, the dyad level, measured the effect of experimental condition and grade level. Grade level was contrast coded so that Grade 4 was coded as -1 and Grade 5 was coded as 1. I specified the *Incorrect-Sequential* condition as the referent condition because this allowed me to test both the effect of seeing incorrect examples relative to only seeing correct ones and the effect of using comparison with incorrect examples relative to seeing such examples sequentially. This resulted in the effect of condition being captured by two variables. One variable indicated the difference between the *Correct-Sequential* and *Incorrect-*

Sequential conditions, and the other variable indicated the difference between the *Incorrect-Compare* and *Incorrect-Sequential* conditions. To test the difference between the *Correct-Sequential* and *Incorrect-Compare* conditions, a Wald test (similar to an incremental F test) was used to examine whether the parameter estimates for these conditions were significantly different from one another. Separate models were run to evaluate the data from the posttest and retention test because standard multi-level models do not accommodate multiple outcome measures in the same model. Also, similar past studies have indicated differences in the effect of incorrect examples on an immediate posttest and a delayed retention test (Durkin & Rittle-Johnson, 2012, in preparation).

In preliminary analyses, I explored several measures of prior knowledge, including pretest accuracy and prevalence of misconception errors at pretest. I examined whether pretest accuracy or prevalence of misconception errors at pretest was a more powerful moderator of condition effects. These two measures were highly correlated (r = -0.79); however, the prevalence of misconception errors was a more consistent moderator variable across time points and knowledge types. Thus, I used the prevalence of misconception errors at pretest to examine the effect of prior knowledge on outcomes. I used the prevalence of misconception errors rather than accuracy at pretest because strong misconceptions can play an important role in learning, particularly in a domain with such widespread misconceptions (Stafylidou & Vosniadou, 2004). To test for whether prior misconception errors moderated the effects of condition, I included two cross-level interaction terms between the prevalence of misconceptions at pretest and each condition indicator variable.

CHAPTER III

RESULTS

First, I overview students' knowledge at pretest. Next, I report the effects of condition and prior misconceptions at posttest and retention test. To help understand effects of condition, I follow this with an exploration of how condition affected performance during the intervention.

Pretest Knowledge

At pretest, students had variable conceptual and procedural knowledge, and misconception errors were prevalent (see Table 3). On average, students were accurate on about 31% of conceptual knowledge items and 22% of procedural knowledge items. Students were also making errors based on misconceptions about 32% of the time. A histogram of the distribution of the prevalence of misconception errors indicated wide variability (see Figure 2). Importantly, there were no significant differences between conditions in terms of conceptual and procedural knowledge scores and prevalence of misconception errors at pretest (p's > 0.500).

Table 3: Performance on Outcome Measures By Condition

	Pretest		Posttest		Retention Test	
Outcome	*M*	*SD*	*M*	*SD*	*M*	*SD*
Percent Correct						
Conceptual Knowledge						
Incorrect-Compare	30	(22)	34	(26)	38	(28)
Incorrect-Sequential	31	(21)	37	(27)	42	(29)
Correct-Sequential	31	(23)	37	(25)	40	(27)
Procedural Knowledge						
Incorrect-Compare	22	(18)	34	(24)	35	(24)
Incorrect-Sequential	22	(21)	37	(24)	38	(28)
Correct-Sequential	22	(21)	36	(26)	37	(26)
Percent Misconception Errors						
Whole Number						
Incorrect-Compare	43	(16)	42	(21)	39	(23)
Incorrect-Sequential	41	(18)	38	(21)	37	(23)
Correct-Sequential	41	(18)	38	(20)	37	(22)
Role of Zero						
Incorrect-Compare	45	(16)	45	(20)	41	(21)
Incorrect-Sequential	46	(17)	43	(21)	40	(21)
Correct-Sequential	45	(14)	42	(19)	40	(19)
Fraction						
Incorrect-Compare	10	(11)	12	(14)	12	(16)
Incorrect-Sequential	10	(11)	12	(15)	12	(14)
Correct-Sequential	11	(11)	12	(12)	12	(12)
Overall						
Incorrect-Compare	33	(10)	33	(13)	31	(14)
Incorrect-Sequential	32	(11)	31	(13)	30	(15)
Correct-Sequential	32	(10)	31	(13)	30	(14)

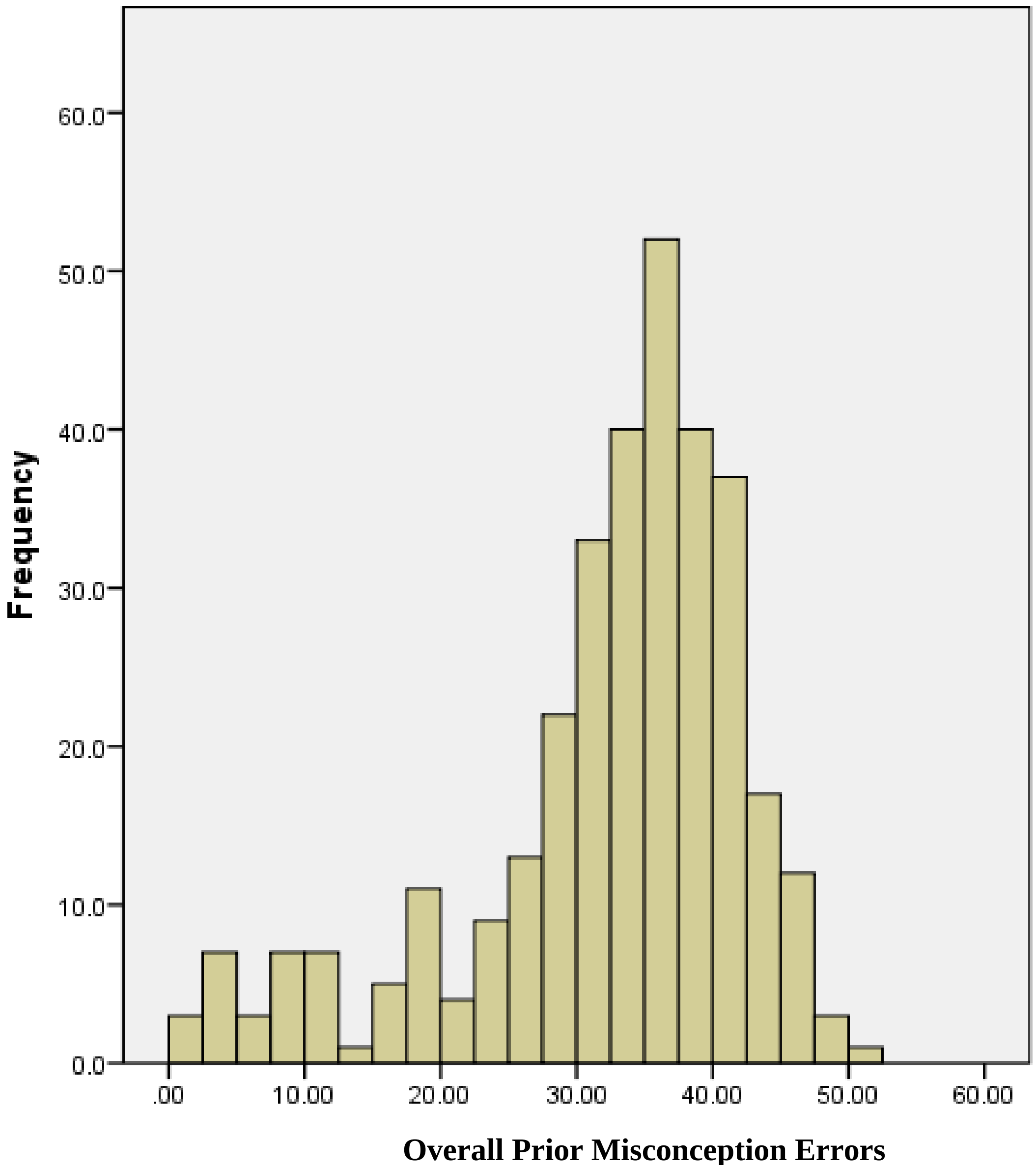


Figure 2: Histogram of the Prevalence of Prior Misconception Errors

In addition to the prevalence of misconceptions, several other measures were used to assess students' misconceptions. First, I examined students' confidence in their responses at pretest using the results from the confidence scales. Overall, when making errors, students were almost as likely to express high confidence (37% of the time) as low confidence (45% of the time). When answering correctly, students expressed low confidence 7% of the time and high confidence 11% of the time. This suggests that many errors may not reflect deeply held misconceptions. On the "hidden decimal" task, overall students were answering correctly 33% of the time, displaying a whole number misconception 53% of the time, and displaying a fraction misconception 12% of the time. Thus misconceptions, particularly the whole number misconception, were prevalent at pretest, with varied confidence when making errors.

Effects of Condition and Prior Misconceptions on Conceptual Knowledge

Inspection of means suggested that condition did not impact performance on conceptual items (see Table 3). Indeed, accuracy did not improve much; accuracy at posttest was about 36%, which was a 5% increase from the pretest. Accuracy at retention test was about 40%, which was a 9% increase from the pretest and a 4% increase from the posttest.

The benefits of comparison or incorrect examples might depend on students' prior knowledge. To explore this possibility, my analysis models included condition by prior misconception errors interaction terms. The two-level linear models included a dummy coded variable for being in the *Correct-Sequential* condition, a dummy coded variable for being in the *Incorrect-Compare* condition, prevalence of prior misconception errors, grade, and the interactions between the conditions and prior misconceptions. Table 4 displays the results for the dependent variables, including conceptual knowledge at posttest and retention test (columns 1 and 2), procedural knowledge at posttest and retention test (columns 3 and 4), and the prevalence of misconceptions at posttest and retention test (columns 5 and 6). These models confirmed that there were no main effects for condition on any measure, but that prior misconceptions moderated the effect of condition on several outcomes. I discuss each outcome in turn.

Table 4: Parameter Estimates for Outcomes

Variable	Posttest Conceptual	Retention Conceptual	Posttest Procedural	Retention Procedural	Posttest Misconceptions	Retention Misconceptions
Intercept	38.04 (1.63)***	41.53 (1.73)***	38.07 (2.11)***	37.35 (2.03)***	30.70 (0.87)***	29.86 (0.91)***
Condition (reference = Incorrect-Sequential)						
Correct-Sequential	0.23 (2.34)	-0.54 (2.48)	-1.19 (3.02)	-0.09 (2.90)	-0.51 (1.25)	-0.53 (1.31)
Incorrect-Compare	-2.25 (2.34)	-1.95 (2.48)	-2.51 (3.00)	-0.09 (2.89)	1.27 (1.25)	-0.15 (1.31)
Condition*Prior Misconceptions						
Correct-Sequential	0.43 (0.22) $^{\tau}$	0.21 (0.23)	0.22 (0.26)	0.29 (0.25)	-0.13 (0.12)	-0.07 (0.12)
Incorrect-Compare	-0.02 (0.23)	0.002 (0.24)	0.45 (0.28)	0.54 (0.27)*	0.05 (0.12)	0.12 (0.13)
Prior Misconceptions	-1.50 (0.16)***	-1.40 (0.16)***	-1.13 (0.18)***	-1.28 (0.17)***	0.76 (0.08)***	0.66 (0.08)***
Grade	10.20 (1.03)***	12.84 (1.09)***	9.63 (1.31)***	10.80 (1.26)***	-4.55 (0.55)***	-5.77 (0.57)***

Note. Unstandardized coefficients are shown with standard errors in parentheses. Frequency of prior misconception errors was grand mean centered and grade was contrasted coded as -1 and +1.
$^{\tau}p \leq .06$, $*p < .05$, $**p < .01$, $***p < .00$

Posttest

At posttest, the effect of condition on conceptual knowledge somewhat depended on the prevalence of students' prior misconception errors (see Table 4, column 1). In particular, students with infrequent misconceptions benefited more from the *Incorrect-Sequential* than the *Correct-Sequential* condition. Then as the frequency of misconceptions increased, the *Correct-Sequential* condition became more effective ($\beta = 0.43$, $p = .053$), although this effect was marginal. For each point increase in the frequency of prior misconception errors, students did about a half percentage point better in the *Correct-Sequential* condition relative to the *Incorrect-Sequential* condition. The prevalence of prior misconception errors also moderated the difference between the *Correct-Sequential* and *Incorrect-Compare* conditions. Post-hoc Wald tests suggested that students who had infrequent prior misconception errors tended to benefit more from the *Incorrect-Compare* condition than the *Correct-Sequential* condition, although this effect was marginal ($\chi^2 = 5.05$, $p = .080$). Students in the *Incorrect-Compare* and *Incorrect-Sequential* conditions performed similarly, regardless of prior knowledge ($\beta = -0.02$, $p = .916$). Also, exploratory analyses indicated that the interactions between conditions and prior misconception errors were stronger for fourth-grade students than fifth-grade students.

To better understand how prior misconception errors moderated the effect of condition, regression lines were created for each condition using the parameter estimates in Table 4 (see Figure 3a). Students with infrequent prior misconception errors (i.e., higher prior knowledge) are shown on the left side of the graph, with students with frequent prior misconception errors (i.e., lower prior knowledge) are shown on the right side. Descriptively, students with infrequent prior misconception errors performed best in the *Incorrect-Sequential* condition, while students with frequent prior misconception errors performed a bit better in the *Correct-Sequential* condition. I also conducted an analysis of the region of significance to identify the upper and lower bounds of the moderator at which the conditions were significantly different (Preacher, Curran, & Bauer, 2006). This preliminary analysis of the region of significance indicated that the frequency of prior misconception errors did not ever fall within the region of significance, and the simple slopes tests were not significant for the the minimum, mean, or maximum values of the moderator (p's > .07). Students in the *Incorrect-Compare* condition performed very similarly to those in the *Incorrect-Sequential* condition.

To further interpret this interaction, I categorized students' prevalence of prior misconception errors into quartiles. The prevalence of prior misconception errors ranged from 0 to 28.4% for the first quartile, 28.5 to 34.7% for the second quartile, 34.9 to 39.3% for the third quartile, and 39.4 to 50.9% for the fourth quartile. I then ran the same model, without the prior misconceptions variable and interaction terms, for students in the first quartile and the fourth quartile separately. When running the model with students in the first quartile (i.e., the students with infrequent prior misconception errors), students were scoring about 9 points better in the *Incorrect-Sequential* condition than in the *Correct-Sequential* condition, although this difference was not significant ($\beta = -9.28$, $p = .132$). When running the model with students in the fourth quartile (i.e., the students with frequent prior misconception errors), students were scoring about 4 points better in the *Correct-Sequential* condition than in the *Incorrect-Sequential* condition, although this difference was not significant ($\beta = 3.92$, $p = .304$). To summarize, students with infrequent prior misconception errors seemed to benefit slightly from viewing incorrect examples; however, this effect was only marginal and very small.

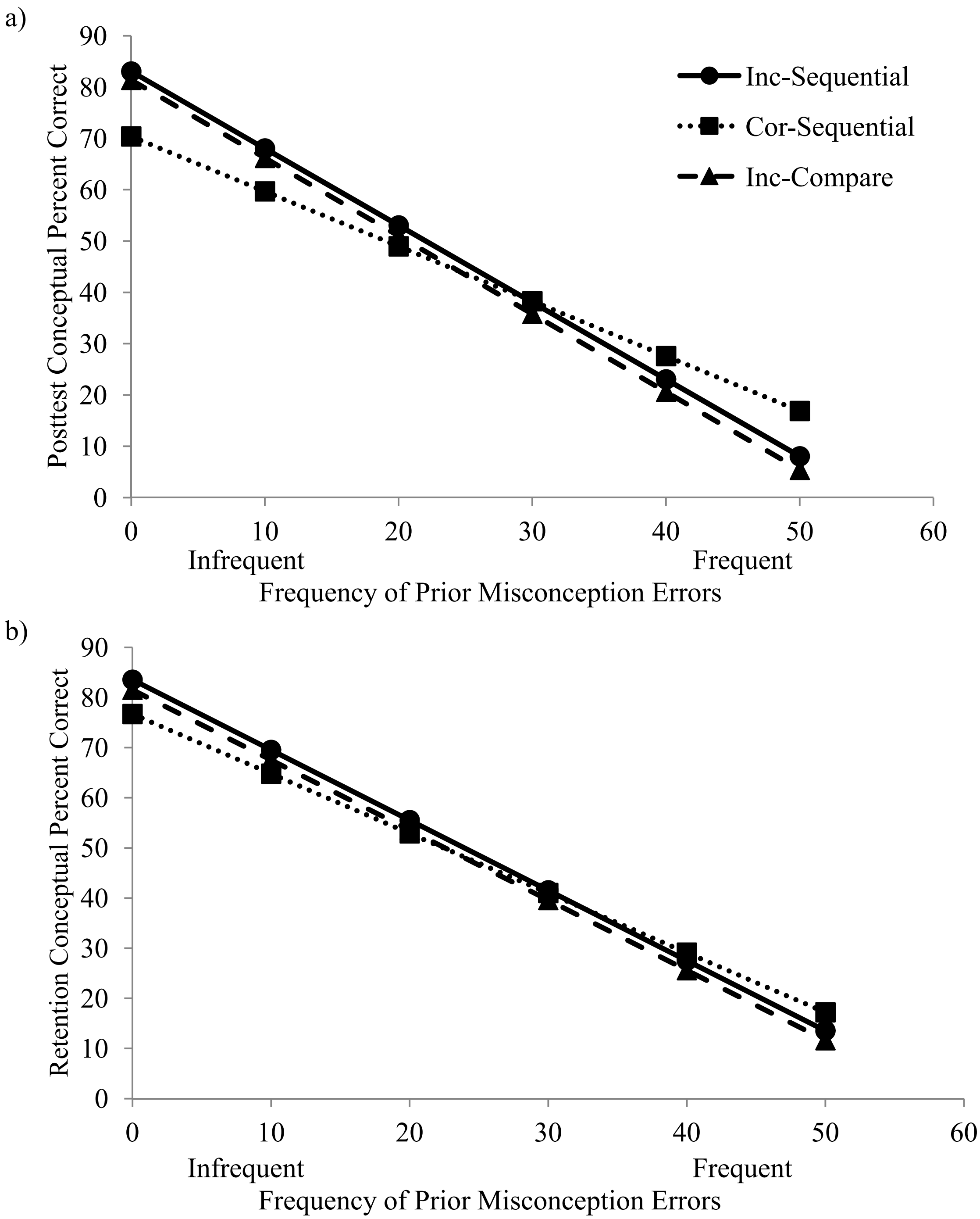


Figure 3: Condition by Prior Misconception Errors Interactions for Conceptual Knowledge on a) Posttest and b) Retention Test

Note: The x-axis has been relabeled from the original mean centered values to the raw values for ease of interpretation. The regression equation for predicting conceptual knowledge at posttest was: Y' = 38.04 – 1.50*PriorMisconceptions + 0.23*CorrectSequential – 2.25*IncorrectCompare + 0.43*CorrectSequential*PriorMisconceptions – 0.02*IncorrectCompare*PriorMisconceptions. At retention it was: Y' = 41.53 – 1.40*PriorMisconceptions – 0.54*CorrectSequential – 1.95*IncorrectCompare + 0.21*CorrectSequential*PriorMisconceptions + 0.002*IncorrectCompare*PriorMisconceptions.

Retention Test

Students completed a retention test three weeks after the intervention. There were no effects for condition nor did prior misconception errors moderate the effects of condition (see Table 4, column 2). For consistency across outcomes, regression lines were created for each condition using the parameter estimates in Table 4 (see Figure 3b). The pattern of findings was the same at retention test as at posttest, but differences based on condition were very small across prior levels of misconception errors.

Effects of Condition and Prior Misconceptions on Procedural Knowledge

In addition to assessing students' conceptual knowledge, I also assessed their procedural knowledge. Inspection of means again suggested that condition did not have a significant impact on accuracy (see Table 3). There were modest gains over time. Accuracy at posttest was about 36%, which was a 14% increase from the pretest. Accuracy at retention test was about 37%, which was a 15% increase from the pretest.

Posttest

At posttest, there were no main effects for condition, nor did prior misconception errors moderate the effect of condition (see Table 4, column 3). For consistency across outcomes, regression lines were created for each condition using the parameter estimates in Table 4. A graph of these regression lines illustrates that students with infrequent prior misconception errors seemed to perform best in the *Incorrect-Sequential* condition, although not significantly so, with few differences between conditions for students who had frequent prior misconception errors (Figure 4a).

a) b)

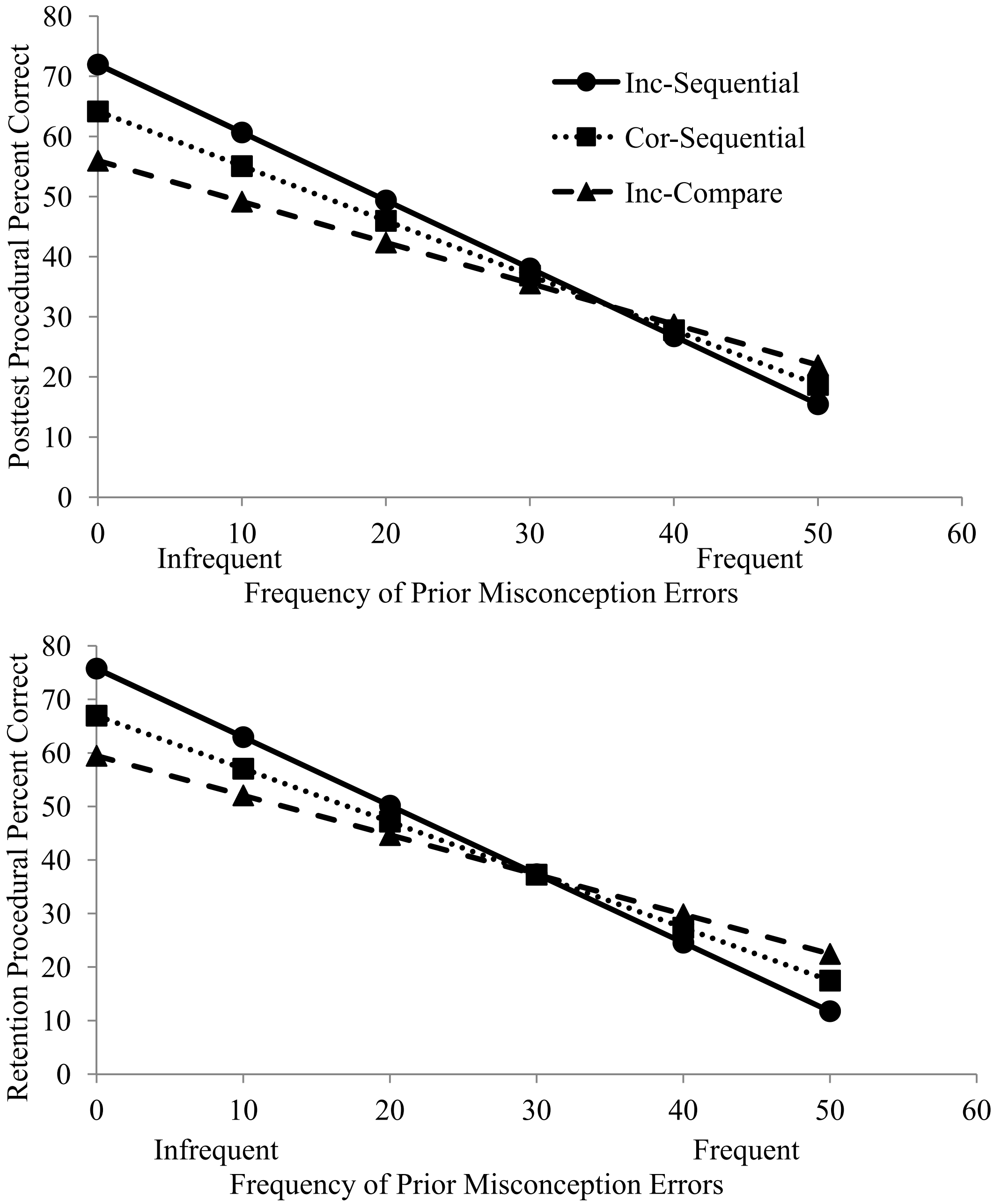


Figure 4: Condition by Prior Misconception Errors Interactions for Procedural Knowledge on a) Posttest and b) Retention Test

Note: The x-axis has been relabeled from the original mean centered values to the raw values for ease of interpretation. The regression equation for predicting conceptual knowledge at posttest was: Y’ = 38.07 – 1.13*PriorMisconceptions – 1.19*CorrectSequential – 2.51*IncorrectCompare + 0.22*CorrectSequential*PriorMisconceptions + 0.45*IncorrectCompare*PriorMisconceptions. At retention it was: Y’ = 37.35 – 1.28*PriorMisconceptions – 0.09*CorrectSequential – 0.09*IncorrectCompare + 0.29*CorrectSequential*PriorMisconceptions + 0.54*IncorrectCompare*PriorMisconceptions.

Retention Test

At retention test, the effect of condition on procedural knowledge was moderated by the prevalence of students' prior misconception errors (see Table 4, column 4). In particular, students with infrequent prior misconception errors benefited more from the *Incorrect-Sequential* than the *Incorrect-Compare* condition. Then as the frequency of misconception errors increased, the *Incorrect-Compare* condition became more effective (β = 0.54, p = .046). For each point increase in the frequency of prior misconception errors, students did a half percentage point better in the *Incorrect-Compare* condition relative to the *Incorrect-Sequential* condition. Relative to the *Correct-Sequential* condition, the effects for the *Incorrect-Sequential* condition were similar, but even more modest. Students who had infrequent prior misconception errors also benefited more from the *Incorrect-Sequential* condition than the *Correct-Sequential* condition, with the *Correct-Sequential* condition becoming more effective as the frequency of prior misconception errors increased, but this was not significant (β = 0.29, p = .254). In addition, post-hoc Wald tests indicated that there were no differences between the *Correct-Sequential* and the *Incorrect-Compare* conditions, nor did prior misconception errors moderate the effect of condition ($\chi^2 < 0.01$, p = .999 and $\chi^2 = 4.08$, p = .130, respectively). Also, exploratory analyses indicated that the interactions between conditions and prior misconception errors were stronger for fourth-grade students than fifth-grade students.

To better understand how prior misconception errors moderated the effect of condition, regression lines were created for each condition using the parameter estimates in Table 4. As shown in Figure 4b, students with infrequent prior misconception errors performed best in the *Incorrect-Sequential* condition, while students with frequent prior misconception errors performed a bit better in the *Incorrect-Compare* condition. However, condition differences were very slight for students with frequent misconception errors. A preliminary analysis of the region of significance indicated that the frequency of prior misconception errors only fell within the region of significance for students with infrequent prior misconception errors. The simple slopes test was marginally significant for the minimum value of the moderator (p = .057), but not for the mean or maximum values of the moderator (p = .976 and p = .083, respectively).

To further interpret this interaction, I categorized students' prevalence of prior misconception errors into quartiles as before. I then ran the same model, without prior misconception errors and interaction terms, for students in the first quartile and the fourth quartile separately. When running the model with students in the first quartile (i.e., the students with infrequent prior misconception errors), students were scoring about 11 points higher in the *Incorrect-Sequential* condition than in the *Incorrect-Compare* condition, although this difference was not significant (β = -10.56, p = .121). When running the model with students in the fourth quartile (i.e., the students with frequent prior misconception errors), students were scoring about 4 points better in the *Incorrect-Compare* condition than in the *Incorrect-Sequential* condition, although this difference was not significant (β = 3.74, p = .540). To summarize, students with infrequent prior misconception errors seemed to benefit from viewing incorrect and correct examples sequentially; however, this effect was very small and unreliable across time points.

Effects of Condition and Prior Misconceptions on Later Misconceptions

As previously mentioned, I wanted to assess students' misconceptions on the post and retention tests in addition to assessing their conceptual and procedural knowledge. This was accomplished by measuring the prevalence of misconception errors students made. I also included two new misconception measures. First, the strength of misconception errors was assessed by examining students' confidence on a subset of items. Second, data from the "hidden decimal" task was used to examine students' misconceptions when generally considering decimal magnitude.

Prevalence of Misconceptions

Inspection of means suggested that the overall prevalence of misconceptions did not change much over time, and condition did not impact the prevalence of misconceptions (see Table 3). At posttest, students were making misconception errors about 32% of the time, which was similar to the pretest. At retention test, students were making misconception errors about 30% of the time, which was a 2% decrease from the pretest and the posttest.

To explore whether prior knowledge moderated the effects of condition, the dependent variables were now the overall prevalence of misconception errors at posttest and at retention test. The overall prevalence of misconception errors was used instead of dividing these errors into conceptual and procedural errors because the pattern of results was the same across both types of misconception errors.

There were no main effects for any condition on the prevalence of misconception errors made at posttest or retention test. Prior misconception errors also did not moderate the effects of condition at either time point (Table 4). A similar pattern of results was found when looking specifically at whole number, role of zero, and fraction misconception errors separately.

These findings suggest that the previously reported conditions by prior misconception errors interactions for conceptual and procedural knowledge were not being driven by differences in the rates of misconception errors. The prevalence of misconception errors was similar across conditions, regardless of the prevalence of prior misconception errors.

Consequently, the previously reported interactions were due to differences in the frequency of random errors students were making. Random errors were incorrect responses that did not fit a particular misconception type (e.g., placing 0.9 in the middle of the number line from 0 to 1).

Strength of Misconceptions

In addition to measuring the prevalence of misconception errors on the assessment, I also wanted to more preciously measure the strength of students' misconceptions. To accomplish this, I classified students' answers on each of the five confidence scale items, as previously mentioned. I then calculated the mean percentage of how often students' answers fell into each of these 4 categories (see Table 5).

Table 5: Percentage of Each Confidence Response Type

	Pretest		Posttest		Retention Test	
Outcome	*M*	*SD*	*M*	*SD*	*M*	*SD*
Percent of Each Response Type						
High Confidence Error	37	(29)	40	(31)	38	(32)
Low Confidence Error	45	(32)	30	(30)	29	(31)
Low Confidence Correct	07	(11)	08	(13)	08	(14)
High Confidence Correct	11	(19)	22	(28)	24	(32)
Percent of Students Making Each Response at Least Half of the Time						
High Confidence Error	32		36		35	
Low Confidence Error	46		29		26	
Low Confidence Correct	< 1		1		2	
High Confidence Correct	5		16		20	

These results indicated that after the intervention, students overall were decreasing their amount of low confidence errors and increasing their amount of high confidence correct answers. However, the percentage of high confidence errors and low confidence correct answers did not change much over time. Examining the percentage of students who gave each response type at least half of the time at pretest, posttest, and retention test, resulted in similar findings (Table 5). This suggests that the intervention was not correcting strongly held misconceptions (i.e., high confidence in errors), and instead the intervention helped decrease guessing (i.e., low confidence in errors).

The "hidden decimal" task was also designed to uncover what misconceptions students might have when comparing decimals' magnitudes. On average, students did not often solve this task correctly at posttest or retention test, and they frequently made whole number misconception errors (Table 6).

Table 6: Percentage of Response Types on Hidden Decimal Task

	Posttest Correct		Posttest Whole Number Misconception		Posttest Fraction Misconception		Retention Correct		Retention Whole Number Misconception		Retention Fraction Misconception	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Overall	35	(48)	51	(50)	15	(35)	41	(49)	41	(49)	18	(39)
Incorrect-Sequential	33	(47)	53	(50)	14	(35)	39	(49)	41	(49)	19	(40)
Correct-Sequential	38	(49)	46	(50)	16	(37)	44	(50)	37	(48)	18	(39)
Incorrect-Compare	34	(48)	53	(50)	13	(34)	38	(49)	44	(50)	17	(38)

Effect of Condition on Intervention Activities

To help understand how condition may have impacted learning, I examined students' work during the intervention. As previously mentioned, students studied packets of worked examples, answered explanation prompts about the examples, and solved practice problems.

Amount of Material Completed

Students spent most of their time during the intervention studying and explaining worked examples with a partner and solving practice problems. The percentage of materials completed is presented in Table 7. Overall, the amount of material completed differed between conditions. In particular, students in the *Incorrect-Compare* condition completed less of the intervention than the other two conditions, which did not differ from one another. Also, the frequency of prior misconception errors did not moderate the effect of condition for the amount of material completed.

In addition to these tasks, students were also asked to fill-in blanks in the worked examples during the intervention to ensure that they were processing the worked examples. These blanks were meant to be easy for students to complete; however, students only filled-in the blanks correctly 74% of the time, and this did not vary much by condition (*M*s = 73%, 74%, and 75% of blanks filled-in correctly for the *Incorrect-Sequential*, *Correct-Sequential*, and *Incorrect-Compare* conditions, respectively). This was perhaps due to students' low prior knowledge, and indicates that students had some difficulty comprehending the intervention materials.

Table 7: Percentage of Intervention Completed By Condition

	Explanation Prompts		Practice Problems	
Condition	*M*	*SD*	*M*	*SD*
Overall	65.53	(22.73)	55.79	(22.62)
Incorrect-Sequential	70.94	(21.17)	60.75	(23.12)
Correct-Sequential	69.62	(21.53)	59.82	(23.01)
Incorrect-Compare	56.23**	(22.62)	47.00**	(22.34)

** Differed from other two conditions at the $p < .005$ level.

Summary

Before the study began, students had varied, but limited, knowledge of decimals and prevalent misconceptions. Condition had minimal impact on learning, but there were some modest effects of condition that depended on students' levels of prior misconception errors. Students with relatively infrequent misconceptions prior to the intervention performed best in the *Incorrect-Sequential* condition. For these students, being in the *Incorrect-Sequential* condition led to slightly better conceptual knowledge than being in the *Correct-Sequential* condition and better procedural knowledge than being in the *Incorrect-Compare* condition. However, these effects were small and were not reliable at both posttest and retention test. As the frequency of students' misconception errors increased, differences between conditions decreased with few detectable differences of condition for students making frequent prior misconception errors. Analyses of misconception errors at posttest and retention test suggested that misconceptions were not changing much after the intervention, regardless of condition and prior misconception errors. Initial analyses of performance during the intervention have not provided many explanatory clues, although there is some indication that students were struggling to comprehend some of the examples across conditions.

CHAPTER IV

DISCUSSION

Students learned about decimal magnitude under one of three conditions: comparing incorrect and correct examples, seeing incorrect and correct examples sequentially, or seeing only correct examples sequentially. Overall, condition did not have a substantial impact on conceptual knowledge, procedural knowledge, or misconceptions. However, prior knowledge did moderate the impact of condition. Students with infrequent misconception errors prior to the intervention seemed to perform best in the *Incorrect-Sequential* condition, although these effects were small and unreliable. However, students with frequent misconception errors seemed to perform more similarly across conditions. In this discussion, I integrate these results with findings from past research, discuss the assessment and correction of misconceptions, and suggest future directions for research in this area.

Integrating with Past Research on Incorrect Examples and Comparison

In the current study, the prevalence of prior misconception errors was an important factor for learning from incorrect examples. There were no significant effects of condition on any of the outcome measures, and knowledge growth was fairly limited across conditions. For example, although low confidence errors diminished after the intervention, high confidence errors did not.

It is important to note that past research on incorrect examples and prior knowledge has not focused on the prevalence of misconceptions as a measure of prior knowledge (e.g., Durkin & Rittle-Johnson, 2012; Große & Renkl, 2007). However, it is necessary to assess the prevalence and strength of misconceptions beyond assessing accuracy because low accuracy does not necessarily mean that students have strongly held misconceptions (McNeil & Alibali, 2005). Therefore, measuring the prevalence and strength of misconceptions is crucial because prior misconceptions seem to play an important role in determining when students are best prepared to learn from incorrect examples and comparison.

For students with frequent prior misconception errors, condition did not matter much for performance. This was possibly due to these students' lack of prior knowledge. If these students did not have sufficient prior knowledge to process the worked examples and give high quality explanations, then the manipulations of what *kinds* of examples were presented (incorrect versus correct) and *how* they were presented (compared vs. sequentially) would most likely not influence performance. Past work on using incorrect examples with novices in science domains has indicated that students learn best when presented with incorrect examples early in the learning process (e.g., Alvermann & Hague, 1989; Diakidoy et al., 2003; Eryilmaz, 2002). Yet in these studies, students were often participating in a structured discussion led by an instructor or using a refutation text in the context of a larger lesson. In the current study, students were provided with limited instruction from the researchers and spent most of their time explaining examples and solving practice problems with a partner in their classroom. Consequently, students with frequent misconceptions might learn more from incorrect examples if they were used in a more structured, extensive lesson, or if they were used in a more individualized, one-on-one tutoring setting. This additional guidance may be necessary for students with frequent misconceptions to make the most of learning with incorrect examples. This would be consistent with evidence that suggests that lower knowledge students benefit from strong external guidance (Kalyuga, 2007).

However, when students had infrequent prior misconception errors, condition did influence performance, although not always consistently or reliably between assessment points and knowledge types. In particular, the *Incorrect-Sequential* condition was better for students who did not have frequent misconceptions. For these students, studying incorrect and correct examples sequentially emerged as more effective than comparing them and slightly more effective than studying only correct examples sequentially. This led to two implications for these students: 1) it was somewhat better to see incorrect and correct examples, but 2) it was better to study these examples sequentially rather than comparing them.

For students with infrequent prior misconception errors, being exposed to incorrect and correct examples was slightly more beneficial than being exposed to only correct examples. This replicated past findings that higher knowledge students perform best when studying incorrect and correct examples (Große & Renkl, 2007). In the current study, the presence of incorrect examples particularly improved these students' conceptual knowledge. Students with infrequent misconceptions, and higher prior knowledge, may have been able to identify the critical features of correct and incorrect examples and identify their previous errors as incorrect (VanLehn, 1999). Importantly, these features may have been illustrating essential domain principles. This may also have helped these students reduce competition between conflicting misconceptions and correct concepts (Van den Broek & Kendeou, 2008).

This finding contradicts past work that suggests all students, regardless of prior knowledge, benefit from studying incorrect and correct examples relative to only correct examples (Durkin & Rittle-Johnson, 2012). The students in both the current and previous study had similar prior knowledge, with similar accuracy and misconception error rates. However, in the prior study, all students compared examples, and it is possible that results similar to the current study's findings would have emerged if students had studied examples sequentially. Additionally, when students *must* compare examples, it is possible that all students can glean more information from comparing incorrect and correct examples than from comparing only correct ones.

Although the presence of incorrect examples was slightly beneficial for students with infrequent prior misconception errors, there was no benefit of comparison. In fact, sequentially viewing incorrect and correct examples was better than comparing them. This was contrary to what might be predicted from past research on comparison (Rittle-Johnson et al., 2009). Students need sufficient prior knowledge to learn from comparison; however, what constitutes sufficient prior knowledge is unclear (Rittle-Johnson et al., 2009). Indeed, past research examining prior knowledge does not often elaborate on what exactly distinguishes high and low knowledge learners. It is possible that most of the current sample did not have sufficient prior knowledge, which may have led to comparison being too overwhelming, even for students with infrequent prior misconceptions. If this was the case, then one would expect that these students would learn better from studying examples sequentially rather than comparing them. These students still had some misconceptions at pretest and did not have very high pretest accuracy. As a result, these students were possibly struggling to properly align examples. As previously mentioned, it is difficult for students to align two unfamiliar examples (e.g., Gentner et al., 2003), and comparison is thought to be beneficial because it can help learners align examples to see their underlying structure (e.g., Loewenstein et al., 1999). If these students were struggling to align examples, then they may not have been able to take advantage of comparison. The completion rate of the intervention materials supports this idea because students who compared completed less of the intervention. In addition, comparison may not have been as useful for students as originally anticipated due to students being unprepared to learn from the current design. Students spent most of their time working with a partner, with limited guided instruction. Consequently, students in the *Incorrect-Compare* condition could have possibly performed better if they were better scaffolded to make meaningful comparisons. Also, for students with infrequent prior misconception errors, viewing these examples sequentially may have been particularly helpful for procedural knowledge because it allowed students the chance to focus on one procedure at time. This could have made it easier for students to process and understand the procedure presented in the worked example because their working memory was not overwhelmed with unfamiliar information from two different sources. If students were overwhelmed by comparison, they may not have been able to notice the critical features of the illustrated correct and incorrect procedures. Finally, students who studied examples sequentially were able to complete more of the intervention materials, so they were exposed to more correct and incorrect procedures during the intervention than students who compared. This exposure to a greater number of procedures may have improved these students' procedural knowledge.

Overall, condition had minimal impact on performance at posttest and retention test. This points to the difficulty in supporting conceptual change when students have misconceptions. When misconceptions were infrequent, sequential study of incorrect and correct examples was best for students' performance. However, even these benefits were not consistent across posttest and retention test, and knowledge did not change much for students overall.

Assessing and Correcting Misconceptions

The current study also illustrates methods for assessing misconceptions and measuring their strength. Past research has generally focused on two ways of diagnosing misconceptions in mathematics domains: 1) through interview techniques, asking individual students their thoughts about problems (e.g., Resnick et al., 1989) and 2) by classifying particular errors on an assessment as being associated with particular misconceptions (e.g., Siegler, 2002). While interview techniques are useful for extensively investigating individual students' misconceptions, they are difficult to do on a larger scale in a classroom full of students. Also, results from interview data can make it difficult to quantify differences between students and look for interactions with condition. The alternative of categorizing students' errors as misconceptions on an assessment is easier to implement on a larger scale, but it does not distinguish between strong misconceptions and guesses. Consequently, providing additional measures to assess the strength of students' misconceptions, such as confidence scales, can create a more complete picture of students' strong, prevalent misconceptions prior to an intervention or lesson.

While confidence scales have been used in past research to see if students differed in overall confidence (e.g., Taylor & Kowalski, 2004), they have rarely been used to specifically target the strength of students' misconceptions. Such scales can provide insight on the impact of educational interventions on changing students' misconceptions. In the current study, students' low confidence errors, which were most likely from guessing, decreased after our intervention, but students' high confidence errors, which were most likely strong misconceptions, did not decrease. This suggests that the intervention was helpful for students who knew they did not understand much about decimals, but the intervention did not correct strongly held misconceptions. In addition, the "hidden decimal" task indicated that students were often displaying a whole number misconception when providing a general rule for decimal magnitude comparisons. The rate of making a whole number misconception error during the "hidden decimal" task was somewhat higher than the rate at which students made whole number misconception errors on the rest of the assessment. This suggests that the "hidden decimal" task is providing different information about the misconceptions students hold. Such insights have important implications for constraints on when different instructional practices may be effective.

Essentially, misconceptions are difficult to change, as illustrated by past research on conceptual change (e.g., Vlassis, 2004). Repeatedly, conceptual change has been shown to be a gradual process, and students' strongly held misconceptions require extended instruction and time to be overcome (Vamvakoussi & Vosniadou, 2004; Vlassis, 2004). This may have been particularly true in the current study because the intervention was only during two classroom lessons, which is a short period of time. In addition, even when students do recognize their misconceptions as incorrect, this does not necessarily result in an understanding of the correct concepts (Van Dooren, Bock, Hessels, Janssens, & Verschaffel, 2004). This lack of understanding can lead students to use different incorrect ideas (Van Dooren, Bock, Hessels, Janssens, & Verschaffel, 2004), and they often revert back to misconceptions. In the current study, students needed to infer correct ideas from examples, with little instruction provided on correct ideas. As a result, it was most likely difficult for students, especially those with frequent misconceptions, to extrapolate the important correct ideas from the examples. Additional instructional supports could help students pull such correct ideas from worked examples. For example, providing students with more feedback on their explanations during the intervention could have encouraged greater knowledge change (e.g., Aleven & Koedinger, 2002). In addition, providing opportunities for structured conceptual change discussions, led by an instructor, may have also improved knowledge change during the intervention (Eryilmaz, 2002).

Future Directions

Given the unexpected results, it is clear that further research needs to be conducted to investigate the role of prior knowledge when learning from incorrect examples and comparison, and the circumstances under which these instructional practices may be most beneficial for students. Past work did not provide clear evidence for the effect of prior knowledge, and the current study had mixed results as well. Consequently, several next steps should be taken to further our understanding of when incorrect examples and comparison may be most beneficial.

An immediate next step is to explore measures of students' activities during the intervention, such as the quality of students' written explanations and their relation to outcome measures. I predict these data will show that students in the *Incorrect-Compare* condition compared more than students in the other two conditions, but that the rate of comparison will be less than expected. I also predict that the prevalence of prior misconception errors will moderate the effect of condition for different explanation features. For example, students with infrequent prior misconception errors may be more likely to talk about correct concepts in the *Incorrect-Sequential* condition relative to other conditions. In addition, I will examine how explanation features might predict outcome measures, and I predict talking about correct concepts will significantly predict outcomes based on past work (Durkin & Rittle-Johnson, 2012).

There are several other steps that could be taken in future research as well. First, the prevalence and strength of misconceptions were important factors in the current study, and future work should incorporate similar measures of misconceptions to assess students' readiness to learn from certain instructional practices. Also, incorrect examples and comparison may have different effects on learning if students are provided with more instructional guidance. Providing students, especially those with strong misconceptions, with more direct instruction or scaffolding during the intervention may improve their ability to process and compare worked examples. By increasing students' knowledge about the domain before the intervention, it might be possible to make comparison less overwhelming and increase attention to important, structural similarities and differences. This attention to critical features of correct and incorrect examples, as opposed to less important surface features, may be particularly important to fully benefit from learning with incorrect examples. In addition, more carefully scaffolded explanation prompts could encourage students studying incorrect examples to spend more time correctly discussing misconceptions, which should improve their attention to critical features of incorrect and correct examples (e.g., Van den Broek & Kendeou, 2008). Consequently, increased supports for students could improve learning from incorrect examples and from comparison.

Conclusion

In conclusion, the prevalence of students' prior misconception errors impacted how much they learned from correct and incorrect examples and from comparison. Students with infrequent prior misconception errors performed best after viewing correct and incorrect examples sequentially. Students with frequent prior misconception errors performed similarly across conditions. Consequently, students' prior misconceptions and knowledge must be considered when using these instructional strategies. Future research should examine potential scaffolds that could improve learning with incorrect examples and comparison.

Appendix A

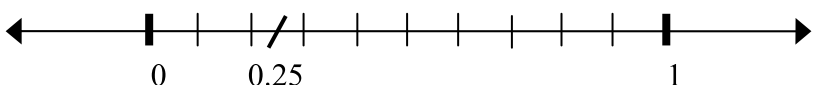
Introductory Lessons from the Intervention

Pretest Day Lesson: Modeling of Partner Work (3 minutes)

Clip digital recorder (labeled with your name) to your shirt/pants and turn it on to tape your lesson.

This week you are going to be working with a partner and studying examples, so I want to show you how you and your partner should work together. When you see a blank in the example, you need to fill it in, and when you see words underlined, you need to circle one of the words. Let's look at one example.

Put the following up on overhead #1, with the answer to the question covered:



Alex said, "I need to put 0.25 on the number line. First, I divided the line into __ pieces. Then, I counted over 2 pieces. And I put my mark for 0.25 a little before/after 2 tenths."

1. Why did Alex divide the line into 10 pieces?

Here you see how a student named Alex placed a number on the number line. Let's imagine that my partner _______ and I are going to try to understand how Alex solved this problem.

First, ___________ and I should look at Alex's solution, and try to see what Alex did, and finish labeling Alex's steps.

(During this section, gesture to the appropriate parts of the example.)

I: Let's see, ___________. Here Alex has to place 25 hundredths on the number line that goes from 0 to 1. What did Alex do? It says Alex thought of the line divided into blank pieces. But how many pieces did Alex break the line into?

Partner: I think Alex divided the line into 10 pieces. So we should write "10" in the blank. ***[Write 10 in the blank.]*** *Then what did Alex do? It says Alex counted over 2 pieces and then put 25 hundredths a little before or after 2. Do you think Alex put it before or after 2?*

I: It looks like Alex put 0.25 a little after 2. So we should circle the word "after". ***[Circle "after".]***

Next, ___________ and I need to answer the questions that are below Alex's solution.

I: It says, Why did Alex divide the line into 10 pieces? Hmmm. What do you think, ___________?

Partner: I think Alex did that because the number line goes from 0 to 1.

I: Oh, I see where you're heading. So Alex divided the line to figure out where 2 tenths would be.

Partner: Yes, and then Alex just marked 25 hundredths a little over to the right.

Then we would write our answer below the question. [*Reveal answer below question.*] So this is how you and your partner will be working together. You'll have a packet of problems to work on, and you'll need to work together to figure out how problems are solved and to answer questions like ___________ and I did. Sometimes you'll also be asked to solve some problems on your own.

Tomorrow we'll be randomly pairing you with a partner, and you will work together just like this.

Does anyone have any questions? OK, we'll see you tomorrow. [C*an turn off digital recorder.]*

Day 1 Lesson: Reminder about place value (5 minutes)

Clip digital recorder (labeled with your name) to your shirt/pants and turn it on to tape your lesson.

We're going to be learning about decimals for the next two days, and you are going to be working with a partner a lot. So let's start by reviewing some things about decimals.

Show overhead #2 with place value chart.

5, 749.123

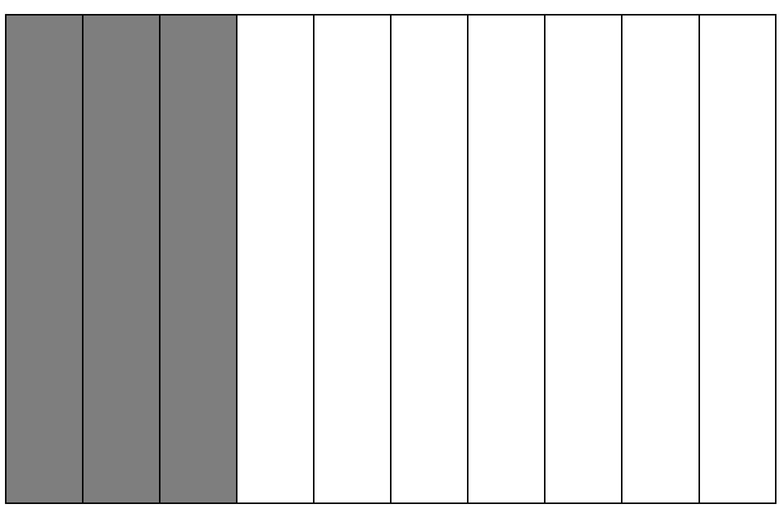
1,000s	100s	10s	1s	.	0.1s	0.01s	0.001s
Thousands	Hundreds	Tens	Ones	.	Tenths	Hundredths	Thousandths
5	7	4	9	.	1	2	3

Decimals are numbers that represent a part of a whole. Here is the number five thousand seven hundred forty-nine and one hundred twenty three thousandths broken into its place values:

You have 5 thousands, 7 hundreds, 4 tens, 9 ones and 1 tenth, 2 hundredths and 3 thousandths. [*Point to place value table as you label it.*]
Each place value is 10 times greater than the place value to its right. So 10 is ten times as much as 1, 1 is ten times as much as 0.1, and 0.1 is ten times as much as 0.01. [*Point to the relevant columns as you speak.*]

Let's do an example in the tenths place.
Put up overhead #3 with box divided into 10 pieces.

0.3

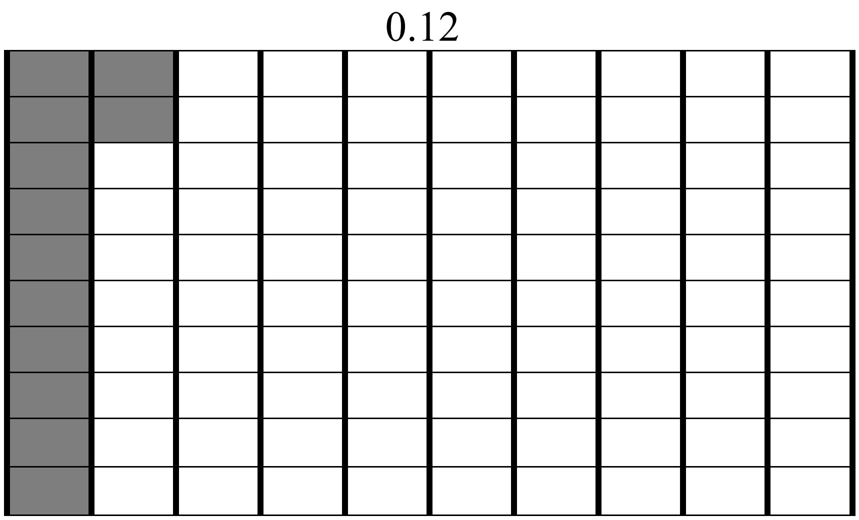


How would you say this number? [*Point to the 0.3 written on the overhead.*]

[If a student says "3 tenths" tell them that's correct. If a student says "zero point three", tell them that's correct but there is a better way to say it and see if they can say the other way. If no student says "3 tenths", then just tell the class it is "3 tenths".]

When a number is in the tenths, it means that it's out of a whole divided into 10 pieces. So 3 tenths is 3 out of 10 pieces. [*Point to the tenths rectangle.*]
This rectangle is a whole divided into 10 pieces. Each piece is 1 tenth [*point to one section*], so 3 of those pieces are colored in to show 3 tenths.

Now, let's do an example in the hundredths place.
Put up overhead #4 with box divided into 100 pieces.



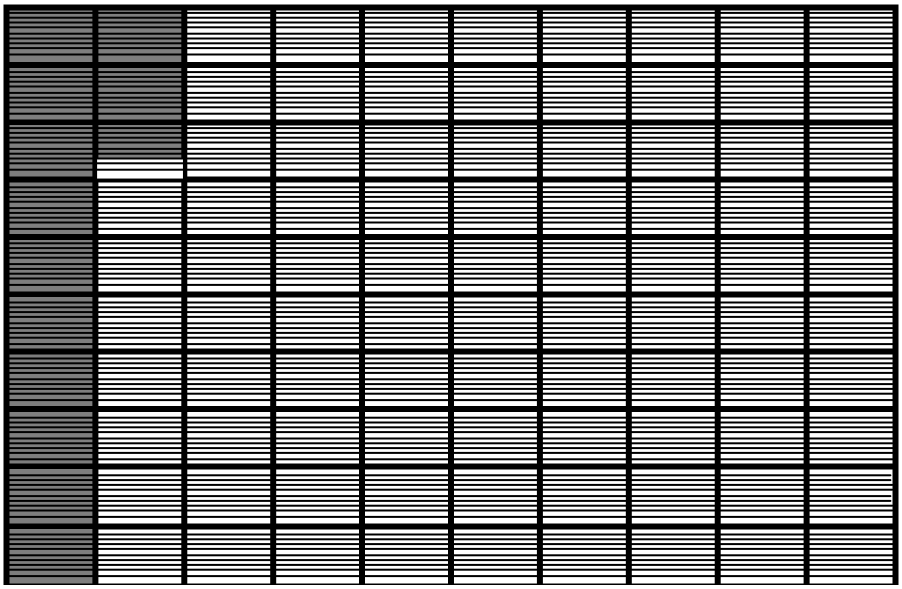
How would you say this number? [*Point to the 0.12 written on the overhead.*]

[If a student says "12 hundredths" tell them that's correct. If a student says "zero point twelve or zero point one two", tell them that's correct but there is a better way to say it and see if they can say the other way. If no student says "12 hundredths", then just tell the class it is "12 hundredths".]

For hundredths, a whole is divided into 100 pieces.
You can think of 12 hundredths as 12 out of 100 pieces. [*Point to the hundredths rectangle.*]
This rectangle is a whole divided into 100 pieces. I took the last rectangle that was divided into tenths and divided each tenth into 10 pieces [*point to first tenth*], to get 100 pieces. Each piece is one hundredth [*point to one hundredth box*], so 12 of those pieces are colored in to show 12 hundredths.

Now, let's do an example in the thousandths place.

0.127



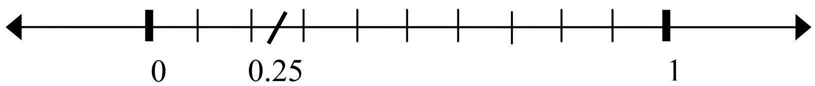
Put up overhead #5 with box divided into 1000 pieces.

How would you say this number? [*Point to the 0.127 written on the overhead.*]

[If a student says "127 thousandths" tell them that's correct. If a student says "zero point one hundred twenty seven or zero point one two seven", tell them there's a better way to say it and see if they can say the other way. If no student says "127 thousandths", then just tell the class it is "127 thousandths".]

For thousandths, a whole is divided into 1000 pieces.
You can think of 127 thousandths as 127 out of 1000 pieces. [*Point to the thousandths rectangle.*]
This rectangle is a whole divided into 1000 pieces. I took the last rectangle that was divided into hundredths and divided each hundredth into 10 pieces [*point to first hundredth*], to get 1000 pieces. Each piece is one thousandth [*point to one thousandth*], so 127 of those pieces are colored in to show 127 thousandths.

Today, you'll see decimals in the tenths, hundredths, and thousandths. Remember, you'll be working with a partner just like we showed you yesterday, and you'll be looking at examples like these. [*Put up overhead #1 from the Pretest Day. It looks like this:*]



Alex said, "I need to put 0.25 on the number line. First, I divided the line into __ pieces. Then, I counted over 2 pieces. And I put my mark for 0.25 a little before/after 2 tenths."

1. Why did Alex divide the line into 10 pieces?

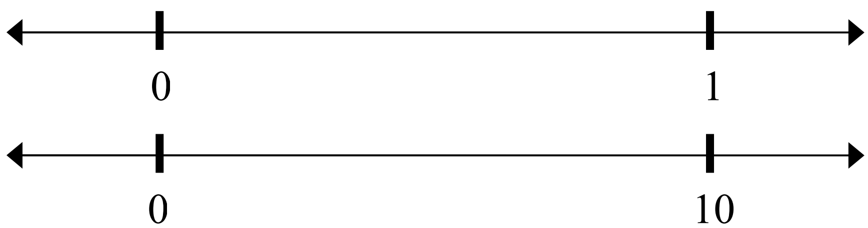
You'll be working with your partner to fill in blanks, circle the correct word when things are underlined, and answer questions. Please work until you get to the stop sign in your packet, and then raise your hand so we can check your answers.

Does anyone have any questions? OK – then we are ready to begin. [C*an turn off digital recorder.]*

Day 2 Number lines to 10 (3 minutes)

Turn on digital recorder to tape the lesson.
Nice work yesterday. Today, we'll start with a reminder about number lines.

Put up overhead #6 with 2 number lines (one from 0 to 1 and one from 0 to 10).



What's different about these two number lines?

(Let the students come up with differences, and students should come up with the difference that one ends at 1 and the other ends at 10. If no student mentions this, you can say that difference.)

Yesterday, you saw number lines that went from 0 to 1. So all the decimals you placed on the number line were greater than 0 and less than 1. Today you're going to see some decimals that are greater than 1. For these decimals, the number line will go from 0 to 10. Pay attention to each number line to see if it goes to 1 or to 10.

Any questions?

Can turn off digital recorder.

<u>Day 2 Wrap-up lesson (3 minutes) (do at end of 2nd intervention day)</u>

Turn on digital recorder to tape the lesson.

Everyone has done a great job with the work we've been doing for the past couple days. Tomorrow, we are going to give you a test to see if the work you've been doing has made sense to you. Today, I wanted to go over some of the things that you may have noticed as you've been looking at problems.

Put up summary overhead #7.

- *You can't think of decimals the same way you think of whole numbers.*
  Decimals and whole numbers are different, and you can't treat them the same way.

- *You need to pay attention to place values.*
  With decimals, it is very important to pay attention to which number is in each place value. You saw numbers that had values in the tenths, hundredths, and thousandths. Looking at the number in each place will help you figure out how big the decimal is.

- In fact, there are at least three different ways you can think about how big a decimal is.
  1) *Look at how many tenths it has.*
     You can see how many tenths a decimal has to figure out where it goes on a number line.

  2) *Look at what number it is out of.*
     You can think of the decimal out of 10, 100, or 1000 pieces to estimate where the decimal goes on a number line.

  3) *See if it is near other numbers you know, like a half.*
     If you know a decimal is near another number you know, like one half, you can place the decimal on a number line near that number.

Thank you again for working so hard today, and we'll see you tomorrow.

Appendix B

Assessment

For each pair, circle the decimal that is greater:

1)	0.24	0.049
2)	0.3	0.92
3)	0.561	0.17
4)	0.87	0.835
5)	0.429	0.7

6) How sure are you that you solved the 5 problems above correctly?
Circle a number from 1 to 5.

1	2	3	4	5
Not at all sure	A little sure	Somewhat sure	Sure	Very sure

Write a number that comes between:

7) 0.3 and 1.0 ______

8) 0.5 and 0.52 ______

9) 0.5 and 0.6 ______

10) 0.76 and 0.77 ______

11) 0.14 and 0.148 ______

12) Circle all the numbers that are worth the same amount as 0.51

a) 0.5100
b) 0.051
c) 0.510
d) 51
e) none of the above

13) Circle all the numbers that are worth the same amount as 0.04

- a) 0.4
- b) 0.40
- c) 0.004
- d) 4
- e) none of the above

14) Circle the number that is **greater** than 0.36

- a) 0.4
- b) 0.360
- c) 0.2
- d) 0.279

15) Circle the number that is **less** than 0.52

- a) 0.6
- b) 0.5
- c) 0.567
- d) 1.4

16) Circle the number nearest to 0.675

- a) 0.98
- b) 0.5
- c) 0.7
- d) 700

17) Circle the number nearest to 0.18

- a) 0.02
- b) 0.1
- c) 0.2
- d) 20

Circle the answer that goes in the blank.

18) 0.4 is _______________ 0.004

a) greater than
b) less than
c) the same as

19) 0.26 is ______________ 0.260

a) greater than
b) less than
c) the same as

20) 0.8 is _______________ 0.08

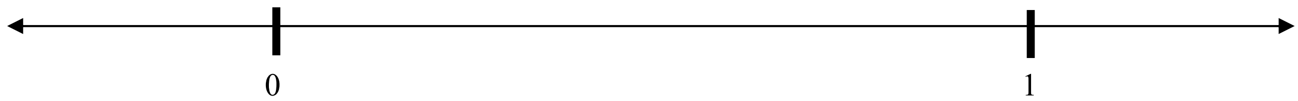
a) greater than
b) less than
c) the same as

21) Clare weighs four packages. They weigh 0.714, 0.32, 0.6 and 0.79 grams. Put the four decimals in order from least to greatest.

Least ___________ ___________ ___________ ___________ Greatest

Mark about where each decimal goes on the number line with a slash (/).

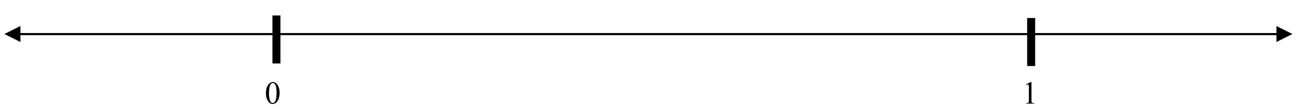
22) 0.3



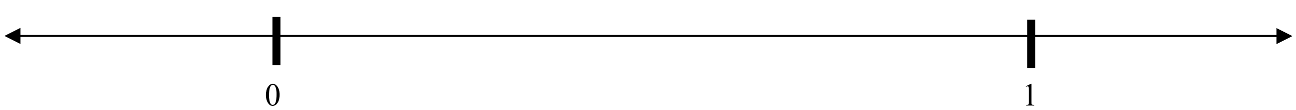
23) 0.256

23a) How did you decide where to put your slash?

24) 0.07



25) 0.83



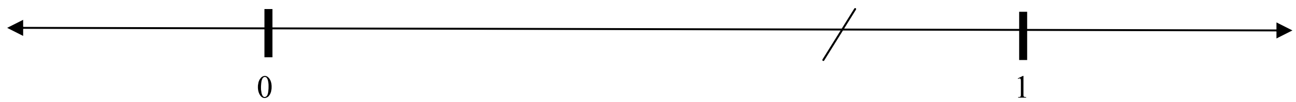
25a) How sure are you that you solved this problem correctly? Circle a number from 1 to 5.

1	2	3	4	5
Not at all sure	A little sure	Somewhat sure	Sure	Very sure

What number tells about where the slash is on the number line? Circle the answer.

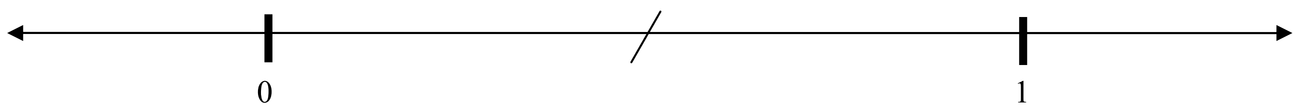
26)

a) 0.76
b) 0.3
c) 0.08
d) 0.401



27)

a) 0.214
b) 0.84
c) 0.489
d) 0.05

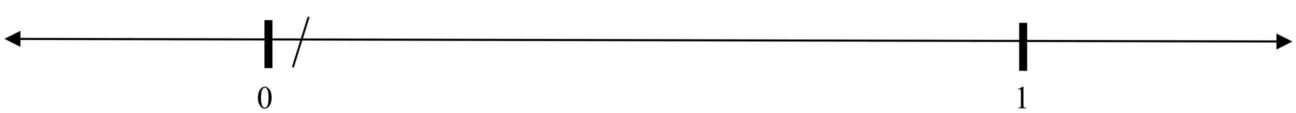


27a) How sure are you that you solved this problem correctly? Circle a number from 1 to 5.

1	2	3	4	5
Not at all sure	A little sure	Somewhat sure	Sure	Very sure

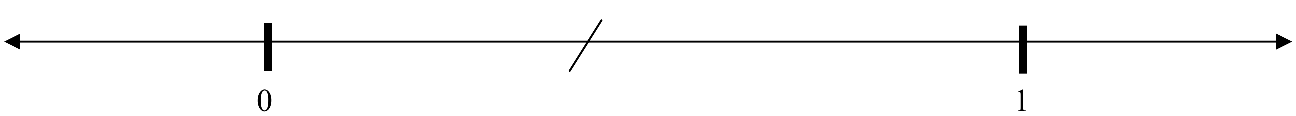
28)

a) 0.534
b) 0.5
c) 0.032
d) 0.80



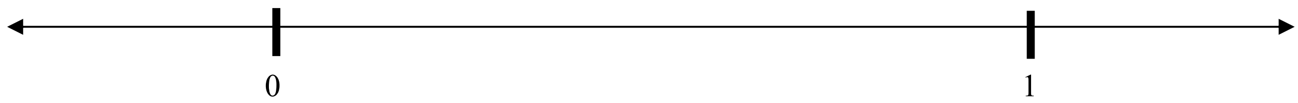
29)

a) 0.189
b) 0.4
c) 0.05
d) 0.87

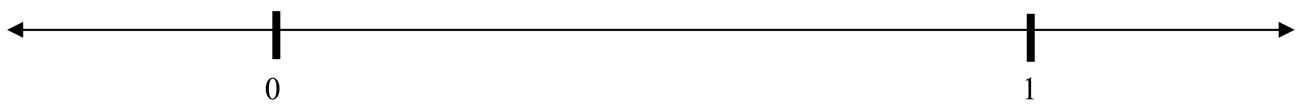


Mark about where each decimal goes on the number line with a slash (/).

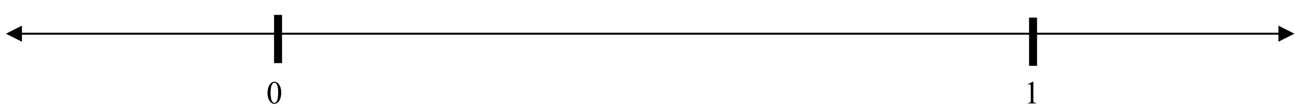
30) 0.3826



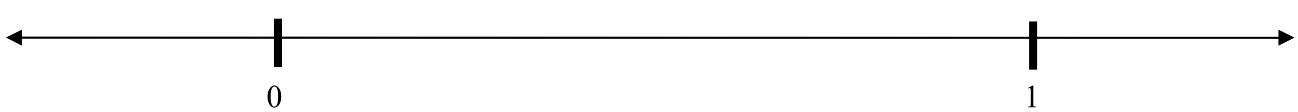
31) 0.0851



32) 0.1473



33) 0.7649



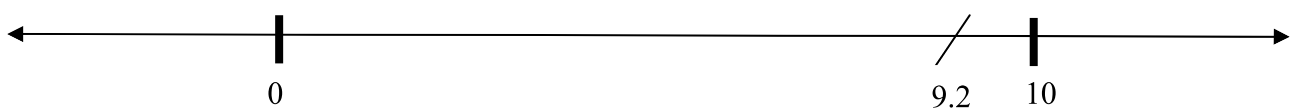
Now the number line goes from 0 to 10
One number is already marked on the number line. Mark about where the other number goes on the number line with a slash (/).

34) 3.52 is marked. Mark where 3.8 goes.

34a) How sure are you that you solved this problem correctly? Circle a number from 1 to 5.

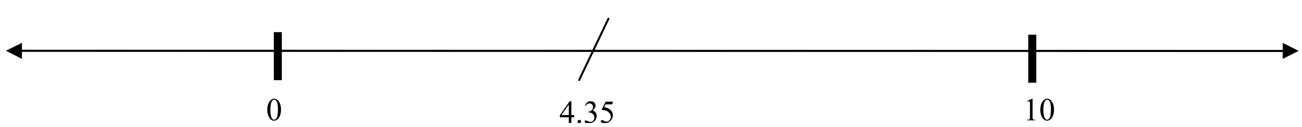
1	2	3	4	5
Not at all sure	A little sure	Somewhat sure	Sure	Very sure

35) 9.2 is marked. Mark where 9.05 goes.



35a) How did you decide where to put your slash?

36) 4.35 is marked. Mark where 4.842 goes.

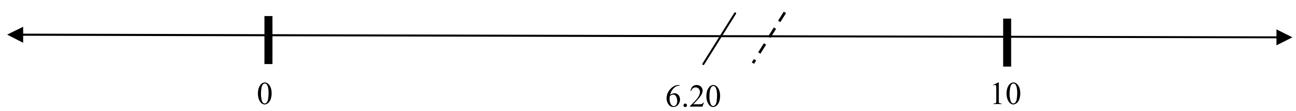


Now the number line goes from 0 to 10

One number is already marked on the number line. What number tells about where the unmarked dashed slash is on the number line? Circle the answer.

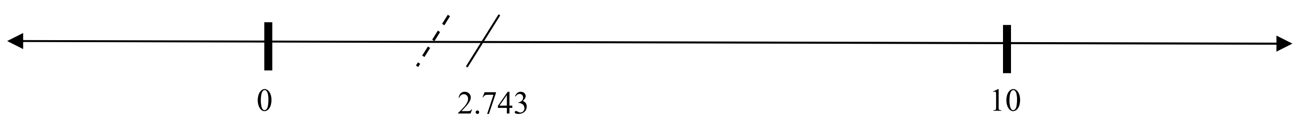
37)

a) 6.173
b) 6.8
c) 6.05
d) 0.45



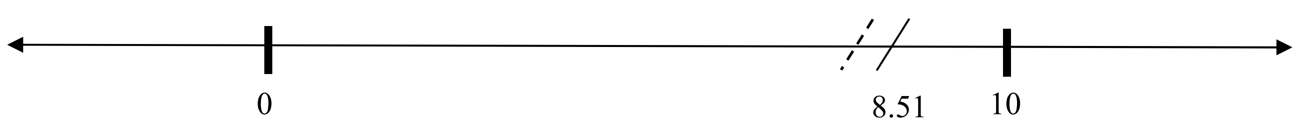
38)

a) 2.814
b) 0.2
c) 2.9
d) 2.09



39)

a) 8.147
b) 8.6
c) 0.8
d) 8.510



39a) How sure are you that you solved this problem correctly? Circle a number from 1 to 5.

1	2	3	4	5
Not at all sure	A little sure	Somewhat sure	Sure	Very sure

40) If you have two decimals, how can you figure out which decimal is bigger?

Please wait for directions before answering the last question.

41) Which of these numbers is greater?

a) 0.□ b) 0.□□□□ c) Can't tell

Explain your reasoning:

42) Please circle one: I am a ….

a) Boy b) Girl

REFERENCES

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*, 147-179.

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

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*, 774-783.

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. B. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education, 29*, 3-20.

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

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

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.

Desmet, L., Gregoire, J., & Mussolin, C. (2010). Developmental changes in the comparison of decimal fractions. *Learning and Instruction, 20*, 521-532.

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.

Durkin, K., & Rittle-Johnson, B. (in preparation). Comparing Incorrect and Correct Examples: The Effects of Practice and Instructional Explanations.

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction, 22*, 206-214.

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*, 1001-1015.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., Karns, K., & Dutka, S. (1997). Enhancing Students' Helping Behavior during Peer-Mediated Instruction with Conceptual Mathematical Explanations. *The Elementary School Journal, 97*, 223-249.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155-170.

Gentner, D. (1989). The mechanisms of analogical learning. In A. Ortony & S. Vosniadou (Eds.), *Similarity and analogical reasoning* (pp. 199-241). New York, NY: Cambridge University Press.

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

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

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

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

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.

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

Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone: Cooperative, competitive and individualistic learning* (4th ed.). Boston, MA: Allyn and Bacon.

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

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.

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

Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction, 19*(158-170).

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin and Review, 6*(4), 586-597.

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).

Murray, T., Schultz, K., Brown, D., & Clement, J. (1990). An analogy-based computer tutor for remediating physics misconceptions. *Interactive Learning Environments, 1*(2), 79-101.

National Center for Education Statistics. (2011). *NAEP Questions Tool*: U.S. Department of Education, Institute of Education Sciences.

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

McNeil, N. M., & Alibali, M. W. (2005). Why Won't You Change Your Mind? Knowledge of Operational Patterns Hinders Learning and Performance on Equations. *Child Development, 76*(4), 883-899.

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

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*, 437-448.

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.

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., 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., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 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*, 529-544.

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101*(4), 836-852.

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*, 147-177.

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.

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.

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14*, 503-518.

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102*, 408 - 426.

Taylor, A. K., & Kowalski, P. (2004). Naive psychological science: The prevalence, strength, and sources of misconceptions. *The Psychological Record, 54*, 15-25.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*(5), 453-467.

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.

Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Verschaffel, L. (2004). Remedying secondary school students' illusion of linearity: A teaching experiment aiming at conceptual change. *14*, 485-501.

VanLehn, K. (1999). Rule-learning events in the acquisition of a complex skill: An evaluation of cascade. *The Journal of the Learning Sciences, 8*, 71-125.

Vlassis, J. (2004). Making sense of the minus sign or becoming flexible in 'negativity'. *Learning and Instruction, 14*, 469-484.