THE EFFECTS OF ERROR REFLECTION AND PERCEIVED FUNCTIONALITY OF ERRORS ON MIDDLE SCHOOL STUDENTS' ALGEBRA LEARNING AND SENSE OF BELONGING TO MATHEMATICS

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
Of the Requirements for the Degree
DOCTOR OF PHILOSOPHY
in Educational Psychology

---

by
Christina Barbieri Doherty
Diploma Date: July 2015

Examining Committee:
Dr. Julie L. Booth, Advisory Chair, Department of Psychological, Organizational, & Leadership Studies
Dr. James P. Byrnes, Department of Psychological, Organizational, & Leadership Studies
Dr. Jennifer G. Cromley, Department of Educational Psychology, University of Illinois at Urbana-Champaign
Dr. Kristie J. Newton, Department of Teaching & Learning
Dr. Mark F. Schmitz, External member, School of Social Work

ABSTRACT

The current study assessed an error reflection intervention on Algebra I students' conceptual and procedural knowledge and sense of belonging to mathematics. Also of interest was whether perceptions of the functionality of errors mediated the effect of condition on learning and sense of belonging to mathematics. Middle school students (*N* = 207) were randomly assigned within classroom to one of four conditions: 1) a Problem-Solving Control group, 2) a Correct Examples Control group, 3) a Correct Examples Error Reflection condition that promoted reflection on hypothetical errors through self-explanation prompts, or 4) an Incorrect Examples Error Reflection condition that promoted reflection on displayed errors within the example through self-explanation prompts. Conceptual and procedural knowledge, sense of belonging to mathematics and perceived functionality of errors were measured pre- and post-intervention.

After controlling for unanticipated clustering effects, results suggest that reflecting on and explaining errors within a worked examples intervention is just as effective at promoting learning as traditional problem solving alone or working with traditional correct worked examples and written self-explanation prompts. Students' sense of belonging to mathematics or perceived functionality of errors for learning were high at the start of the study and remained so throughout the intervention. Perceptions of the functionality of errors were unrelated to learning and sense of belonging to mathematics. The limited size of the minority population in the sample did not allow for exploration of differential effects of condition for underrepresented minority (URM) students. However, these students reported lower feelings of belonging to mathematics than non-URM students. Implications for theory and practice are discussed.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Julie Booth, who provided not only invaluable professional advice but also personal guidance, who helped me grow as a researcher, and who believed in me when I didn't.

Thanks are also due to the members of my committee, Dr. Jennifer Cromley, Dr. James Byrnes, Dr. Mark Schmitz, and Dr. Kristie Newton, for willingly sharing their guidance and expertise.

I'd also like to thank my level-headed labmates and friends, Dr. Kelly McGinn and Laura Young, who were always ready to offer a word of encouragement, a pragmatic perspective, or a sensible solution no matter how busy they were.

Finally, I'd like to thank a dear friend and colleague, Dana Miller-Cotto, for helping me realize my passion for research and being my most supportive cheerleader since we met six years ago.

Each of those listed here played an integral role in the completion of this dissertation as well as my professional development and for that I extend my deepest gratitude.

DEDICATION

To my mother,

who instilled in me the value of hard work, who always put me first, and who supported me in any crazy new endeavor I pursued, as long as it made me happy. Bendición ma.

To my husband,

whose unceasing faith in me gave me the strength to persist, whose patience and understanding was probably unwarranted, and whose dedication in the kitchen not only kept me well fed but allowed me even more hours at the computer. We did this together.

TABLES OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The Research Problem

Algebra I, commonly taken in the ninth or eighth grade, is often considered a gatekeeper course for higher level math courses as well as select science courses (Matthews & Farmer, 2008). Students who take higher level math courses are more likely to be admitted into a four-year college (Schneider, Swanson, & Riegle-Crumb, 1998) and to major in a STEM field (Chen, 2009). Unfortunately, there is a well-documented decline in mathematics achievement as well as achievement motivation for mathematics during adolescence, particularly between seventh and eighth grade (e.g., Wang & Pomerantz, 2009). Achievement and motivation in mathematics are closely linked especially during adolescence (Wang, 2013). Adolescents across the country struggle with mathematics. According to the 2012 Programme for International Student Assessment (PISA), which measured 15-year-olds' performance on several academic subjects across 65 countries, the United States ranked below average in mathematics in comparison to other nations who participated in the assessment (Kelly, Xie, Nord, Jenkins, Chan, & Kastberg, 2013). Further, this report suggests that 50% of US students have low levels of interest in mathematics.

The decline within US students’ math achievement is even more prevalent for students of racial and ethnic minority status and students from impoverished neighborhoods (Hemphill & Vanneman, 2010; Vanneman, Hamilton, Baldwin Anderson, & Rahman, 2009). In the United States, Black and Hispanic students generally do not perform as well as their White and Asian counterparts on achievement tests, particularly in mathematics and science (National Center for Education Statistics, NCES, 2013). Prior achievement in mathematics may eventually lead to decisions to pursue a STEM degree. In 2012, only 6.3% and 6.5% of all doctorates were awarded to Black students and Hispanic students respectively (NSF, 2012). Most of these doctorates awarded to Black students were in a non-STEM-related field such as education; for Hispanic students, most doctorates awarded within STEM were restricted to the social sciences, with very few doctorates being awarded in life sciences, physical sciences, or engineering. Within these hard science fields, doctorates were predominantly awarded to Asian students followed by White non-Hispanic students (NSF, 2012).

Despite increased acknowledgement of these issues, the US continues to perform below average in international mathematics assessments and ethnic and income achievement gaps persist. Some scholars suggest that the achievement gap between White and Black students has been slowly narrowing and that the gap between low-income and high-income students has been increasing dramatically (Reardon, 2013). Although eighth grade math achievement has increased slightly across all racial and ethnic groups, the gap in math scores between Black and White students and Hispanic and White students has not narrowed since 2008 (NCES, 2013). As economic disparities mainly coincide with racial differences in the US, treating the income achievement gap as an isolated research issue unrelated to racial inequalities may result in fragmented knowledge and even more questions than answers. How to remedy the situation is undoubtedly a complex social issue, but the preponderance of findings suggests prior achievement and motivation as strong determinants of later achievement and intent to pursue a STEM degree for both minority students and nonminority students (Mau, 2003; Wang, 2013). To remedy the problem, changes to our education system are needed that facilitate student learning. Achievement that is lower than expected for students of varying ethnic and socioeconomic backgrounds is likely related to a lack of resources (Darling-Hammond, 2006) which may include opportunities to learn and to be motivated to do so.

Increasing student learning while simultaneously improving students' motivation may be critical for promoting lasting change in mathematics achievement. Middleton and Spanias (1999) suggest that achievement motivation research should consider research on learning and classroom discourse and should also consider the domain of mathematics from a social perspective. Though there are many motivation constructs prevalent in the literature, students' sense of belonging to a math community is particularly important for learning in math (Good, Rattan, & Dweck, 2012). Unfortunately, students' sense of school belonging (Anderman, 2003) has been found to decline during adolescence. One potentially useful method for altering students' sense of belonging as well as their view on learning while also promoting actual learning in algebra may be the use of an error reflection intervention in the classroom.

Teachers differ greatly in how they handle their students' errors (Schleppenbach, Flevares, Sims, & Perry, 2007). In many classrooms, especially in the United States, errors are explicitly discouraged. Some teachers discourage errors due to concern that students' consideration of these errors may reinforce their incorrect procedures or faulty knowledge (Santagata, 2004). This is particularly true of teachers in the United States (Stigler & Perry, 1988). In contrast, errors are frequently discussed in Japanese classrooms, where they are thought to be an integral part of learning (Stigler & Hiebert, 1999). In China, teachers seem to utilize errors to provoke students' discussion of mathematical concepts and to relieve any shame that students may have about making mistakes (Stevenson & Stigler, 1992; Wang & Murphy, 2004). Though a number of scholars suggest that focusing instructional time on student errors can be a prime opportunity to correct student misconceptions (Borasi, 1994; Henderson & Harper, 2009; Yerushalmi & Polingher, 2006), little empirical evidence exists that supports the causal nature of this claim. Middle school students' perceptions ofhow errors are evaluated in the classroom, or the *perceived error climate*, predicts their adaptive reactions to struggles in mathematics problem-solving (Steuer, Rosentritt-Brunn, & Dresel, 2013). Further, how consideration of errors influences students' beliefs about their ability and their classroom learning experience has rarely been considered. The purpose of the current study is to determine the influence of an error reflection intervention on student sense of belonging to mathematics and algebra learning.

Purpose

The current study had several purposes. The main purpose of the current study was to determine the influence of an error reflection intervention on student sense of belonging to mathematics and algebra learning. Specifically, the current study assessed whether reflection on errors within a worked examples intervention improves middle grade students' algebra learning. It was hypothesized that error reflection would lead to increases in conceptual knowledge by refining knowledge of concepts targeted in the examples and self-explanation prompts. Promoting error reflection was also expected to lead to increases in procedural knowledge by increasing students' negative knowledge (i.e., what *not* to do) or weakening knowledge of incorrect procedures. The current study also investigated whether there are differential effects of error reflection in the form of incorrect worked examples combined with self-explanation prompts in comparison to error reflection in the form of correct worked examples combined self-explanation prompts focused on errors.

Second, the study examined whether reflection on errors within a worked examples intervention improves middle grade students' sense of belonging to mathematics. It was expected that the error reflection intervention would lead to increases in sense of belonging to mathematics through normalization of errors as part of the learning processes.

Third, the current study investigated whether increases in students' perceived functionality of errors mediates the influence of the error reflection intervention on learning and sense of belonging to mathematics. Promoting error reflection was expected to lead to increases in students' perceived functionality of errors for learning. This was expected to be a causal mechanism for increases in learning and sense of belonging to mathematics. If students see errors as an important part of the learning process, they may be more likely to benefit from the opportunities to learn from errors that are present during error reflection. These general hypothesized relationships are presented in Figure 1.1.

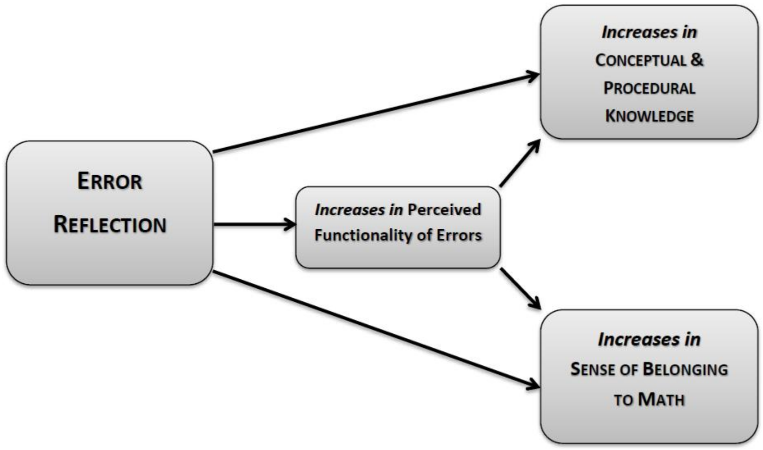


Figure 1.1. Basic Theory of Change

The current study also considered individual differences in these relationships. In a previous study (Barbieri & Booth, under review), a worked examples intervention which included promotion of error reflection in the form of incorrect examples led to increases in sense of belonging to mathematics specifically for URM students. However, due to a lack of a proper control condition and a small sample size, the exact cause of this effect is unable to be specified. The current study aimed to address whether the relationship between the error reflection intervention and student learning and sense of belonging to math differs by URM status as suggested in the previously mentioned study. The predicted model also considered whether differences found could be explained by differences in prior knowledge or prior sense of belonging to mathematics. This more nuanced approach is reflected in Figure 1.2 below.

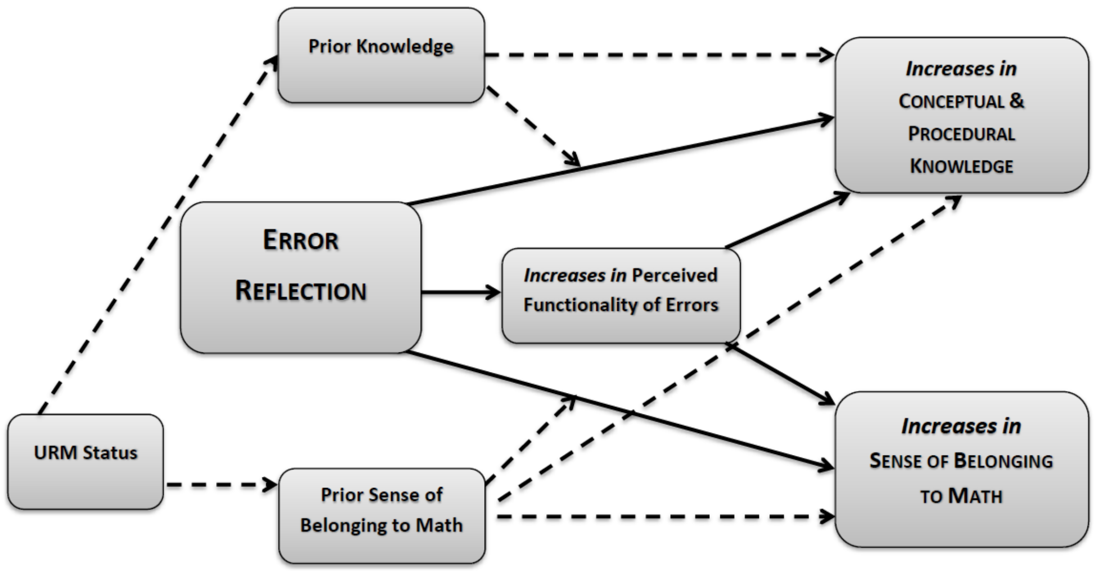


Figure 1.2. Theory of Change with consideration of individual differences

The dotted lines in Figure 1.2 demonstrate hypothesized effects of prior knowledge and sense of belonging to mathematics. It was expected that students' prior knowledge would moderate the influence of the error reflection intervention on learning. Specific hypotheses by condition will be further explained after explanation of methodology. However, a main effect of prior knowledge on learning was also expected regardless of condition. More prior knowledge may allow students to be more ready to learn and result in greater learning in general than those with low prior knowledge. It was further expected that a prior sense of belonging to mathematics may moderate the influence of the error reflection intervention on learning. Students with a lower sense of belonging to mathematics may find error reflection more salient and may be impacted to a greater extent in terms of increases in sense of belonging to mathematics. However, prior sense of belonging to mathematics may also demonstrate a main effect on change in sense of belonging and learning regardless of condition. Students who have a high sense of belonging to mathematics may be easier to influence in terms of motivation and may demonstrate higher sense of belonging post-intervention. They also may be more ready to learn and may show greater learning across conditions.

Research Questions

The specific research questions addressed in the current study are listed below. The literature underlying the conceptual framework is described in detail in Chapter 2.

1. Does promoting error reflection within a worked examples intervention improve middle grade students' algebra learning? Are there differential effects of promoting error reflection with the error displayed in comparison to the error simply mentioned?
    a. If so, does prior knowledge moderate the effect of condition?
2. Does promoting error reflection within a worked examples intervention improve middle grade students' sense of belonging to mathematics?
3. Is the influence of error reflection on the outcome variables mediated by changes in students' perception of the functionality of errors?
    a. Is the influence of error reflection on students' post-test conceptual scores mediated by changes in students' perception of the functionality of errors?
    b. Is the influence of error reflection on students' post-test procedural scores mediated by changes in students' perception of the functionality of errors?
    c. Is the influence of error reflection on students' post-intervention sense of belonging mediated by changes in students' perception of the functionality of errors?

4. Do the relationships between the error reflection intervention and students' learning and sense of belonging to math differ by URM status? If so, can prior knowledge and prior sense of belonging explain these differences?

Significance of the Study

As previously suggested, the current study sought to make several important contributions. US students have a history of struggles in Algebra I, a "gatekeeper course" for higher level math courses and indirectly for college admittance and pursuing STEM majors (Chen, 2009; Matthews & Farmer, 2008; Schneider, Swanson, & Riegle-Crumb, 1998). Declines in adolescent math achievement and motivation coincide (Wang & Pomerantz, 2009; Wang, 2013). Declines in achievement are more severe for underrepresented minority students and students living in poverty (Hemphill & Vanneman, 2010; Vanneman et al., 2009). There is also a well-documented achievement gap between Black and White students, and Hispanic and White students (NCES, 2013). These students are often underrepresented in STEM fields (NSF, 2012). Despite increased acknowledgement of these issues, the US continues to perform below average in international mathematics assessments and ethnic and income achievement gaps persist. Achievement that is lower than expected for students of varying ethnic and socioeconomic backgrounds is likely related to a lack of resources (Darling-Hammond, 2006) which may include both opportunities to learn and to be motivated to do so. The currently current study will seek ways to increase learning opportunities for Algebra I students that may simultaneously foster improvements in students' motivation.

Both cognitive and motivational interventions have the potential to increase student learning in a variety of domains. The influence of motivation on achievement has been explored for many years as has the influence of cognition. However, they are not often studied together. Achievement and motivation in mathematics are closely linked, especially in adolescence (Wang, 2013), yet few interventions in the education literature seek to address both cognition and motivation. Many students may have the skill to achieve in a particular area but don't necessarily do so. This may be explained by a variety of factors including resources and opportunities to learn (Byrnes, 2003) as well as instructional contexts that encourage student achievement motivation (Stipek, 2002). Cognitive interventions often seek to design instruction to be more suitable for students' cognitive capabilities (Sweller, 2012). Deriving interventions that use what we know about cognitive science can and has provided us with many useful learning tools (i.e., worked examples within cognitive load theory, self-explanations and the generation effect, distributed practice and spacing, scaffolding and fading, etc.). Motivation interventions often seek to increase student engagement or alter student beliefs in order to increase the effectiveness of traditional instruction (e.g., Blackwell, Trzesniewsky, & Dweck, 2007). Deriving interventions that use what we know about student motivation can and has provided us with many useful tools to alter students' perceptions about themselves and learning and altered motivational beliefs (i.e., utility value writing prompts, instruction on malleability of intelligence, altering attributions of feelings of belonging, writing about future academic selves, etc.). However, some motivation interventions focus solely on increases in the particular motivational construct targeted but do not simultaneously address learning (e.g., Falco, Summers, & Bauman, 2011).

Even when considering all that is seemingly “known”, there are inconsistencies in the effects found. For example, the effectiveness of the highly cited worked examples effect does not always lead to learning and effects differ based on factors such as design and structure of the worked example as well as students’ prior knowledge, etc. (Atkinson, Derry, Renkl, & Wortham, 2000; Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, & Sweller, 2001). Students still struggle to succeed in algebra. Many factors influence students’ ability to achieve. Focusing on either cognition or motivation alone may not be enough to promote lasting changes in learning and motivation. Designing a method that can potentially lead to changes in both motivation and cognition may be more effective and provide important information that is missing from the literature. We currently know very little about whether there are or can be motivational benefits from cognitive interventions such as worked examples, written self-explanations, or interventions promoting error reflection. This missing piece in the literature may provide a more thorough explanation of the effects found on learning. Both cognitive and motivation interventions have provided useful information about student learning and useful tools for increasing motivation. However, considering both the motivation and learning difficulties that students experience in mathematics during adolescence, and the particular struggles that URM students may experience, the current study suggests a more comprehensive approach is needed.

Focusing directly on increasing both student motivation and student learning may be critical to address the complex issues discussed. The current intervention of error reflection is unique in that it was designed to be cognitively stimulating and motivating. The experimental design of the study allows for the isolation of the effect of promoting error reflection on both students' learning and motivation, while the in-vivo aspect provides higher levels of ecological validity than is often seen in research on learning and cognition. Rather than simply assessing whether students' motivation influences the effectiveness of the intervention, the current study allows for the assessment of changes in motivation due to the intervention as well. The current study is one of the first to consider methods for altering middle school students' sense of belonging to the mathematics domain (see Barbieri & Booth, under review, for an exception). Further the current study is the first worked examples intervention to consider students' perceptions of a characteristic of the classroom climate, perceived functionality of errors, on both learning and motivation. The social-cognitive approach taken in the current study allows for consideration of multiple facets of students' learning experience in the Algebra classroom. The few contributions listed here are expanded upon in Chapter 2 which provides a detailed report of the existing state of knowledge applicable to the current study. The following section clarifies the boundaries of the current study.

Delimitations

Math cognition and motivation are two broad areas of research each with countless questions that can and should be answered. As previously described, the purpose of the current study was to determine the influence of an error reflection intervention on middle school students' sense of belonging and algebra learning. The current study investigated whether increases in students' perceived functionality of errors mediates the influence of the proposed error reflection intervention on learning and sense of belonging to mathematics. The current study also aimed to consider individual differences in these relationships. In particular, the current study aimed to address whether the relationship between the error reflection intervention and student learning and sense of belonging to math differs by URM status as suggested in a prior study (Barbieri & Booth, under review). If differences were found, consideration was to be given to whether these differences can be explained by differences in prior knowledge or prior sense of belonging to mathematics. However, a smaller than expected proportion of URM students in the sample did not allow this question to be answered. This is discussed in further detail in Chapters 4 and 5. The following is a list of delimitations that specify the scope of the current study:

1. The current study uses worked examples which are typically considered a cognitive load effect, derived from the cognitive load literature. The current study does not test cognitive load theory. This, for one, would require a measure of load and many other modifications to the design. Rather, the current study uses worked examples as a tool to promote error reflection, as much prior research has already tested the effectiveness of worked examples (e.g., Atkinson, et al., 2000; Booth et al., 2013; Booth et al., in press) and explored the nuances of cognitive load theory (e.g., Kalyuga, 2011; van Gog, Paas, & Sweller, 2010).
2. Similarly, although the current study uses self-explanation prompts, the goal of the current study is not to test the effectiveness of written self-explanation prompts. This would also require considerable modifications to the design. Rather, the current study is using self-explanation prompts as a tool to promote error reflection. Prior research has already explored the effectiveness of aural self-explanation prompts (e.g., Chi, De Leeuw, Chiu, & LaVancher, 1994; Rittle-Johnson, 2006) and self-explanations with pull-down menus prompted by

computers (Aleven & Koedinger, 2002) and written self-explanation prompts and another manuscript is in preparation on a study assessing the effectiveness of written self-explanation prompts used on classroom worksheets (Miller-Cotto, Booth, & Barbieri, in prep).

3. The sample is restricted to middle school students taking Algebra I in a midwestern public school district. This sample was chosen based upon prior work with this district. While this limited sample may restrict generalizability of the results, the participating school district includes a high proportion of underrepresented minority students. The expectation of a high proportion of URM students should have allowed for a large enough sample of URM students to answer research question 4. Although this did not end up being the case, this explains this design feature of the study.
4. To allow for more manageable data analyses and interpretation, self-report measures use Likert scales as a response format and do not include open-ended response items or interviews. However, each of the scales utilized have been previously validated. These are detailed in Chapter 3.
5. Although there are many other important achievement motivation constructs that are related to mathematics achievement in middle school, the current study focuses primarily on sense of belonging to mathematics. This construct was chosen for several reasons which are further discussed in chapter 2. Briefly, of primary concern was exploring a motivation construct that may be particularly relevant to those underrepresented in the domain of mathematics. Walton and Cohen (2007; 2011) suggest that students who are more likely to experience

negative stereotypes may also be more concerned with feelings on belonging than students who are not negatively stereotyped. Prior research demonstrates that sense of belonging to mathematics is a contributor to female college students' math course grades and intent to pursue a STEM career (Good, Rattan, & Dweck, 2012). Sense of belonging to the domain of math at the middle school level has not been explored. As URM students are particularly underrepresented in math (NSF, 2012), and as Algebra I poses particular problems for students in comparison to other subjects, sense of belonging to math was deemed to be a relatively new, interesting, and potentially fruitful area to explore in terms of understanding other factors that may contribute to the pervasive achievement gap in the US between URM and non-URM students.

6. One factor that has been found to contribute to underachievement of stereotyped individuals is stereotype threat, the fear of confirming a negative stereotype about one's group that can lead to underperformance in settings in which it is made salient (Steele, 1997). It is important to note that the current study in no way manipulated the environment to induce stereotype threat nor is stereotype threat measured. Stereotype threat is often considered a feeling that is linked to a particular experience in which one is presented a task in a way that reminds one of the stereotypical underperformance of one's group on that task. Sense of belonging to math is a separate construct and is defined as a general feeling of belonging or acceptance to the domain. While it is likely that URM students experience a greater history of stereotype threat experiences in their academic career, and while it is even possible that repeated experiences of stereotype threat

may lead to lowered sense of belonging, measuring stereotype threat was not deemed vital to the research questions of the study. It is important to note that the experience of stereotype threat is often considered to be more characteristic of students in *environments* in which they are particularly underrepresented (i.e., high achieving Black college student in a predominantly White Ivy League university). The current study was conducted in a very diverse middle school in the midwest, which is not a typical setting for experiencing the negative effects of stereotype threat. This district also has a notable achievement gap in algebra. Much research has been conducted on stereotype threat and several interventions have been created to reduce feelings of stereotype threat (Aronson, Fried, & Good, 2002; Cohen, Garcia, Apfel, & Master, 2006; Cohen, Garcia, Purdie-Vaugns, Apfel, & Brzustoski, 2009). While it is possible that stereotype threat is related to sense of belonging to math (see Good, Rattan, & Dweck, 2012), the current study focused on designing and testing a new error reflection intervention that may modify one's sense of belonging to math, without considering stereotype threat. Measuring or manipulating stereotype threat is beyond the scope of this study but may be an interesting area to explore in future research.

7. The current study measured students' entity view of math ability along the lines of Carol Dweck's theory of intelligence research (1999) as well as math self-concept and values along the lines of Eccles' and colleagues' work with the expectancy-value model of achievement motivation (Wigfield & Eccles, 2002). It was planned that if error reflection did indeed lead to changes in students' sense of belonging to mathematics, measuring these would allow the author to rule out

alternative hypotheses that might arise. It is important to note that the goal of the current study was not to test the expectancy-value model. Rather, these constructs were measured to allow the current author to ensure that whatever changes found in sense of belonging to mathematics that seem to be related to perceived functionality of errors were not better accounted for by students' entity view or their values and expectancies for success. As previously stated, the purpose of the current study was to assess the effect of the error reflection intervention on sense of belonging to mathematics and learning and to consider whether changes in perceived functionality of errors can account for some of these changes.

Definition of Terms

The following is a listing of definitions for key terms used throughout the proposed dissertation.

- *ability beliefs:* one's self-evaluations of how competent one is in a particular area (Wigfield & Eccles, 2002)
- *achievement gap*: the difference frequently observed on standardized measures of mathematics achievement between non-Asian ethnic minority students and their White and Asian peers.
- *attainment value:* how important one deems doing well on a particular task (Wigfield & Eccles, 2002).
- *cognitive load theory*: the theory proposed by John Sweller (e.g., Sweller, 1988) that takes into consideration limitations of working memory capacity to explain why certain instructional strategies are more effective than others.

- *conceptual knowledge (algebra)*: an understanding of the meaning of features in an algebra problem (Booth, 2011).
- *entity theorist:* one who believes that intelligence is a fixed trait, or something that cannot be changed much (Dweck, 1999).
- *expectancies for success:* students' beliefs on how well they will do on a particular task or in a particular domain (Wigfield & Eccles, 2002).
- *expertise reversal effect:* the reversal of cognitive load effects after learners develop expertise; the phenomenon that instruction that is effective with inexperienced learners or learners with low prior knowledge is particularly *ineffective* with experienced learners (Kalyuga, Ayres, Chandler, & Sweller, 2003).
- *factoring*: a method of separating an equation into terms that when the separate terms are multiplied, they produce the original equation.
- *interest in math:* intrinsic value for math; the enjoyment one attributes to the activity of doing math (Simpkins, et al., 2013).
- *intrinsic value:* how enjoyable or interesting one judges a particular task to be (Wigfield & Eccles, 2002).
- *math self-concept:* one's expectancies for success in math and one's beliefs about their ability in math (Wigfield & Eccles, 2000)
- *motivation*: a construct that explains the initiation, strength, and direction of an individual's behavior (Eccles, Wigfield, & Schiefele, 1998).

- *perceived error climate*: a multidimensional characteristic of a learning environment which represents the manner in which errors are perceived to be evaluated and used as an integral part of the learning process (Steuer et al., 2013).
- *perceived functionality of errors for learning*: one's perception of how errors are used as starting points for learning in the classroom (Steuer et al., 2013).
- *perceptions of the importance of math:* how important one deems mathematics material and performance to be overall; includes utility value and attainment value (Simpkins, et al, 2013).
- *procedural knowledge (algebra)*: the ability to use learned procedures to solve algebraic problems (Booth, 2011)
- *quadratic equation*: an algebraic equation in which the highest exponent is a square (2), normally solved by either using the quadratic formula, factoring, taking the square root, or graphing the quadratic function.
- *quadratic formula*: a formula used to find the roots of a quadratic equation using its coefficients. The formula is $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ where *a*, *b*, and *c* are real number and *a* does not equal zero.
- *self-explanation prompts*: prompts that ask students to explain how and why something occurred, or to explain a strategy used within one’s own problem-solving or the problem-solving of a fictitious student within a worked example.
- *sense of belonging to mathematics*: the feeling that one fits in or is a member of the academic community of mathematics; comprised of one's feelings of membership, acceptance, affect, trust, and a desire to fade or avoid active participation in mathematics (Good, Rattan, & Dweck, 2012).

- *sense of school belonging*: sense of school belonging: students' feelings of social belonging, or respect and comfort, in their school (Anderman, 2003; Goodenow, 1993).
- *socioeconomic status (SES)*: a measure representing one's economic and social position in relation to others in one's society, normally based off of one's income, education, and occupation. The current study will have limited information on students' SES, so a proxy will be used. The proxy often used in educational research is whether the student qualifies for free or reduced lunch (FRLP), based primarily upon family income.
- *theory of intelligence:* one's belief that intelligence is either a malleable or stable trait (e.g., Dweck & Leggett, 1988, Dweck, 1999).
- *underrepresented minority (URM) students*: members of racial and ethnic populations that are underrepresented in the domain of mathematics relative to their numbers in the general population. Non-Asian ethnic minority students, mainly Black and Hispanic students, are underrepresented in the field of mathematics in the United States (NSF, 2012). *utility value:* how useful one deems a particular task to be either for a particular goal or more generally
- *worked examples*: written examples of a solution to a problem presented procedurally or with explanations of the process, often presented as being completed by a fictitious student; said to be a prototypic example of cognitive load theory.
- *working memory*: using information in short-term memory in conjunction with information stored in long-term memory to complete a task or solve a problem

(Baddeley & Hitch, 1974). Working memory capacity is said to be limited in the amount and complexity of information to be processed.

Organization of the Remaining Chapters

The remainder of this document is organized into four additional chapters, a reference section and appendices. Chapter 2 presents a review of relevant literature on error reflection and worked examples, students' sense of belonging, and perceived error climate. Chapter 3 explains the methodology used in the current study. This section explains procedures and materials for the study. Chapter 4 presents the results of the data analyses conducted to answer the posed research questions. Chapter 5 presents a discussion of the findings, including limitations and suggestions for future research.

CHAPTER 2

LITERATURE REVIEW

Overview

Student cognition and beliefs about themselves and learning are closely related when it comes to learning middle school mathematics. Therefore, focusing on increasing student learning with consideration to these beliefs may be critical for promoting lasting change in mathematics achievement. Thus, the current study assessed an instructional tool that was hypothesized to increase learning and motivation: Promoting error reflection with the use of worked examples and self-explanation prompts. The following chapter is organized thematically into six major segments to provide a review of relevant literature. This chapter begins with an introduction of error reflection and reviews research comparing having students study either their own errors or teacher or researcher composed errors. The second segment suggests benefits that promoting error reflection may have on gains in conceptual and procedural knowledge in algebra and how worked examples might be used to reach this goal. The third segment provides suggestions as to why students may develop a stronger sense of belonging to math after error reflection. The fourth segment reviews literature on perceived error climate and suggests how and why increases in one particular aspect of classroom error climate, students' perceived functionality of errors, were expected to mediate the influence of the error reflection intervention on both learning and sense of belonging to mathematics. The fifth section provides suggestions on why the error reflection intervention was expected to be particularly useful to underrepresented minority (URM) students. The final segment presents research questions in the context of the current study.

Error Reflection

A central premise of the current study is that promoting error reflection in the classroom would foster algebra learning and a sense of belonging in mathematics. Prior research has suggested the benefits of having students reflect on teacher or researcher-composed errors as opposed to their own. Reflecting on errors may allow students to confront incorrect procedures and concepts and promote correction of these errors and refinement of knowledge. Reflecting on errors may also increase beginning algebra students' sense of belonging to mathematics. As algebra is a particularly troubling topic for many students, failure experiences may be more common within algebra than in other academic subjects that middle school students study. Reflection on errors may foster the idea that errors are made by many students and thus are a normal part of the learning process in math. If making errors is not viewed as an indicator of one not being a 'math person', all students can feel like they are a valuable part of their math class community. This may be especially true for those who feel underrepresented in the domain. The following segment will summarize research on reflection of one's own errors as well as teacher or researcher-composed errors.

Reflecting on actual student errors or researcher composed errors.

A number of scholars suggest that having students correct their errors on assignments such as quizzes and exams increases students' engagement and ability to detect errors and also supports learning (Cherepinsky, 2011; Henderson & Harper, 2009; Yerushalmi & Polingher, 2006). Having students study teacher- or researcher-composed errors may be even more effective than having students correct their own errors (Yerushalmi & Polingher, 2006). Defective prior knowledge can make error detection challenging and newly acquired knowledge difficult to accommodate. Therefore, one might assume that a struggling student is not likely to detect his or her own errors. However, in a naturally cumulative subject such as algebra, avoiding correction of errors simply due to their difficulty would likely prevent students from developing a deep understanding of the material. Highlighting an error for the student may help to alleviate the students' dependence on prior knowledge, while still allowing them the opportunity to refine overgeneralized knowledge, thus leveling the playing field in an algebra classroom. By having students study the same errors, teachers can ensure that all students are exposed to the same instructional content. Studying others' errors may be beneficial because it also provides exposure to multiple perspectives rather than just one's own perspective (Siegler & Chen, 2008). The following is a brief overview of some of the most relevant studies on error reflection in the classroom.

Most research on error reflection in education is qualitative in nature and differs by whether students are asked to detect and correct their own errors or that of a fictional student, formulated by the researcher or teacher. Studies also differ in the purpose of error reflection. While the current study proposes the use of error reflection to refine algebra knowledge, some prior work uses error reflection as a kind of metacognitive strategy or self-assessment. Reflection of errors may allow for correction of one's own incorrect notions within a particular subject matter depending upon how it is used. Two qualitative classroom studies reported improvements in undergraduate physics students' conceptual knowledge when students either explained why they answered an assessment item incorrectly or explained why the error itself is indeed an error (Henderson & Harper, 2009). Both conditions reflected on errors but for different purposes. However, these qualitative comparisons were made by comparing these participants' performance to the performance of students taking the course in prior semesters. No experimental control group was used. Therefore, the influence of instructor/experimenter effects could not be controlled for. Although students reported assessment corrections as being the most helpful class component, the actual influence that these error identification and corrections had on student learning is impossible to ascertain. Further, is it important to note that the two "manipulations" require the participants to engage in two very different tasks; one being metacognitive (i.e., Why did you answer this incorrectly?) and the other being focused on refinement of knowledge (i.e., Why is this answer incorrect?). The current study is particularly concerned with the use of error reflection for refining faulty knowledge.

Cherepinsky (2011) conducted a qualitative study on a grading method utilizing error-correction to foster learning. Undergraduate calculus students completed optional assignments in which they were encouraged to find errors in exam problems marked incorrect. Once found, students made judgments on the severity of the errors and explained and corrected them. Students reported better understanding of the material after completing the error correction assignments. However, as in prior qualitative classroom research (i.e. Henderson & Harper, 2009), causal claims cannot be made due to the lack of a control group and the non-experimental design of the study. Students in this study also reported spending more time on reviewing exams when completing error corrections. It is uncertain whether these error identification and correction assignments promoted learning. It is plausible that spending more time reflecting on errors could allow for refinement of knowledge. However, it is also possible that much of that time is spent trying to locate the error within the problem marked incorrect. One might question the usefulness of this search. Cognitive load theory, as proposed by Sweller (1988; 2006; 2012), might suggest that this search is unnecessary for learning. By clearly labeling which step taken during problem-solving is incorrect, students could spend more effort on correcting that error and refining their knowledge. This may allow students to overcome errors quickly and more easily.

Yerushalmi and Polingher (2006) conducted a classroom study with high school physics students in an attempt to assess the effectiveness of these two methods for guiding students in evaluating errors. One group was presented with fictional students' statements and was asked to identify mistakes in fictional students' work, explain why they were incorrect, and correct the mistakes. Another group identified, explained, and corrected errors made by students on previous exams. Most students correctly identified errors in fictional students' work whereas very few did so for actual student work. Of the students who were able to identify and correct errors, very few in either condition correctly explained why the answers were incorrect. It seems likely that teacher-constructed or researcher-constructed errors may be more effective in fostering learning, because they can be strategically formulated to address common student misconceptions. However, the authors did not assess the effect of these different error-identification conditions on overall learning in the course. It is possible that being prompted to reflect on researcher-composed errors based on commonly made mistakes in algebra problem-solving might be beneficial to both students' learning.

Several suggestions can be made based on these qualitative classroom studies. Firstly, due to the amount of time it can take to detect an error (Cherepinsky, 2011), bringing students' attention to specific errors may save time and also reduce cognitive effort exerted to find the error which may not be helpful to student learning. Second, asking students why a step is incorrect may be more helpful in supporting learning than asking students why they took an inappropriate step during problem-solving. Lastly, reflecting on researcher composed errors may be more helpful than reflecting on an actual student error, either one's own or a classmate's error; reflecting on a researcher (or teacher) composed error may also allow the researcher more control over which errors are being addressed. The next segment will provide a rationale as to how promoting error reflection might foster learning in algebra.

Error Reflection and Learning

The current study is based on the assumption that error reflection can foster learning of both procedural and conceptual knowledge. Procedural skill is, simply, the ability to use learned procedures to solve problems (Booth, 2011). In contrast, conceptual knowledge is an understanding of the meaning of features in a problem (Booth, 2011). Reflecting on errors within a solved problem in algebra may force students to attend to the components of the problem that make those procedures ineffective. Further, reflecting on more conceptually-based errors, or errors that represent a deeper misconception than something more procedurally based, such as a misunderstanding of the meaning of a variable, may assist refinement of conceptual knowledge as well. Although errors are a daily occurrence in students' schoolwork at all levels, not much is known about the potential to learn from these errors. Teachers differ greatly in how they handle their students' errors (Schleppenbach, Flevares, Sims, & Perry, 2007). In many classrooms, errors are explicitly discouraged. Some teachers express the belief that consideration of errors may reinforce students' incorrect procedures or faulty knowledge (Santagata, 2004; Stigler & Perry, 1988). This idea was also proposed early in the education literature from a classic behaviorist perspective. It was thought that students might adopt the errors they were exposed to in incorrect examples (Skinner, 1961). In contrast, research in experimental psychology has suggested that the generation of errors in a practice testing environment is a more effective encoding strategy than further reading of material to be tested (Potts & Shanks, 2014). Errors are frequently discussed in Japanese classrooms, where they are thought to be an integral part of learning (Stigler & Hiebert, 1999). Some researchers suggest that student errors can be prime opportunities to correct student misconceptions (Borasi, 1994; Henderson & Harper, 2009; Yerushalmi & Polingher, 2006).

Ohlsson's (1996) theory of learning from errors purports that explaining one's own incorrect solutions may help learners to identify what features of the problem make the specific step taken incorrect, which can subsequently be used to correct faulty knowledge and fine-tune problem-solving. Ohlsson explains that the very definition of an expert is someone who makes little errors in his or her domain, as opposed to a novice or student who is learning relatively new material. Errors are described as stemming from overgeneralized knowledge structures that are activated at inappropriate times. Ohlsson proposed that errors are detected when there is a discrepancy between what one believes to be true and what one is currently learning. One's prior knowledge is what allows a learner to self-regulate and detect what is most likely to be an error. These errors should be corrected through specialization of knowledge. Reflecting on and explaining a solution may help to identify what features of the problem make the specific step taken incorrect.

This may be used to correct faulty knowledge and fine-tune problem-solving. Further, Oser and Spychiger (2005) proposed that knowing what does not work, or having negative knowledge, may have positive effects on performance.

Students first learning new concepts and procedures are likely to make many errors. Therefore, using errors as a learning device may be most relevant for novices. Students transitioning from arithmetic to algebra often do not understand important concepts which they are expected to have learned prior to beginning algebra, and which are necessary for being successful. Misconceptions that students hold, most commonly about the equals sign, the negative sign, and variables, are important to overcome in order to do well in algebra (Booth, 2011). Certain errors that students make while learning algebra are predictive of later struggles with math achievement (Booth, Barbieri, Eyer, & Paré-Blagoev, 2014).

Many uncertainties exist regarding the use of errors in learning, though mistakes are undoubtedly unavoidable in the learning process. Ohlsson purports that the very definition of an expert is someone who makes few errors in their domain, as opposed to a novice or student who is learning relatively new material. Particularly in the field of mathematics education, students transitioning from arithmetic to algebra often do not understand important concepts which they are expected to have learned prior to beginning algebra, and are necessary for being successful. Misconceptions that students hold, most commonly about the equals sign, the negative sign, and variables, are important to overcome in order to do well in algebra (Booth, 2011). Further, procedural and conceptual knowledge are continuous skills that develop together (Rittle-Johnson & Alibali, 1999; Star, 2005), so it follows that errors made regarding conceptual understanding influence procedural skills and vice versa. As Ohlsson's (1996) theory on learning from errors suggests, faulty prior knowledge can make error detection challenging and newly acquired knowledge difficult to specialize. Highlighting an error instead of relying on students' abilities to detect it themselves, may help to alleviate the students' dependence on prior knowledge, while still allowing them the opportunity to refine overgeneralized knowledge.

As previously discussed, some researchers suggest the importance of utilizing teacher-composed errors focused on common student errors to promote learning (Yerushalmi & Polingher, 2006; Booth et al., 2013), as opposed to actual student errors. There are logistical limitations of having each student within a classroom focus on his or herown errors. Having students correct their own errors may not be as effective at reducing common student errors that are important to overcome in order to progress one's algebraic thinking and learning. Booth and colleagues (Booth, Barbieri, Eyer, & Paré-Blagoev, 2014) suggest that some student errors are more indicative of future struggles in algebra learning than others. Studying teacher- or researcher-composed errors would ensure that all students are focusing on the most important and instrumental errors that need to be overcome in order to progress in algebra learning.

A potentially useful way to enable error reflection is to interleave common student errors in algebra into worked examples labeled incorrect. Targeting errors with self-explanation prompts in addition to the worked examples may promote deeper levels of processing needed to grasp a more conceptual understanding of the error itself. Though much research has demonstrated the benefits of worked examples research, focusing specifically on the benefits of incorrect examples is scarce. The following segment will review worked examples with a specific focus on incorrect worked examples and explain how worked examples may be a prime opportunity to promote error reflection for learning in algebra.

Worked Examples.

As previously stated, a potentially useful way to promote error reflection in algebra would be to use worked examples to display common student errors. By studying incorrect worked examples or worked examples focused on common errors, students have the opportunity to realize that the strategies that they might have used in prior problem-solving are actually incorrect. Relevant research on worked examples will be presented next. First, the origins of worked examples will be presented. Next, research demonstrating the variety of ways that worked examples have been used to influence learning in both laboratory and classroom settings will follow. Then the use of self-explanation prompts in conjunction with worked examples will be discussed. Lastly, the segment will culminate with a review of research involving the use of incorrect worked examples as a learning tool.

As previously mentioned, worked examples will be used as a tool to promote error reflection in middle school algebra students.

Origins of worked examples.

Sweller (2006; 2012) suggests that human cognitive architecture should be considered while developing instructional practices used to enhance learning in the classroom. He uses human cognitive architecture to explain why certain instructional practices are effective. Cognitive load theory, as proposed by Sweller (1988), was inspired by the lack of evidence that problem-solving fosters learning. Sweller suggests that humans' cognitive capacity, as characterized by working memory, is limited in what it can process and that traditional problem-solving overloads that capacity, leaving little room for the acquisition and development of new knowledge structures. Working memory, as proposed by Baddeley and Hitch (1974), can be described as working with or using information in short-term memory in conjunction with information stored in long-term memory to complete a task or solve a problem.

The worked examples effect (Sweller & Cooper, 1985) is often considered a prototypic example of cognitive load theory. Worked examples are written examples of a solution to a problem presented procedurally or with explanations of the process. They are often presented as being completed by a fictional student. Studying worked examples, or worked out problem solutions, is said to reduce extraneous load normally present in traditional problem-solving; while students study these examples, they are not required to generate random processes to solve the problem or attend to irrelevant features. Instead, they can devote their cognitive effort to studying only the relevant aspects for solving the problem and the correct process to acquire and develop schemata for the problem-type studied (Sweller, 2006). The goal is that, after studying multiple worked examples, the student can extract the features of or rule for a problem that are common to other problems of that type, to assist in avoiding cognitive overload, leading to automation of problem-solving in that domain.

The use of worked examples in improving learning in STEM areas has been supported in laboratory studies and classroom studies that will be discussed in the next section. One of the earliest forms of support for the theory is the classic and heavily cited study conducted by Sweller and Cooper (1985) that reported that worked examples save time and reduce errors. The next segment will go into further detail on the various functions of worked examples in laboratory and classroom research.

Effects of worked examples on learning.

The use of worked examples in improving learning in STEM areas has been supported in laboratory studies with students of a wide age range and using various formats and structures. However, the research on worked examples conducted in classroom settings is less extensive. Findings from classroom research present a variety of effects depending upon the topics covered, the presentation format, and the structure of the example itself. Previous research has shown that implementing worked examples, or worked out problem solutions, increased learning in classroom work (Caroll, 1994) and, in some cases, through homework assignments (Ward & Sweller, 1990) in well-structured areas such as mathematics (Zhu & Simon, 1987; Rittle-Johnson & Star, 2007) and science (Pol, Karskamp, Suhre, & Goedhart, 2009). One of the earliest and most heavily cited studies that supports the use of worked examples was conducted by Sweller and Cooper (1985) and demonstrated that worked examples save time and reduce errors. Studying worked examples allows students the opportunity to study only the relevant aspects for solving the problem (Sweller, 2006). This is in contrast to problem-solving, which Sweller suggests requires the generation of random processes to solve the problem or attend to irrelevant features. After studying multiple worked examples, students can extract the features of or rule for a problem that are common to other problems of that type and refine their problem-solving in that domain. Worked examples have been used to scaffold independent problem-solving with backward fading of steps (Renkl, Atkinson, Maier, & Staley, 2002). Worked examples have also been used to promote flexibility in problem-solving by having students compare a pair of worked examples demonstrating the same problem solved in two different ways (Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2009).

A number of factors influence the effectiveness of worked examples. Prior knowledge is an important factor to consider when implementing new strategies into the classroom. Caroll (1994) demonstrated the worked examples effect with low and high achievers. However, Kalyuga and colleagues (Kalyuga, Ayres, Chandler, & Sweller, 2003) demonstrated an expertise reversal effect; worked examples were most effective in promoting learning for learners with low prior knowledge. Thus, worked examples are an effective tool to use for novice learners.

Another critical factor is the structure of the worked examples. When a worked example uses various sources of information in different formats, learning may not occur. In these cases worked examples may sometimes even be deleterious to learning resulting from a split-attention effect (Ward & Sweller, 1990). The split-attention effect occurs when students are required to split their attention and attend to two different sources of information to solve a problem, thus decreasing learning. Atkinson and colleagues (Atkinson, et al., 2000) provide a thorough review in which they make several suggestions of how to effectively structure and use worked examples. The authors emphasize the importance of integrating features within worked examples to avoid imposing a heavy extraneous load such as is found in the split-attention effect. Another important structural feature to consider is the effectiveness of clearly indicating subgoals by either labeling each step or having each step presented visually separate within the worked example. The relationship between the examples within a particular lesson should also be considered. Atkinson and colleagues suggest presenting at least two worked examples per problem-type within a lesson as well as pairing each worked example with a practice problem to be solved. These suggestions will be applied to the design of the worked examples for the current study. An additional structural feature that influences the effectiveness of worked examples is the use of self-explanation prompts. The next segment will briefly review the literature on the use of self-explanations in combination with worked examples.

The use of worked examples with self-explanation prompts.

Worked examples are often supplemented with self-explanation prompts. Self-explanations are prompts that ask students to explain how and why something occurred, or to explain a strategy, and have been shown to reduce errors (Siegler & Chen, 2008). Self-explanations force learners to make their knowledge explicit, and in doing so, they are required to generate and integrate new knowledge with prior knowledge (Chi, 2000). While aural self-explanation prompts have been shown to be effective on their own (e.g., Roy & Chi, 2005), they are also effective in written form when used in combination with worked examples (Berthold, Eysink, & Renkl, 2009). More specifically, worked examples in combination with written self-explanation have been found to be particularly useful in boosting conceptual knowledge in algebra (Booth, Lange, Koedinger, and Newton, 2013). As can be expected, the quality of learners' self-explanations is related to the learning that has occurred (Renkl, 1997) and spontaneous self-explanations are more characteristic of high-ability students (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). One example of what would be considered a high-quality explanation is that of a learner focusing on the deep structure of the problem, such as principled knowledge, rather than superficial features of the problem. However, Chi and colleagues have also demonstrated that prompting students to self-explain results in more learning regardless of prior ability in comparison to not prompting students to self-explain (Chi, de Leeuw, Chiu, & Lavancher, 1994). The current study employs written self-explanation prompts in an attempt to encourage integration of new knowledge gained through the worked examples with prior knowledge.

The use of incorrect worked examples.

As previously described, there has been much support for the use of worked examples from both laboratory and classroom studies (e.g., Carroll, 1994; Sweller & Cooper, 1985; Pol, Karskamp, Suhre, & Goedhart, 2009; Ward & Sweller, 1990; Zhu & Simon, 1987). The effectiveness for worked examples is often increased by supplementing the examples with self-explanation prompts (Berthold, Eysink, & Renkl, 2009; Booth, Lange, Koedinger, & Newton, 2013; Siegler & Chen, 2008). Some prior research has suggested the benefits of comparing correct and incorrect worked examples (Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2009). However, the influence of incorrect worked examples alone on learning and motivation has not been thoroughly explored. Booth and colleagues (Booth et al., 2013) have suggested that studying incorrect worked examples benefits encoding of algebraic equations. Heemsoth and Heinze (2014) only saw benefits of incorrect worked examples for students with high prior knowledge. In contrast, Adams and colleagues (Adams, McLaren, Durkin, Mayer, Rittle-Johnson, Isotani, & van Velsen, 2014) found that studying incorrect worked examples within a cognitive tutor was equally beneficial for both students with low and high prior knowledge. More recently, Barbieri and Booth (Barbieri & Booth, under review) found that effectiveness of incorrect worked examples did not vary by prior knowledge. Rather, the authors found learning benefits of the incorrect worked examples for students with certain motivational beliefs. Particularly, students who had low expectancies for success in mathematics demonstrated the most conceptual learning gains in when studying incorrect worked examples as opposed to correct examples or traditional problem-solving. These inconclusive results regarding the effectiveness of incorrect worked examples on learning, and whether they are equally effective for students with low and high prior knowledge, suggest that further exploration of both the cognitive and motivational benefits of using incorrect worked examples as a learning tool. Error reflection through the use of incorrect worked examples could be a prime opportunity to allow students to learn from errors. The following segment provides a brief overview of some of the most relevant studies including incorrect worked examples.

Most recent research that includes incorrect worked examples uses them in combination with correct worked examples (Durkin & Rittle-Johnson, 2012; Grosse & Renkl, 2007) making the effect of error reflection within the incorrect worked examples themselves difficult to ascertain. Rittle-Johnson's earlier work focused on using comparisons of correct and incorrect worked examples to promote the development of procedural flexibility (Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2009). More recently, Durkin and Rittle-Johnson (2012) found that comparing incorrect and correct worked examples in which the same problems were solved, promoted students' learning of the correct procedures and concepts of decimal magnitude, regardless of students' prior knowledge. Students exposed to both correct and incorrect worked examples also discussed correct concepts more often than those in the correct only condition.

Interestingly enough, improvements in decimal magnitude knowledge for students in the combination condition (i.e., correct and incorrect) were not demonstrated until a delayed post-test. Durkin and Rittle-Johnson suggest that students may have needed more time to process what they learned during the individual researcher-implemented intervention. The current study planned to employ a delayed post-test in addition to a post-test to assess the effectiveness of the intervention immediately as well two weeks after the intervention. However, due to unforeseen circumstances during data collection, the data analyses presented in Chapter 4 do not include delayed post-tests as they were not completed during the allotted time for the study. Implications are discussed in Chapter 5.

As previously noted, it is important to consider that the aforementioned study focused on displaying the benefit of comparison between correct and incorrect procedures. While Durkin and Rittle-Johnson do provide evidence for the usefulness of incorrect worked examples, this is only demonstrated in combination with correct worked examples. It is unclear whether error reflection alone is what led to these benefits or whether a combination of reflection on both correct and incorrect procedures and concepts was the root cause. The current study is concerned with the use of incorrect worked examples alone to promote error reflection and changes in both conceptual and procedural knowledge.

Several recent studies that have examined the influence of incorrect worked examples alone vary in methodology and have presented inconclusive findings. Booth and colleagues (Booth, Lange, Koedinger & Newton, 2013) conducted two classroom studies, using the Algebra I Cognitive Tutor curriculum and compared an all-incorrect worked examples group to a correct worked examples group and a correct-incorrect combination group. Incorrect worked examples were designed to target student misconceptions about features within the problems. Each of the worked examples conditions utilized highly structured and scaffolded self-explanation prompts for which students used pull-down menus to answer. In Experiment 1, students who worked with all correct, all incorrect, and a combination of correct and incorrect worked examples all increased their conceptual understanding, in comparison to a problem-solving control group. However, in Experiment 2 the combined incorrect and correct examples group demonstrated more learning than the all correct and all incorrect worked examples groups. Nevertheless, the all-incorrect group made less encoding errors of algebraic equations than the correct only group. Booth and colleagues suggest that studying and explaining incorrect worked examples encourages students to refine their own faulty knowledge. The current study makes a similar suggestion; that reflecting on errors through the use of incorrect worked examples and self-explanation prompts might promote the refinement of students' knowledge in algebra.

More recent work on incorrect worked examples has demonstrated that their effectiveness depends upon several student factors. Heemsoth and Heinze (2014) conducted an experiment to determine the effect of reflecting on errors in the form of incorrect worked examples on sixth grade students' fraction knowledge. The authors found that only students with high prior conceptual and procedural knowledge were able to show learning gains with reflecting on incorrect worked examples. The authors explain that students may benefit more from incorrect worked examples once they have developed a stronger foundational knowledge of the to-be-learned material. They suggest this be achieved through the use of correct worked examples earlier on and incorrect worked examples later on in the learning process. This suggestion is also made by Grosse and Renkl (2007) in earlier work comparing the effectiveness of an all-correct worked examples condition to a correct-incorrect worked examples combination condition. However, the conditions may not have been comparable in regards to level of difficulty for the incorrect condition in Heemsoth and Heinze's study or for Grosse and Renkl's (2007) correct and incorrect combination condition.

In Heemsoth and Heinze's correct condition, students were shown correct worked examples and asked to describe the solutions and why they are correct before solving analogous problems next to them. In contrast, students in the incorrect condition were shown incorrect worked examples and asked to first identify the solution step that was incorrect and explain why it was correct. They also were asked to explain why they thought the fictitious student made the error and then to solve the problem correctly. As is evident, students in the incorrect worked examples group not only had to locate the error themselves and explain it but also to provide reasoning as to why the student made the error before correctly solving it. This adds a metacognitive component to the intervention that students in the correct worked examples condition were not asked to engage in. While students with high prior knowledge may have benefitted from this intricate intervention, students with low prior knowledge may have experienced high levels of cognitive load that did not contribute to their learning of the correct procedures. Rather than solely reflecting on errors and focusing their attention on the key features of the problem that make a specific step incorrect, students also had to consider why a fictitious student made the error. Thus, these experimental manipulations were not solely focusing on promoting the reflection of the errors themselves but also on the students making the errors. A similar issue is presented in Grosse and Renkl's (2007) study.

In Grosse and Renkl's (2007) incorrect and correct combination condition, the self-explanation prompts varied as well. For correct worked examples, participants were asked to explain the procedure used to solve the problem and why it is correct to use that procedure. Participants were also asked to explain how one would avoid making errors. Incorrect worked examples varied by whether the incorrect step was highlighted or not, so participants were first asked to either locate the incorrect step and explain it, or explain the incorrect step highlighted. They were then to explain what the correct solution would be and to provide an example of a case in which the incorrect step taken would be an appropriate step to take. Whether this latter prompt is of equal difficulty in comparison to the former prompt explained for the correct worked examples is questionable. As in the Heemsoth and Heinze (2014) study, it is possible that the prompts Grosse and Renkl used for incorrect worked examples required students to possess more refined knowledge or more abstract forms of thought to engage in using the incorrect worked examples than what the prompts for the correct worked examples require. These more intricate self-explanation prompts may also explain why Grosse and Renkl did not find them to impact learning. Participants with low prior knowledge may not have been equipped to engage in the generative processing that might lead to the refinement of knowledge. The authors did demonstrate that highlighting errors within an incorrect worked example was more helpful for students with low prior knowledge than requiring students to find the error themselves. The conditions in the current study are designed so that they only vary on whether students are prompted to reflect on errors displayed in incorrect worked examples or hypothetical errors that could occur in a correctly solved worked example. This should make the conditions more comparable and allow for more precise analysis of causality. This is further described in the chapter on methodology. However, the influence of prior knowledge on the effectiveness of the conditions are factored into the analyses. Although interactions with prior knowledge should be taken into account when implementing any instructional intervention, an error reflection intervention may be helpful for students of varying levels of prior knowledge if presented in a manner that is accessible to learners at varying levels.

A recent study conducted by Adams and colleagues (Adams, McLaren, Durkin, Mayer, Rittle-Johnson, Isotani, & van Velsen, 2014) found that using incorrect worked examples within a computer-based intervention on decimal knowledge was just as effective for students with low prior knowledge as it was for those with high prior knowledge. Though participants did not experience significant learning gains immediately, the incorrect worked examples group outperformed the problem-solving control group on a delayed post-test. This was found for students with high prior knowledge and low prior knowledge. A limitation of this study is that Adams and colleagues compared their incorrect worked examples intervention to a problem-solving control group rather than a correct worked examples group. Thus, based on these results, one cannot assume that incorrect worked examples would also be more beneficial than correct worked examples. Further, there were several differences between the problem-solving control group and the incorrect worked examples group that makes zeroing in on exactly which aspect of the intervention, either alone or in combination to the other factors, results in the learning gains found. For example, although the problems displayed were similar, students in the incorrect worked examples condition were asked to explain what the fictitious student had done incorrectly, correct the error, received feedback on their correction, and then explained the correct solution. In contrast, students in the problem-solving control group were given a problem to solve, received feedback on their answer, and then were asked to explain the correct solution. Due to these differences, isolating causal variables within the intervention would be difficult. The authors argue that consideration of cognitive load led them to the decision of excluding a correct worked examples condition. It was thought that the two conditions used were more comparable from a cognitive load standpoint whereas the correct worked examples condition would require much less effort on behalf of the student. However, whether or not the conditions used were comparable in regards to the amount of effort or the level of processing that the students engaged in is open to debate. Students in the incorrect worked examples condition may have had access to more knowledge to be learned than students in the problem-solving control group; they were asked to explain both the incorrect and correct procedure whereas the students in the problem-solving control group were only asked to explain the correct answer. Adams and colleagues acknowledge this difference when explaining that the incorrect worked examples condition may have led to greater learning due to the greater generation of knowledge that resulted from the experimental condition. Were the learning gains found by Adams and colleagues due to reflecting on incorrect procedures themselves which then initiates the refinement of knowledge? If not, may the benefits stem from greater opportunities to generate explanations? If so, can this also be done with correct worked examples? If not, is there something distinct about incorrect worked examples that allows this learning to occur?

The current study employs conditions that are designed to be more comparable in the amount of effort required to engage in the task. The conditions vary on the type of task (i.e., reflection on errors versus reflection on correct procedures) to ensure that findings such as those found in prior research on incorrect worked examples (e.g., Adams et al., 2014; Heemsoth & Heinze, 2014) are not simply artifacts of inadequate control or comparison conditions.

One study that was conducted to compare a solely incorrect worked examples condition to a solely correct worked examples condition and a problem-solving control condition found no overall main effect of condition on learning (Barbieri & Booth, under review). However, students with certain characteristics benefitted differentially from the worked examples conditions. Specifically, students with low expectations of their own competence in mathematics demonstrated more learning gains in conceptual understanding in the incorrect worked examples group than in the correct worked examples group and the problem-solving control group. Students who were members of an underrepresented ethnic minority group also demonstrated some increases in motivation due to the worked examples condition. This is discussed in a later segment. Nonetheless, the design of this particular study also seems to have conditions that have more than one varying factor. The self-explanation prompts in the aforementioned study were not completely comparable in that sometimes students were asked to explain different concepts displayed in the examples due to the correctness or incorrectness of the example. It is apparent that results are inconclusive on whether incorrect worked examples are just as or more effective than correct worked examples or traditional problem solving in boosting procedural skill and conceptual knowledge in algebra. The worked examples designed for the current study will ensure that conditions will clearly vary on specific factors that can be isolated in order to enable and justify claims of causality about promoting error reflection.

Error Reflection and Sense of Belonging to Math

Another potential benefit from having students study researcher composed errors as opposed to their own errors is for motivational purposes. Reflecting on errors may alter students' motivation. Motivation is a broad psychological construct that refers to the initiation, direction, intensity, and persistence of behavior (Eccles, Wigfield, & Schiefele, 1998). During adolescence, there is a decline in math achievement accompanied by a decline in motivation for students in the United States (Wang & Pomerantz, 2009). Educational researchers often want to know what motivates students as well as how students' motivation influences their cognition. Beliefs that students hold about themselves and about learning may impact their own learning. Due to the impact that student motivation has on learning, particularly in mathematics, a cognitive intervention aimed at adolescent mathematics students that does not consider the influence of motivation would be incomplete.

One motivational construct that may be particularly important for students learning mathematics is their sense of belonging to the *math* community (Good, Rattan, & Dweck, 2012). The importance of a sense of belonging to the *school* community has been established in terms of its influence on students' school experience (for a review, see Osterman 2000); students' sense of belonging to a *math* community in particular (Good, Rattan, & Dweck, 2012) has also been found to be a potentially important factor in mathematics learning. Unfortunately, students' sense of school belonging in general has been found to decline during adolescence (Anderman, 2003). Changes in sense of belonging to the *math* community during adolescence have not been studied. Since research on the sense of belonging to the math community is relatively new and understudied, the following section will also include research on sense of belonging in general as well as sense of *school* belonging.

A sense of belonging and connectedness within a given domain influences students' motivation to achieve (Baumeister & Leary, 1995; Ryan & Deci, 2002). Increases in sense of school belonging are related to increases in achievement motivation (Goodenow, 1992; 1993; Goodenow & Grady, 1993), particularly increases in an interest of learning and understanding (Anderman & Anderman, 1999; Seifert, 1995). Sense of *school* belonging has been found to decrease during adolescence (Anderman, 2003; Gutman & Midgely, 2000). Most research on students' sense of belonging in an academic domain has focused on *social* belonging or feelings of social connectedness during achievement experiences (Pickett, Gardner, & Knowles, 2004; Walton & Cohen, 2007). While the importance of a sense of belonging to the school community has been established as having an influence on students' school experience (for a review, see Osterman 2000), the influence of students' sense of belonging to the academic domain of mathematics on learning has not been thoroughly explored. Social connectedness has been found to influence students' academic achievement, especially for minority students (e.g., Walton & Cohen, 2007). However, the current study is concerned with students' feelings of belonging to the academic domain of mathematics.

Good and colleagues (Good, Rattan, & Dweck, 2012) found a domain-specific sense of belonging to mathematics predicted female college students' course performance and intent to pursue a STEM degree. More specifically, Good and colleagues found that decreases in sense of belonging to mathematics, due to a stereotype threat manipulation that emphasized mathematics as a fixed trait, was particularly predictive of women's intent to further their mathematics education. Women who experienced a protective manipulation against stereotype threat, that mathematics ability could be developed, maintained their high sense of belonging and intent to further their mathematics education. As mentioned previously, much research has been conducted on stereotype threat and several interventions have been created to reduce feelings of stereotype threat (Aronson, et al., 2002; Cohen, et al., 2006; Cohen, et al., 2009). While it is likely that URM students experience a greater history of stereotype threat experiences in their academic career, and while it is even possible that repeated experiences of stereotype threat may lead to lowered sense of belonging, the current study does not attempt to manipulate stereotype threat in an effort to alter sense of belonging to math. Though exploring the relation between stereotype threat within an error reflection intervention may provide useful information relevant to URM student achievement, it is assumed that experiences of stereotype threat may not be especially high in the sample for the current study, as the students recruited are from a diverse middle school and district. Rather, it is assumed that a sense of belonging to math may be particularly important to students who are underrepresented in the domain of mathematics, namely underrepresented minority students, due primarily to their underrepresentation status and the influence this may have on their perceptions of the domain itself. Smith and colleagues (Smith, Lewis, Hawthorne, & Hodges, 2012) found that when presented as a male-dominated domain, female students judged a fictional graduate program to require more effort on their behalf and also deemed it a less interesting field of study to pursue. However, when effort was presented as a normal part of achievement, female students' sense of belonging to a STEM field increased. In a similar way, the current study sought to assess changes in sense of belonging to math for all students as well as for URM students in particular as a result of error reflection. It is assumed that an intervention focused on promoting error reflection may alter students’ perceptions of errors to one in which errors are seen as a normal part of the learning process. This in turn may then lead to students perceiving the domain of mathematics as one that is more inclusive and one in which anyone may succeed in. This rationale is further detailed below.

These findings suggest that one’s perception of mathematics ability and learning may be related to their sense of belonging to mathematics. It is important to note that findings from the previously reviewed studies (i.e., Good et al., 2012; Smith et al., 2012) were found for female undergraduate students. Younger students, particularly young adolescents, may have an even greater need for a sense of belonging, given the achievement and motivational declines found in the US during the transition to middle school, particularly in mathematics (Wang & Pomerantz, 2009). As suggested earlier, it is likely that failure experiences within a given domain may lower feelings of belonging in that domain. As algebra is a particularly troubling topic for many students, failure experiences are probably more common for many students in the early stages of algebra learning, than in other academic subjects that middle school students study. Thus, it was expected that sense of belonging to math in particular would be generally low on average for middle school algebra students in the US. Students’ sense of *school* belonging declines during adolescence. Students’ sense of belonging to the academic domain of *mathematics* in particular may also decline greatly during adolescence as well. If this is the case, reflection on errors may be particularly beneficial to sense of belonging to mathematics for students in the early stages of algebra learning.

Having students study their own errors and increasing the salience of these errors may present the potential risk of adopting negative views of their own ability. Tulis and Ainley (2011) demonstrated that students react differently to failure experiences. Students who view their mistakes in mathematics in a positive manner and who see feedback on errors as a learning opportunity experience more positive emotions after a failure experience. Mastery oriented students are less likely to experience negative emotions after a failure experience, especially when students attributed success in mathematics to increased effort (Tulis & Ainley, 2011). Therefore, students may vary greatly in whether they benefit from studying and correcting their own errors. These variations may be due to prior knowledge, motivational orientations and emotions. Thus, studying researcher-composed errors presented as errors of a fictitious student may help students to focus on the error as a learning tool as opposed to a reflection of their ability. An error reflection intervention may foster the idea that errors are made by many students and thus are a normal part of the learning process in math. If making errors is not viewed as an indicator of one not being a ‘math person’, all students can feel like they are a valuable part of their math class community. Similarly, changes in students’ beliefs about errors themselves may influence the impact of promoting error reflection on belonging as well. As URM students may be more at risk for the negative influences of stereotype threat (Steele & Aronson, 1995), sense of belonging to math may be even more important for these particular students. The next segment includes a review of relevant literature on perceptions of errors in the classroom.

Perceptions of Errors

Prior research has demonstrated that the way students react to errors and failure is highly influenced by their personal achievement motivation (Wentzel & Wigfield, 2009). While an error reflection intervention may have motivational and learning benefits, students' perceptions of errors are also important to consider. A recent study demonstrated that middle school math students' perceptions of how errors are evaluated in the classroom by their teacher, classmates, and themselves, or the *perceived error climate,* influences how they react to struggles in problem-solving (Steuer, Rosentritt-Brunn, & Dresel, 2013). More specifically, *perceived error climate*, predicts students' adaptive reactions to struggles in mathematics problem-solving to a greater extent than does students' achievement motivation (Steuer, et al., 2013). The authors suggest that teachers consider how they can create classroom environments that promote the functionality of errors. However, little research has actually supplied supporting evidence or suggestions on why or how to create a classroom in which errors are perceived as vital to the learning process.

At the time of writing, Steuer and colleagues are the only researchers who have assessed students' perceptions of errors and their relationship to achievement motivation and task behavior. Though some research suggests that how teachers' react to errors influences their students' reaction to errors (Tulis, 2013), how students' perceptions of errors influences their own learning, and whether this can be altered, has not been addressed. The current study expected that error reflection would alter students' perceptions of errors to that of a functional learning tool. This change was also expected to coincide with learning in an algebra classroom.

Most research on error climate comes from the field of organizational management and is often focused on worker productivity or employee collaboration of group projects (e.g., van Dyck, Frese, Baer, & Sonnentag, 2005). Research from this field has demonstrated that adults in work settings are often defensive of their own errors unless they work in a collaborative environment focusing on problem-solving; analyzing errors in this environment was shown to be an effective strategy in reducing future errors (Tjosvold, Yu, & Hui, 2004). How these findings apply to the classroom is still unknown. The few studies that have explored error climate in classrooms have not measured learning, but have instead focused on students' motivation and reaction towards errors. Tulis (2013) examined the impact that teachers' attitudes towards making mistakes had on their students' emotions in the classroom. Tulis found teachers' negative or maladaptive reactions towards student errors, such as criticizing the student or not allowing the student enough time or assistance to correct their own error, was predictive of students' own negative attitude towards making mistakes in the classroom. Interestingly, teachers' maladaptive reactions to errors were more common in mathematics classrooms than in German language classes or economics classes. It is important to note that Tulis used more objective measures of teachers' reactions to errors, not measurements of students' perceptions of how errors are handled in the class. Also, Steuer and colleages (2013) demonstrated that perceived error climate varies both between and within classrooms. As it is likely that students react differently to errors based on individual differences such as prior achievement and motivation, it may be more informative to measure individual students' *perception* of errors made in the classroom. Even for students with the same teacher, this perception may vary from student to student within a classroom. Thus, there is a particular need for research exploring error climate in mathematics classrooms with consideration of the effect of students' perceptions of the functionality of errors in their classrooms on their actual learning.

Functionality of errors as a mediator of the effect of an error reflection intervention on learning and sense of belonging to math.

Steuer and colleagues (2013) found that students' perceptions on how errors are evaluated in the classroom, or the *error climate*, predicts students' adaptive reactions to struggles in mathematics problem-solving to a greater extent than does students' achievement motivation (Steuer, Rosentritt-Brunn, Dresel, 2013). As these beliefs seem to play an important role in the mathematics classroom, it is important to consider how these beliefs might alter the effectiveness of cognitive interventions. The perceived error climate scale, created by Steuer and colleagues, measures various dimensions of classroom error climate. These subscales measure a variety of views including students' report of how errors are tolerated by the teacher and classmates, the students' willingness to risk being wrong, and the use of errors as learning tools. However, as the current study is mainly concerned with learning from errors and how students' views of errors within a mathematics learning environment might change after an error reflection intervention, the current study focuses on the *perceived functionality of errors* subscale. This subscale was suggested by Dr. Gabriele Steuer, the first author of the aforementioned study, in personal correspondence, due to its high consistency with the overall measure as well as its high predictive utility in predicting personal achievement motivation and adaptive reactions toward failure experiences.

Students who are more focused on learning have been found to interpret failure in a positive way or demonstrate constructive or adaptive responses (Clifford, Kim, & McDonald, 1988; Meyer & Turner, 2006). Tulis and Ainley (2011) found that students who felt positive emotions after making errors were more likely to be mastery oriented. Although the current study will not assess emotions after failure experiences, it is likely that students' subjective experience or views of the functionality of errors may change after an error reflection intervention. Promoting error reflection in the classroom may present students with the idea that errors are functional or useful for learning. Increases in the usefulness of errors as a learning tool may mediate the influence of the error reflection intervention on both learning and sense of belonging to mathematics. Reflection on displayed (incorrect examples) or hypothetical (correct examples) common errors was expected to promote the idea that errors are indeed an important part of the learning process. This expected change in students' perceptions of errors was then expected to foster greater utilization of error reflection itself which could lead to the refinement of conceptual and procedural knowledge. Increases in students' perception of the functionality of errors for learning in the classroom were also expected to foster greater feelings of belonging to mathematics. As previously suggested, the error reflection intervention was expected to foster the idea that errors are a normal part of mathematics learning which may then foster the idea that all students can belong to the mathematics community regardless of the struggles they face. Further, if students perceive errors as a more functional aspect of the algebra learning environment, this may provide an added benefit of the error reflection intervention on sense of belonging to mathematics.

URM student achievement

As previously noted, there is also a well-documented achievement gap between Black and White students, and Hispanic and White students. In the United States, Black and Hispanic students generally do not perform as well as their White and Asian counterparts on achievement tests, particularly in mathematics and science (National Center for Education Statistics, NCES, 2013). These students are often underrepresented in STEM fields (NSF, 2012). Despite increased acknowledgement of these issues, ethnic and income achievement gaps persist. There have been several explanations of the differences found between White and URM students. These explanations include genetic deficit views (e.g., Herrnstein & Murray, 1994), inequality in school resources and quality (Darling-Hammond, 2004; Kozol, 2005; Leventhal & Brooks-Gunn, 2000; Oakes, 2004), differences in family and neighborhood environments (Brooks-Gunn & Markman, 2005), and African American students' academic disengagement (Ogbu, 2003). Genetic deficit views have been largely out of favor due to the lack of evidence that supports and explains the differences found between White and URM students' achievement. Further, views proposing academic disengagement or a devaluing of school characteristic of minority students have been contradicted with findings that minority students and their parents report highly valuing school in general (Stevenson, Chen, & Uttal, 1990). The current understanding is that these differences arise mainly after infancy, suggesting the importance of different environmental factors within the home and school on achievement (Dickens, 2005).

Achievement that is lower than expected for students of varying ethnic and socioeconomic backgrounds is likely related to a lack of resources (Darling-Hammond, 2006; Duncan & Magnuson, 2005; Fryer & Levitt, 2004) which may include both early opportunities to learn and to be motivated to do so. There are large disparities in socioeconomic status between Black and White families in the US (Reardon, 2003). García Coll and colleagues (García Coll et al., 1996) suggest the importance of considering the negative effects of social stratification on the development of underrepresented minority children. While the comprehensive approach that García Coll and colleagues suggest to take when considering minority student achievement is beyond the scope of the current study, the previously cited research presents the complex situation that arises when trying to consider the many factors that influence URM students' achievement and specifically learning in mathematics.

As ethnic minority status is mainly confounded with low socioeconomic status in the United States, differences in achievement due to lack of resources and opportunities not confounded by SES are difficult to distinguish. However, Burchinal and colleagues (Burchinal, McCartney, Steinberg, Crosnoe, Friedman, McLoyd, & Pianta, 2011) demonstrated that differences in early mathematics skills between low-income Black and low-income White children are mainly due to differences in family life, child care, and school experiences. The authors found that instructional quality was even more impactful for Black children than White children. The authors demonstrated that Black children from low-income homes have a number of different experiences than their White low-income counterparts. When considering factors such as school quality and other environmental variables, the gap in math achievement was accounted for. This suggests that Black and White children may have very different school experiences which lead to differences in preparation for later learning, particularly for children from low-income homes. The current study expected particular benefits from the error reflection intervention for URM students who often begin school without the necessary skills to succeed due to the previously mentioned social issues.

Benefits for URM students from an error reflection intervention.

It was expected that any differential effects that error reflection on URM students would be accounted for by sense of belonging to mathematics and prior knowledge. URM students were expected to have a lowered sense of belonging to mathematics due to being a member of an underrepresented group. URM students are also at risk of having lower prior knowledge due to the issues of social stratification in the US previously reviewed including the differences in quality instruction during early childhood.

Although much research has found that there is a general decline in motivation during adolescence (Wang & Pomerantz, 2009), as well as a development of a more realistic judgment of one's expectations of success that matches actual competence in a domain, especially in students considered gifted, (e.g., Pajares & Graham, 1999), these particular findings have been suggested not to apply to African-American students (Graham, 1994). After reviewing countless studies on African-American achievement motivation, Graham reported that, contrary to popular belief at that time, African American students do have high expectations of success. African American students have been reported to be somewhat optimistic in their competence beliefs in comparison to higher-achieving White students (Stevenson, Chen, & Uttal, 1990). Thus, African American students' expectancies for success are not a particularly useful motivation construct in terms of predictive power for actual achievement. However, more recently, Wang (2013) found that 12th grade minority students' math self-efficacy was influenced to an even greater extent by their 10th grade achievement than White and Asian students (Wang, 2013), which then influenced their intent to pursue a STEM degree. Thus, it may be imprudent to assume that URM students' motivation is not related to their own achievement. Therefore, it may be important to not only focus on increasing learning of the material but on simultaneously increasing students' sense of belonging in mathematics.

Good and colleagues (2012) suggest that stereotyped individuals may have a lower sense of belonging within the corresponding domain leading to lower levels of interest and desire to pursue future careers in that domain. It is suspected that many underrepresented minority middle school students in the US may have a lowered sense of belonging to math in general, in comparison to White and Asian students, due to underrepresentation of Blacks and Hispanics in mathematics (NSF, 2012) and perhaps even a general awareness of the existing achievement gap or racial discrimination.

There have been several interventions aimed at increasing students' sense of belonging to school in general or to a particular academic domain. Much like Good and colleagues, Walton and Cohen (2007; 2011) also suggested that students who experience negative stereotypes may be more concerned with feelings on belonging than students who are not negatively stereotyped. These researchers found that Black college students who were presented with information on college students' altering sense of school belonging during the difficult transition to college showed increases in GPAs compared to those who were not presented with information on belonging. A focus on mastery and high expectations held by teachers has been found to positively influence Hispanic middle students' sense of school belonging (Stevens, Hamman, Olivárez, 2007). However, Stevens and colleagues did not measure actual achievement. These two sample interventions did not measure belonging to a particular domain, so it is unclear of whether this increase in sense of belonging translated to increases in learning and achievement particularly for mathematics. In general, research on school belonging has found that respectful student-teacher relationships have been found to influence sense of school belonging (Anderman, 2003; Fine, 1991; Morrison, Cosden, O'Farrel, & Campos, 2003). While feelings of connectedness are undoubtedly important for school success, sense of belonging to the particular domain of mathematics may involve additional influential factors. Considering the difficulties that many students experience when transitioning from arithmetic to algebra, increasing a sense of belonging to math may require more focus on learning processes and overcoming challenges than increasing school belonging in general.

This lowered sense of belonging may differentially influence the way that they experience failure experiences or view the functionality of errors for learning. Minority students’ sense of belonging to mathematics may be influenced by their underrepresentation status. Students’ underrepresentation status and awareness of their groups’ underperformance in math in general may influence the manner in which URM students perceive errors and respond to failure experiences. Rather than presenting students with information about other students’ belongingness in hopes that they internalize these feelings themselves, the current study focused on promoting error reflection to foster the idea that making errors in mathematics is part of the learning process. If making errors is not viewed as an indicator of one not being a 'math person', all students can feel like they are a valuable part of their math class community. This was expected to be true for those who feel underrepresented in the domain. It was also expected for non-URM students but thought to be particularly beneficial for non-Asian minority students as members of an underrepresented and stigmatized group.

At the time of writing, there has been no empirical research specifically exploring URM students' perceptions of errors in the classroom. The current study aimed to provide an initial exploration into the relationship between URM status and perceptions of errors, although the unexpectedly few number of URM students did not allow for this analysis (see Chapter 4). While it would be interesting to consider whether the influence of changes in students' perceptions of errors on learning and motivation differs by URM status, no research to date directly suggests reasons to believe that differences will be found based on URM status. If differences were found, it would be important to consider the influence of prior knowledge and sense of belonging which may be found for students in general, regardless of URM status, but may be particularly prevalent for URM students considering social stratification and achievement differences found between URM and non-URM students historically in the US.

The Current Study

The current study sought to answer several research questions detailed below.

Research Question 1.

The first research question was twofold:

(1) Does promoting error reflection within a worked examples intervention improve middle grade students' algebra learning?

(2) Are there differential effects of promoting error reflection with the error displayed in comparison to the error simply mentioned? If so, would prior knowledge moderate the effect of condition on learning?

First, the current study assessed whether reflection on errors within a worked examples intervention improves middle grade students' algebra learning. Reflecting on errors may be potentially useful in promoting the refinement of students' knowledge in algebra. The current study assessed whether a combination of *incorrect* worked examples and self-explanation prompts focusing on errors can be just as beneficial to learning in comparison to a worked examples intervention that has only *correct* worked examples coupled with self-explanation prompts that ask students to explain a common error that is not displayed. Conditions varying these aspects of the intervention are further discussed in Chapter 3.

Reflecting on errors while explaining incorrect examples or common errors in algebra is thought to provide two types of support for learning: it should allow students to fully accept that the strategies they might have used are wrong, increasing negative knowledge (Oser & Spychiger, 2005) and it should force them to attend to the components of the problems that make the solution incorrect (Booth et al., 2013). Highlighting an error instead of relying on students' abilities to detect it themselves, may help to alleviate the students' dependence on prior knowledge, while still allowing them the opportunity to refine overgeneralized knowledge. Further, using researcher composed errors as opposed to individual students' own errors allows one to control the errors that are targeted for learning. By presenting students with common procedural or conceptual errors embedded within worked-out problems and clearly labeling these errors, students could focus their effort on correcting the error and refining their knowledge. This may allow students to overcome errors quickly and more easily than in a situation in which students must locate and correct errors, as was suggested in previously reviewed qualitative studies on learning from errors (Cherepinsky, 2011; Henderson & Harper, 2009). This reflection of errors may improve both procedural and conceptual knowledge, both of which are valued components of mathematics proficiency (Kilpatrick, Swafford, & Findell, 2001; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

However, promoting error reflection through the use of self-explanation prompts alone paired with correct worked examples may also be beneficial. Previous research demonstrating the expertise reversal effect (Kalyuga, et al., 2003) suggests that traditional correctly-solved worked examples may be particularly useful for students with low prior knowledge. Some prior work also demonstrates that incorrect worked examples may be most beneficial for students with high prior knowledge (e.g., Heemsoth & Heinze, 2014). Although it was hypothesized that the error reflection intervention will be useful for all students, the effectiveness of the format used to promote error reflection may vary based on prior knowledge. It was hypothesized that students with high prior knowledge may demonstrate particular benefits from error reflection in the form of incorrect worked examples. Conversely, students with low prior knowledge were expected to benefit most from error reflection in the form of self-explanation prompts paired with correct worked examples. Correct worked examples were thought to provide the support needed to learn accurate procedures and concepts. At the same time, the self-explanation prompts focused on potential errors by problem-type were thought to provide opportunities to support refinement of knowledge and help students avoid using ineffective strategies as previously hypothesized.

As previously explained, the original use of worked examples during practice sessions was to boost problem-solving skills, particularly because studying correct examples relieves dependence on students' working memory and allows them to spend time deeply processing the different steps in the problem solution rather than applying procedures by rote (Sweller & Cooper, 1985). However, reflecting on errors by studying and explaining incorrect examples is expected to have an additional benefit to procedural knowledge in that it should extinguish students' ineffective strategies and prepare them to seek out more effective strategies (e.g., Siegler, 2002).

In addition, explanation of incorrect worked examples is thought to support conceptual knowledge in two ways. First, the process of explanation alone is known to increase learners' reflection on the connections between the present information and their previous knowledge, leading to critical refinement of their existing knowledge (Chi, 2000; Roy & Chi, 2005). Second, and perhaps more importantly, focusing this explanation on incorrect examples allows students a greater opportunity to reflect on errors, and to notice the components of the problem that make those procedures ineffective; students exposed to incorrect examples demonstrate better conceptual encoding of problem features than those exposed to solely correct examples (Booth et al., 2013). Practice sessions involving incorrect examples have been shown to lead to increases in conceptual knowledge about those problem features (Durkin & Rittle-Johnson, 2012; Booth, Oyer, Paré-Blagoev, Elliot, Barbieri, Augustine, & Koedinger, in press) and the extent to which students discuss concepts in math class (Durkin & Rittle-Johnson, 2012).

Research Question 2.

The second question posed focused on whether promoting error reflection within a worked examples intervention improves middle grade students' sense of belonging to mathematics.

As previously discussed, students' sense of belonging to a math community (Good, Rattan, & Dweck, 2012) has been found to be a potentially important factor in mathematics learning at the college level, yet little is known about middle school students' sense of belonging to math and its relation to learning. Students' perceptions of their mathematics ability and learning may be related to their sense of belonging to mathematics. Young students, particularly young adolescents, may have a greater need for a sense of belonging, given the achievement and motivational declines found in the US during adolescence for mathematics (Wang & Pomerantz, 2009) and their increased focus on peer relationships (Rodkin & Ryan, 2012; Wentzel, 2005) Failure experiences within a given domain may lower feelings of belonging in that domain and these experiences may be more common when students are just beginning to learn a more advanced form of mathematics than they are accustomed to. The error reflection intervention was expected to be particularly beneficial to students' sense of belonging to mathematics for those in the early stages of algebra learning. The error reflection intervention was expected to normalize making errors as part of the learning process and help students to identify with mathematics.

Research Question 3.

The third question was posed to address whether the influence of the error reflection intervention on the outcome variables is mediated by changes in students' perception of the functionality of errors.

Specifically, the current study investigated whether increases in students' perceived functionality of errors mediates the influence of the error reflection intervention on learning and sense of belonging to mathematics. Changes in students' beliefs about errors themselves were expected to influence the impact of the error reflection intervention on learning and belonging. Students' perceptions on how errors are evaluated in the classroom predict students' adaptive reactions to struggles in mathematics problem-solving (Steuer et al., 2013). Students' subjective experience or views of the functionality of errors were expected to change after the error reflection intervention. Promoting error reflection in the classroom may have presented students with the idea that errors are a functional tool for learning. Increases in the usefulness of errors as a learning tool were expected to mediate the influence of the error reflection intervention on both learning and sense of belonging to mathematics. It was expected that reflection of errors in the form of either those displayed within incorrect worked examples or those hypothetical errors considered in combination with correct worked examples would promote the idea that errors are indeed an important part of the learning process. This positive "mastery-oriented" change in students' views on errors may have promoted more adaptive and effortful learning strategies on behalf of the study. This could lead to the refinement of conceptual and procedural knowledge. Increases in students' views on the usefulness of errors were also expected to promote greater feelings of belonging to mathematics through normalization of errors or the idea that the math community is a more inclusive community in which the student might actually take part. If students perceive errors as functional, this may provide an added benefit of the error reflection intervention on sense of belonging to mathematics.

To rule out alternative hypotheses, the current study also measured students' entity view of math ability, math self-concept, perceptions of math importance or values, and interest in math. This was done to provide evidence of the unique contributions of perceived functionality of errors to post-intervention measures. Entity view of math ability is conceptualized and measured based on Dweck's longstanding research (Dweck, 1999; Rattan, Good, & Dweck, 2012). Math self-concept, perceptions of math importance, and interest in math were conceptualized and measured under the theoretical framework of Eccles' and Wigfield's well-validated Expectancy-Value model of achievement motivation (Wigfield & Eccles, 2000). Each of these constructs and the rationale for their inclusion are briefly summarized below.

One's theory of intelligence is defined as one's belief that intelligence is either a malleable or stable trait (e.g., Dweck & Leggett, 1988, Dweck, 1999). Students who believe that intelligence is a fixed trait, or something that cannot be changed much, are referred to as *entity theorists*. As previously explained, changes in students' theory of intelligence due to a stereotype threat manipulation have been found to predict changes in sense of belonging to math (Good et al., 2012). The current study did not assess stereotype threat. However, a potential alternative hypothesis that could arise may be that the new construct of perceived functionality of errors may be synonymous with, or very similar to, one's theory of intelligence. Dweck suggests that students' persistence in the face of difficulty is predicted by their theory of intelligence (Blackwell, Trzesniewski, & Dweck, 2007). One might suggest that an error reflection intervention may result in changes in one's entity view, rather than one's perceived functionality of errors, which in turn may predict increases in sense of belonging to math or learning. Similarly, one might even suggest that a students' perceived functionality of errors for learning is more dependent upon their theory of intelligence than either their sense of belonging to math or any experimental manipulations with errors that we choose to implement. However, in the current author's prior research experience with middle school Algebra I students, there has been very little variability in students' theories of intelligence, no relation to learning or other motivation constructs within similar worked examples interventions, and no change from pre- to post-worked example interventions. Still, its inclusion would allow us to rule out these and other alternative hypotheses or explanations related to students' theory of intelligence.

Another alternative explanation that might arise is that students' perceived functionality of errors is another representation of their ability beliefs, values, or even interest in math, and that learning through error reflection may be better predicted by students' ability beliefs, values for, or interest in math, rather than changes in perceived functionality of errors. To rule this out, math self-concept, perceptions of the importance of math, and interest in math will also be measured. These will be explored in the event that the relatively new and understudied construct of perceived functionality of errors seems to predict learning.

Eccles' and colleagues' conceptualization of math self-concept includes both expectancies for success and ability beliefs (Wigfield & Eccles, 2000; Wigfield & Eccles, 2002). Expectancies are students' beliefs on how well they will do on an upcoming task. Ability beliefs are students' self-evaluations of how competent they are in a particular area. Eccles' and colleagues' conceptualization of math self-concept was chosen for use in the current study as a measure of ability beliefs as it is a domain-general measure of students' beliefs about their own ability in math and how they compare to others. It has been found to predict both achievement and motivation in a variety of subjects during adolescence. As the current study is concerned more generally with factors that may or may not predict learning and motivation in the domain of math more generally, rather than on specific math tasks assigned in class, and math is presumably a domain in which comparison to others is prevalent, this particular conceptualization of *math* self-concept seemed more appropriate than other more task-specific perceptions of ability such as self-efficacy (Bandura, 1986; 1997) or more general perceptions of ability such as *general* self-concept or *academic* self-concept (Byrne, 1996). Prior related work has found that students' ability beliefs as measured by a related construct, competence expectancy (Elliot & Church, 1997), predicted higher post-test scores overall but did not interact with condition based on worked example type (i.e., correct versus incorrect worked examples; Barbieri & Booth, under review). So while it was expected that perceived functionality of errors would be a more suitable construct for predicting changes in motivation and learning based on the error reflection interaction, math self-concept may also account for changes in this construct. Thus, it was measured as a potential explanatory variable should this new construct of perceived functionality of errors did indeed demonstrate mediating effects of condition.

Several different types of values make up what is referred to as subjective task values under the Expectancy-Value model. Throughout Eccles' and colleagues earlier research, the main components of subjective task values include *attainment value* or importance, *utility value*, and *intrinsic value* (Eccles et al., 1983; Wigfield & Eccles, 1992). Attainment value is defined as how important one deems doing well on a particular task, or in this case, in math class (e.g., “For me being good at math is . . . .”). Utility value is defined as how useful one deems a particular task to be either for a particular goal or, in the current study, more generally (e.g., “In general, how useful is what you learn in math?”). Intrinsic value is defined as how enjoyable or interesting one judges a particular task to be (e.g., “How much do you like doing math?”). More recently, Eccles and colleagues have treated intrinsic value for math as a measure of *interest in math*, separate from attainment and utility value for math, which they term as *perceptions of the importance of math* (Simpkins, et al, 2013). Thus, the current study adopts this recent modification and terminology and considers both utility and attainment value as a measure of perceptions of importance of math. The current study refers to intrinsic value as interest in math. Our prior work has not found predictive power in measures of students’ interest in math on learning or motivation overall or by condition. As such, students’ interest in math is not expected to mediate the effect of error reflection on learning or post-intervention motivation. However, perceptions of math importance have not been previously assessed in our work. This, interest, and math self-concept were all measured to allow testing of alternative hypothesis testing in relation to perceptions of the functionality of errors, if deemed necessary.

Research Question 4.

The fourth and final research question was posed to determine whether the relationships between the error reflection intervention and students' learning and sense of belonging to math differ by URM status. If so, the author sought to determine whether prior knowledge and prior sense of belonging can explain these differences.

As previously explained, Barbieri and Booth (under review) found that a worked examples intervention which included promoting error reflection through the use of incorrect examples led to increases in sense of belonging to mathematics specifically for URM students. The causal mechanism is unknown. This effect could not be explained by differences in prior knowledge, competence expectancy, or SES. However, the sample had very few students receiving free or reduced lunch (~5%) and included a relatively small group of URM students (~ 25%). Thus, the current study aimed to replicate and further explore this finding in a district with a larger proportion of minority students. During interpretation however, it would be important to consider other contextual features of the district and classrooms completing the study that may influence the results. Some of these factors are proportion of minority students in each classroom, tracking methods used to assign students to specific math classes, and differences in SES that may or may not coincide with URM status in the participating district. Unfortunately, as previously mentioned, the analysis sample had fewer than promised URM students, which complicated analyses. This is discussed in Chapters 4 and 5.

The current study aimed to address whether the relationship between the error reflection intervention and student learning and sense of belonging to math differed by URM status as previous findings suggest. If similar differences were found in the current study, consideration was to be given to whether these differences could be explained by differences in prior knowledge or prior sense of belonging to mathematics. For example, students' sense of belonging to mathematics may be influenced by their representation status within mathematics. Representation status may influence the manner in which underrepresented students perceive errors and respond to failure experiences. Due to the history of lower performance of URM students combined with their underrepresentation status, URM students were expected to have a lower sense of belonging to mathematics than non-URM students. Increases in *actual* competence and exposure to reflection on errors may be particularly salient to students with a lowered sense of belonging. The error reflection intervention was expected to promote the idea of the mathematics community as a more inclusive community. This may have been especially useful for students who feel underrepresented in a particular domain, as URM students are in math. However, if the expected moderating effect of URM status on the influence of the error reflection intervention on belonging could be accounted for by prior knowledge, this could have been interpreted in a number of ways. This could include the possibility that URM students often come from low-income backgrounds or experience lower quality instruction which leads to discrepancies in prior knowledge between students. The following chapter describes the methodology that was used to address the research questions above.

CHAPTER 3

METHODOLOGY

Overview

The following chapter begins with a review of the research design. This is followed by a description of the sample and sampling procedures as well as materials used for the study. Data collection procedures are then detailed followed by a plan of analysis. The methodology was designed to assess the relationship between reflection of errors, perceived functionality of errors in the classroom, and students' sense of belonging to math. Specifically, the study addresses four key questions: 1) whether promoting error reflection through the use of incorrect worked examples and self-explanation prompts fosters middle grade students' algebra learning; 2) whether the error reflection intervention leads to increases in students' sense of belonging to mathematics; 3) whether increases in students' perceived functionality of errors for learning mediate the influence of the error reflection intervention on learning and sense of belonging to mathematics; and 4) whether the influence of the error reflection intervention on learning and sense of belonging differs by URM status was also to be explored. It was initially proposed that if differences arose, consideration would be given to whether these differences can be explained by differences in prior knowledge or prior sense of belonging to mathematics.

Research Design

A field experiment was employed to determine the influence of error reflection interventions, in the form of worked examples and self-explanation prompts, on students' conceptual and procedural knowledge in algebra as well as their sense of belonging to mathematics. Also of interest was whether the impact of error reflection interventions is mediated by students' perceived functionality of errors and whether any of these relationships differ by URM status. Several measures were administered before and after the intervention. Sense of belonging to mathematics as well as several other motivation constructs were measured pre- and post-intervention in Survey A. A pre-, post-, and delayed post-test were also planned to measure students' conceptual and procedural knowledge in solving quadratic equations. Students' perceived functionality of errors for learning was also measured pre- and post-intervention in Survey B; Surveys A and B were not measured at delayed post-test due to the relatively short timeframe in which the current study was to be completed. This decision was made based on prior experience with middle-school participants who expressed frustration when administered the same surveys more than twice during similar short unit studies. This often led to incomplete surveys.

Table 3.1
Data collection timeline

Time	Measure	Constructs measured
1	Survey A: T1	Students' sense of belonging to math, math self-concept, perceptions of importance/value, interest in math
2	Pre-test	Conceptual and Procedural knowledge for solving quadratic equations
3	Survey B: T1	Students' perceived functionality of errors for learning, entity view of math ability
4	Intervention	(4 worksheets)
5	Survey B: T2	Students' perceived functionality of errors for learning, entity view of math ability
6	Survey A: T2	Sense of belonging to math, math self-concept, perceptions of importance/value, interest in math
7	Post-test	Conceptual and procedural knowledge for solving quadratic equations
8	Delayed Post-test	Conceptual and procedural knowledge for solving quadratic equations

Students were randomly assigned to a condition within classrooms which was done to control for teacher effects and allow for causal inferences. The intervention itself consists of four worksheets completed over the course of several weeks that varied on a number of features, mainly on whether students are asked to reflect on an error and how they are asked to do so. Ecological validity is expected to be high due to the classroom setting and the administration of the error reflection intervention worksheets by the regular classroom teacher. Table 3.1 displays the timeline of the current study and timing of data collection.

Specific times are not detailed in this table because each participant completed the study at their own pace. For example, even if the classroom teacher administered the worksheets on certain days, if a student did not finish a particular worksheet within the class period that it was assigned, they were given time to complete it in class at another time before moving on to the next measure. However, the ordering of data collection was the same across all participants. Eight classes completed the study up to and including the post-test in the time allotted for the study (approximately two months). However, these classes did not complete the delayed post-test during the allotted time. Also, three classes did not finish up to the post-test within the allotted time, and this is reflected in the data analyses. Discussion with the coordinating mathematics teacher revealed that although each participating teacher planned to finish the study in the agreed upon time, this was not possible for some due to unforeseen circumstances such as unclear testing schedules, class trips, surprise assemblies, and school closings and delays due to inclement weather. This is further detailed in Chapter 4.

Population and Sample

The population for this study includes middle school algebra classrooms in a midwestern public school district. This school district was selected due to their willingness to participate after a previous partnership. The participating school district is part of the Minority Student Achievement Network (MSAN) which is a partnership between school districts and researchers committed to eliminating achievement gaps between URM and non-URM students through collaborative research, professional development, program evaluation, and other means. The participating district has a marked achievement gap between Black and White students. The sample for the study included 11 classrooms. According to the US Census 2011-2013 American Community Survey (United States Census Bureau/American FactFinder), the population of town that the participating school is in is 34.9% Black or African American and 56.1% White. The rest of the population consists of other or mixed races. Of the population, 2.5% is of Hispanic or Latino origin. Of particular interest was whether the relationship between the error reflection intervention, sense of belonging, and learning differs by students' URM status. This would require an oversampling of URM students. In the domain of mathematics, Black and Hispanic students are particularly underrepresented. The sample for the current study includes Algebra I students from a midwestern school district. According to the school district website (www.shaker.org), African-American students made up 50% of the student body in the 2011-2012 academic year. This is one of the reasons that this district was chosen. Unfortunately, after data collection, it was realized that this percentage was not represented in the Algebra I classrooms. This complicated planned analyses and is discussed in more detail in Chapters 4 and 5.

Participants included middle school algebra students (*N* = 207; 104 Female) from eleven Algebra I classrooms of four different teachers. Eight of the eleven classes in the sample were considered accelerated Algebra I classrooms. Of those classes that completed the post-test, six of them were accelerated and two were average Algebra I classes. The overall sample included White (62.8%), Black (23.2%), Asian (6.8%), multi-racial (5.8%), and Hispanic (1.4%) students, as categorized by their school district. As the achievement gap in the US and in MSAN districts in particular is most notable between White and non-Asian minority students, and with consideration of the restricted sample size, White and Asian students were classified as non-URM students (69.6%) and all other students were classified as URM students (30.4%). Students receiving free or reduced lunch, a common proxy of socio-economic status, made up a small amount of the sample (14%). Most students were in the eighth grade, although specific grade levels for participants were not made available to the present author. Students were assigned within classroom to one of four conditions: Problem-Solving control (*n* = 51), Correct Example control (*n* = 51), Correct Example Error Reflection condition (*n* = 53), and Incorrect Example Error Reflection condition (*n* = 52). These conditions will be detailed later on in the chapter. Comparisons of conditions at pre-test as well as missing data analyses are presented in the beginning of Chapter 4.

Materials

Students' conceptual and procedural knowledge of solving quadratic equations was measured pre- and post-intervention in the form of a pencil-and-paper test (Appendix A). Two self-report survey measures were utilized to measure a variety of motivation constructs and students' perceived functionality of errors for learning. Survey A

(Appendix B) measured students' sense of belonging in mathematics, math self-concept, expectancies and values, and interest in math. Survey B (Appendix C) measures students' perception of the functionality of errors in the classroom and entity view of math ability. Though the constructs of interest within the surveys are sense of belonging to mathematics and perceived functionality of errors, other constructs are measured to rule out potential alternative explanations and provide evidence of the unique contributions of sense of belonging to mathematics and perceived functionality of errors to post-intervention measures. The intervention includes the administration of four worksheets on the topics of solving quadratic equations. These worksheets are displayed in Appendices D through G.

The current study has two control conditions and two experimental conditions. One control condition is a problem-solving only condition. Another control condition is a worked examples condition with only correctly solved worked examples and written self-explanation prompts paired with similar practice problems. The two experimental conditions include two worked examples conditions that vary in how they promote error reflection. One experimental condition has students reflect on errors displayed in incorrectly-solved worked examples and answer written self-explanation prompts targeting the errors displayed. Another experimental condition hasstudents reflect on hypothetical errors through the use of written self-explanation prompts as they study correctly-solved worked examples. Each worked example in the three worked examples conditions are paired with similar practice problems. These worksheets will be further detailed below and are displayed in Appendices D through G. The following sections explain each of the measures listed above.

Procedural and conceptual knowledge.

The pre- and post-tests assess students' ability to solve quadratic equations using the methods demonstrated in the worksheets (i.e., the quadratic formula, factoring, square root, graphing). Students' conceptual and procedural scores are the percent of each problem type answered correctly. The assessment includes 23 conceptual items and 16 procedural items which are displayed and labeled in Appendix A. Conceptual and procedural knowledge have been defined in a variety of ways by researchers in the domain of mathematics. The current study adopts the definition of conceptual knowledge as an understanding of the meaning of features in a problem and procedural knowledge as knowing how to solve a problem (Booth, 2011). To be successful in mathematics, one must be able to take the appropriate steps to solve a problem as well as have an understanding of the meaning of features in that particular problem. Because the development of conceptual and procedural knowledge in the domain of mathematics is an iterative process, these two skills are often thought of as being on a continuum rather being discrete (Star, 2005; Rittle-Johnson, Siegler, & Alibali, 2001). This makes measurement of these skills difficult. A description of their differentiation in the present study follows.

Conceptual knowledge is demonstrated by a students' understanding of features in a problem. For example, Question 3 on the pre-post unit test asks, "*State whether each of the following is true for the quadratic function* $y = -x^2 - 2x + 3$". One of the statements to be responded to as "Yes" or "No" is, "g. *The vertex is a maximum*". To respond to this statement, a student must presumably have an understanding of the meaning of a particular feature of the problem; the coefficient of x^2, or *a*. This feature determines whether the vertex is a maximum or minimum value on the graph of a quadratic function. The student must also understand that if a < 0, the graph for the quadratic function will be a downward-facing parabola and the vertex will be a maximum value on the graph. Thus, in this particular problem, the coefficient of x^2, or a, is -1 which is less than 0, which means that the vertex will be a maximum. So the correct answer for this conceptual problem will be "Yes", the vertex is a maximum. As is evident, there are very few steps one could take to solve this problem. Rather, understanding the feature a is presumably what enables a student to correctly respond to this item on the measure.

In contrast, procedural knowledge will be demonstrated by knowing how to solve a problem. For example, Question 4 on the pre-post unit test asks students to, "Solve the quadratic equation by factoring. Show all of your work. $0 = x^2 - 6x - 27$." This problem would be considered mainly procedural because students need to be able to take the appropriate steps to factor this quadratic function into $(x - 9)(x + 3)$. Although a conceptual understanding of the features of the problem could assist students in solving this problem, they could arguably not fully understand all features in the problem and still solve the problem correctly based on taking appropriate steps. All test items can be found in Appendix A. Internal consistency for procedural scores at pre- and post-test were sufficient (α = .766; α = .820). Internal consistency for conceptual scores at pre- and post-test was also sufficient (α = .696; α = .820).

Sense of belonging to mathematics.

Students' sense of belonging to mathematics was measured using five statements for which the student provided his or her level of agreement on a Likert scale, on their sense of belonging in a math community *when they are in math class*. The five statements were chosen based on a factor analysis of pilot study data which supported the five subscales of Membership, Acceptance, Affect, Trust, and Desire to Fade that make up a sense of belonging as proposed by Good, Rattan, and Dweck (2012). Items with the highest loadings in each of the factors were be used: "I feel like I am part of the math community" (Membership), "I feel accepted" (Acceptance), "I feel comfortable" (Ease), "I trust my teachers to help me learn" (Trust), and "I enjoy participating" (Desire to Fade). These items are included in Survey A and displayed in Appendix B. This survey was administered at the start of the study and then after the intervention prior to the post-test. Items were reworded to meet the reading level of middle school students, as the original scale was used with undergraduates. The scale ranges from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*) as it is standardly measured in the extant literature. Participants' scores were computed by averaging the items. Higher means represent a stronger sense of belonging in math. Internal consistency for sense of belonging to math pre- and post-intervention was sufficient ($\alpha = .879$; $\alpha = .849$).

Sense of belonging to math was measured along with several other motivation constructs in Survey A which include students' math self-concept, perceptions of math importance or values, and interest in math. Entity view of math ability was measured in Survey B along with perceived functionality of errors. Each of these was measured to enable ruling out of several alternative hypotheses as detailed in Chapter 2, if deemed necessary. Entity view of math ability was conceptualized and measured based on Dweck's longstanding research (Dweck, 1999; Rattan, et al., 2012). Math self-concept, perceptions of math importance, and interest in math were conceptualized and measured under the theoretical framework of Eccles' and Wigfield's well-validated Expectancy-Value model of achievement motivation (Wigfield & Eccles, 2000). Measurement of these constructs is detailed below.

Entity view of math ability.

One's theory of intelligence is defined as one's belief that intelligence is either a malleable or stable trait (e.g., Dweck & Leggett, 1988, Dweck, 1999). Students who believe that intelligence is a fixed trait, or something that cannot be changed much, are referred to as *entity theorists*. Students' entity views were measured using their level of agreement using a Likert scale with four statements within Survey B which were selected from Dweck's (1999) Implicit Theories of Intelligence Scale. These statements include, "You have a certain amount of math ability, and you can't really do much to change it", "Your math ability is something about you that you can't change very much", "To be honest, you can't really change your ability in math", and "You can learn new things, but you can't really change your basic math ability." The Likert scale ranges from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). Participants' scores were computed by averaging the items. Higher means represent a stronger entity view of math ability. These items are included in Survey B and displayed in Appendix C. This survey was administered at the start of the study and then again prior to the post-test.

Math self-concept.

Eccles' and colleagues' math self-concept was chosen for use in the current study. This measure of math self-concept includes both expectancies for success and ability beliefs (Wigfield & Eccles, 2000; Wigfield & Eccles, 2002). Expectancies are simply students' beliefs on how well they will do on an upcoming task, and in this case, in a particular domain (i.e., math class). Ability beliefs are students' self-evaluations of how competent they are in a particular area, and in this case, math. Eccles' and colleagues' conceptualization of math self-concept was chosen for use in the current study as it is a domain-general measure of students' beliefs about their own ability in math and how they compare to others. As the current study is concerned more generally on factors that may or may not predict learning and motivation in the domain of math more generally, rather than on specific math tasks assigned in class, and math is presumably a domain in which comparison to others is prevalent, this particular conceptualization of self-concept seemed most appropriate.

Students' math self-concept was measured with five statements for which students used a 7-point Likert scale to respond. These statements were derived from well-validated scales (Anderman, et al., 2001; Eccles, et al., 1993; Jacobs, et al., 2002). Although early work in development of the scales assumed that theoretical differences between competences beliefs and expectancies should result in distinct scales (Eccles et al., 1983), later work by Eccles and colleagues demonstrated that these measures load on the same factor (Eccles & Wigfield, 1995; Eccles et al., 1993). Thus, work under this model includes both expectancies and ability beliefs to measure the overarching construct of math self-concept. The original items validated by Eccles and colleagues (Eccles, et al., 1993) have been slightly modified and validated for use with middle and high school math students (Simpkins, Davis-Kean, & Eccles, 2013). Each of the statements has different options for values 1 and 7, but 1 generally represents low math self-concept and 7 generally represents high math self-concept. The statements include, "How good at math are you?" (1 – *Not at all good*, 7 – *Very good*), "If you were to list all of the students in your class from worst to best in math, where would you put yourself?" (1 – *One of the worst*, 7 – *One of the best*), "Compared to most of your other school subjects, how good are you at math?" (1- *A lot worse,* 7 – *A lot better*), "How well do you expect to do in math this year?" (1 - *Not at all well,* 7 – *Very well*), and "How good would you be at learning something new in math?" (1 – Not at all good, 7 – Very good). These items are included in Survey A and displayed in Appendix B. This survey was administered at the start of the study and then again prior to the post-test.

Perceptions of the importance of math.

Several different types of values make up what is referred to as subjective task values under the Expectancy-Value model. Throughout Eccles' and colleagues earlier research, the main components of subjective task values include *attainment value* or importance, *utility value*, and *intrinsic value* (Eccles et al., 1983; Wigfield & Eccles, 1992). Attainment value is defined as how important one deems doing well on a particular task, or in this case, in math class (e.g., "For me being good at math is . . . ."). Utility value is defined as how useful one deems a particular task to be either for a particular goal or, in the current study, more generally (e.g., "In general, how useful is what you learn in math?"). Intrinsic value is defined as how enjoyable or interesting one judges a particular task to be (e.g., "How much do you like doing math?"). More recently, Eccles' and colleagues has treated intrinsic value for math as a measure of *interest in math*, separate from attainment and utility value for math, which they term as *perceptions of the importance of math* (Simpkins, et al, 2013). Thus, the current study adopts this recent modification and terminology and considers utility and attainment value items as a measure of perceptions of importance of math and intrinsic value scale items as a measure of interest.

Students' perceptions of the importance of math were measured with three statements for which students used a 7-point Likert scale to respond. These statements were derived from well-validated scales (Anderman, et al., 2001; Eccles, et al., 1993; Jacobs, et al., 2002) and recently modified for use with middle and high school math students (Simpkins, et al., 2013). Each of the statements has different options for values 1 and 7, but 1 generally represents low importance and 7 generally represents high importance. The statements include, "In general, how useful is what you learn in math?" (1 – Not at all useful, 7 – Very useful), "For me being good at math is . . . ." (1 – Not at all important, 7 – Very important), and, "Compared to most of your other activities, how important is it to you to be good at math?" (1 – Not as important, 7 – A lot more important). These items are included in Survey A and displayed in Appendix B. This survey was administered at the start of the study and then again after completion of the intervention and prior to the post-test.

Interest in math.

Students' interest in math can be defined as the intrinsic value or enjoyment that one experiences when partaking in a math task (Simpkins, et al., 2013). Students' interest in math was measured with three statements for which students must use a 7-point Likert scale to respond. These statements were derived from well-validated scales (Anderman, et al., 2001; Eccles, et al., 1993; Jacobs, et al., 2002) and recently modified for use with middle and high school math students (Simpkins, et al., 2013). Each of the statements has different options for values 1 and 7, but 1 generally represents low interest in math and 7 generally represents high interest in math. The statements include, “In general, I find working on math assignments . . . .” (1 – Very boring, 7 – Very interesting), “How much do you like doing math?” (1 – A little, 7 – A lot), and, “Compared to most of your other activities, how much do you like math?” (1 – Not as much, 7 – A lot more). These items are included in Survey A and displayed in Appendix B. This survey was administered at the start of the study and then again after the completion of the intervention but prior to the post-test.

Perceived functionality of errors.

Students’ perceptions of the functionality of errors for learning in the classroom were measured using a subscale of a previously validated perceived error climate scale (Steuer et al., 2013). While the overarching perceived error climate scale measures a student’s perception of how their classroom environment supports or restrains learning from errors, the subdimension of perceived functionality of errors for learning measures students’ perceptions of how errors made in the classroom may be used as starting points for learning in the classroom. Functionality of errors for learning was measured with four statements for which the student provided his or her level of agreement on a Likert scale. The items include, “In our Math class the mistakes students make are often used to make sure you really understand Math”, “In our Math class we learn a lot from assignments that were not done correctly”, “In our Math class wrong answers on assignments are used to learn something”, and “In our Math class wrong answers are often a good opportunity to really understand the material”. The scale ranges from 1 (*Strongly disagree*) to 5 (*Strongly Agree*). Participants’ scores were computed by averaging the items. Higher means represent more positive perceptions of the functionality of errors for learning.

These items are included in Survey B and displayed in Appendix C. Survey B was administered immediately before administration of the first worksheet and immediately after administration of the fourth worksheet within the intervention.

Worksheets.

The experimental manipulation was within the four worksheets that the students completed. It is important to note that the four worksheets used are not simple or brief practice problems. Rather, they were designed to be cognitively demanding and require both time and effort on behalf of each student. In prior work, teachers using similar worksheets have reported their students working on them for 25 to 30 minutes at a time and have normally taken between four and six weeks to complete the worksheets along with the specified unit. In the current study, completion of the worksheets seemed to take a similar amount of time. Most importantly, four worked examples worksheets have been found to be effective on both learning and changes in motivation (e.g., Barbieri & Booth, under review). Worksheets for each condition are displayed in Appendices D through G.

To test the effect of an error reflection intervention on learning and changes in sense of belonging to mathematics, there were two experimental worked examples conditions that promote error reflection and two control groups, one of which is a problem-solving condition (i.e., *Problem-Solving Control* condition) and another that is a worked examples control condition. The three worked examples conditions varied in the correctness of the procedure displayed as well as whether students are prompted to reflect on correct procedures or on errors. Each of the worked examples in the three worked examples conditions were paired with a similar problem to be completed after providing written responses to the self-explanation prompt displayed for the matching example.

Rather than utilizing aural self-explanation prompts as was done in early laboratory research on self-explanation (Chi, 2000; Rittle-Johnson, 2006), written self-explanation prompts were used as a more appropriate logistical activity to be completed individually in a classroom setting. This form of self-explanation has been found effective in classroom studies in combination with worked examples (Booth, et al., in press).

Each of the four worked examples worksheets, regardless of condition type, includes three example-problem pairs. This results in each student in the worked examples groups studying a total of 12 worked examples and completing a total of 12 practice problems. The problem-solving control group completes practice problems identical to the problems in the worked examples conditions but without the worked examples or self-explanation prompts. To compensate for loss of time on task differences with the removal of the worked examples or self-explanation prompts, students in the problem-solving control group completed a total of 24 practice problems instead of 12. These features are displayed in Table 3.2.

The worked examples control group was a correct worked examples condition which included self-explanation prompts that focus students' attention solely on correct concepts and procedures. This control is referred to as the *Correct Example Control* condition. The first of the experimental conditions prompted students to reflect on errors through the use of incorrect worked examples and self-explanation prompts that targeted errors displayed in the worked examples. This condition is referred to as the *Incorrect Example Error Reflection* condition. Another experimental condition prompted students to reflect on errors through the use of self-explanation prompts by asking students to consider if an error had been made in the correctly solved worked example with which it is paired. This condition will be referred to as the *Correct Example Error Reflection* condition. These conditions varied in a way that enabled the determination of whether students benefit from promoting error reflection in general just by prompting them to consider a hypothetical error or whether reflecting on an error displayed within the worked example is necessary for students to benefit from error reflection.

The self-explanation prompts in these two error reflection conditions were designed to direct attention to common errors made by algebra students in the topic of solving quadratic equations. The feature specified in the self-explanation prompts was the same for each worked examples condition regardless of the correctness of the example or whether the student was asked to reflect on an error or a correct procedure.

The four worksheets focus on solving quadratic equations by using (1) the quadratic formula, (2) factoring, (3) the square root, and (4) by graphing. Each worksheet had three sets of example-problem or problem-problem pairs. In total, the unit included 12 example-problem pairs for the worked examples condition in total, or 12 problem-problem pairs for the problem-solving control group.

Data Collection Procedures

At the start of the study, Day 1, students completed Survey A.1, which measured students' sense of belonging to mathematics and a variety of other motivational constructs that were detailed above (see Table 3.1). After completing Survey A.1, students took a pre-test measuring procedural and conceptual knowledge of solving quadratic equations. After completion of the pre-test but before starting the intervention, students completed Survey B.1, which measured students' perceived functionality of errors for learning and another motivation construct. Students began the intervention after

Table 3.2.
Design of conditions

Condition	Practice problems[3]	Worked examples	Self-explanation prompts	Correct worked examples	Incorrect worked examples	Error reflection
	Worksheet features					
Problem-Solving Control Group	X					
Correct Example Control Group	X	X	X	X		
Error Reflect[1] Self-Explanation Group	X	X	X	X		X
Error Reflect[2] Incorrect Example Group	X	X	X		X	X

[1] Error reflection promoted in self-explanation prompts only
[2] Error reflection promoted in self-explanation prompts plus errors displayed in example
[3] Worked example conditions have 12 practice problems; problem-solving control has 24 to control for time-on-task

completion of Survey B.1.

Students were randomly assigned within classroom to one of the four worksheet conditions (i.e., *Problem-Solving Control*, *Correct Example Control*, *Incorrect Example Error Reflection*, or *Correct Example Error Reflection*). Worksheets focused on solving quadratic equations through the use of the quadratic formula, through factoring, through the use of the square root, and by graphing. Students began the study with general knowledge of graphing. Therefore, participating teachers decided to start all students off by first working on the solving quadratics through graphing worksheet. The order of completion for the remaining three worksheets varied by teacher. Order effects were not of concerned as any effects of order would be distributed across conditions.

In the *Problem-Solving Control* group, students solved 12 pairs of similar practice problems. In the worked examples conditions, students studied 12 worked-example problem pairs that included a written self-explanation prompt. In the *Correct Example Control* group, students studied clearly marked correctly solved worked examples and provided written responses to self-explanation prompts focusing on the correct procedures. They also completed a similar practice problem paired with each worked example. In the *Incorrect Example Error Reflection* condition, students studied clearly marked incorrectly solved worked examples and provided written responses to self-explanation prompts that asked students to reflect on the incorrect procedures. They also completed a similar practice problem paired with each worked example. In the *Correct Example Error Reflection* condition, students studied clearly marked correctly solved worked examples and provided written responses to self-explanation prompts asking students to reflect on hypothetical errors that could be made within that particular type of problem. They also completed a similar practice problem paired with each worked example.

Students completed the worksheets independently at a pace that their classroom teacher deemed acceptable according to her or his normal classroom practices. Teachers were advised to use the worksheets in a way that they would normally use worksheets in their classroom, which would likely include practice after instruction on solving quadratic equations. This could have also included discussion of the problems seen in the worksheets after students have completed them. Teachers were instructed to not discuss differences between conditions but whole-class discussions of the problems themselves in the worksheets were permitted. Teachers took special care to keep students separate when they worked on the worksheets. Students were aware that they had different worksheets from each other but did not seem to be aware of how they differed exactly. Class discussions were acceptable in order to maintain ecological validity but avoiding discussion of specific differences in worksheets by condition, such as worked examples or errors, enables the maintenance of fidelity of the experimental manipulation. Although some classrooms and even students took longer than other classrooms to complete the unit and discussions potentially varied by classroom, randomization within classroom was expected to assist in avoiding potential confounds of differences in time-on-task or instruction by condition.

After completing all four worksheets in the unit, students took survey B.2 which measured students' post-intervention perceived functionality of errors for learning and another motivation construct (see Appendix C). After completing Survey B.2, students completed Survey A.2, which measured students' post-intervention sense of belonging to math and other motivation constructs (see Appendix B). Students completed an immediate post-test after completing the intervention and both post-intervention surveys. This post-test was identical to the pre-test. A delayed post-test was scheduled to be administered two weeks after the post-test date to measure retention.

Methodological Limitations

The previously described methodology allows the proposed research questions to be answered. However, a major limitation of the current study is the relatively small sample size. Eleven classrooms within one school district agreed to participate in the current study. So although students were randomly assigned within classroom, the study sample is a convenience sample, not a true random sample. This small sample size limits the complexity of the analyses that can be conducted. Further, as will be discussed in Chapter 4, the number of URM students actually present in the analysis sample as well as the overall sample was unexpectedly quite minimal which makes answering research question 4 impossible. Implications are discussed in Chapter 5.

Another limitation of the current methodology is the use of self-report survey measures using Likert scales for sense of belonging to mathematics and perceived functionality of errors. Likert-scale measures allow for certain biases in reporting such as demand characteristics or certain response types. However, as the data collection was anonymous with no incentive for participation, participants may have been less likely to provide self-reports unrepresentative of their actual beliefs. As with any field experiment, there was a potential for diffusion of treatment. However, teachers were asked to ensure that students were unaware of the differences between conditions for the duration of the study. Even with these limitations, the data collected in the current study should allow for causal claims to be made due to the random assignment within classrooms and the careful design of the worksheets for the four different conditions. These limitations should be considered when interpreting the findings presented in the upcoming Chapter 4. A more thorough analysis and suggestions for future directions are presented in Chapter 5.

CHAPTER 4

RESULTS

Overview

The following chapter is organized into three major parts. First, the plan of analysis is detailed, which includes information on how missingness and clustering are handled in the data from the present study. This section will also necessarily include descriptive statistics that helped to guide decisions made regarding analyses. After presentation of the plan for analysis, models predicting learning outcomes of conceptual and procedural knowledge will be detailed. This will be followed by models exploring students' sense of belonging to math. The chapter will close with a brief discussion of the limitations of the presented analyses.

Plan of Analysis

The current study was conducted to assess the influence of the error reflection intervention on students' algebra learning and sense of belonging to mathematics. Also of concern was whether students' perceived functionality of errors for learning mediates the influence of the error reflection intervention on learning and changes in sense of belonging. A series of path analyses was conducted in MPlus Version 7.3 (Muthén & Muthén, 1998 - 2014) using full information maximum likelihood (FIML) estimation to answer the research questions of the current study and are detailed below. Using FIML for all analyses allowed missingness to be handled consistently while increasing statistical power (Shafer & Graham, 2002).

Dummy coding.

Several dummy coded variables were created, including condition, URM status, and low socioeconomic status (SES). More specific analyses comparing each of the four conditions required a set of three dummy-coded variables. These are termed Incorrect Example Error Reflection, Correct Example Error Reflection, and Correct Example Control. A zero on all of these variables represents the Problem-solving Control condition. URM status is a dichotomous variable as well. As the achievement gap in mathematics performance in the US is most notable between White and non-Asian minority students, White and Asian students are classified as non-URM students and all other students are classified as URM students. Low SES is also a dichotomous variable with students receiving free or reduced lunch, a common proxy of socio-economic status, being coded as low SES.

Descriptive statistics.

Descriptive statistics for the entire sample on key observed variables prior to grand mean centering are presented in Table 4.1. Prior to running prediction models, chi square analyses and one-way ANOVAs were conducted to ensure equivalence between conditions on demographic variables, pre-intervention Sense of Belonging to Math, and pre-test Procedural and Conceptual scores. No differences were found at pre-test amongst participants across the four conditions. These initial tests along with descriptives prior to centering are presented in Table 4.2. Variables were centered to reduce the risk of multicollinearity that may arise particularly when including interaction terms in analyses.

Table 4.1. *Sample statistics (N = 207)*

Continuous Variable	Mean	*SD*	*Min*	*Max*	*n*
Pre-test Conceptual Score	45.96	17.26	.00	86.96	203
Post-test Conceptual Score	71.26	19.45	17.39	100	110
Pre-test Procedural Score	17.06	16.41	.00	83.33	203
Post-test Procedural Score	54.18	25.40	.00	100	110
Pre-intervention Sense of Belonging to Math	4.54	1.07	1.00	6.00	204
Post-intervention Sense of Belonging to Math	4.36	1.13	1.00	6.00	117
Pre-intervention Perceived Functionality of Errors	3.59	.83	1.00	5.00	197
Pre-intervention Perceived Functionality of Errors	3.68	.87	1.00	5.00	134
Categorical variables		%			*n*
Low SES		14%			207
URM		30%			207
Female		50%			207

Missingness.

FIML estimation in MPlus assumes data are missing at random (MAR). It is important to note that a number of participants did not yet complete the study at the time of analysis and the remainder of the missing data will be collected by the end of the academic year. Discussions with teachers of participating classroom revealed that although they had planned to complete the study by the date agreed upon, unforeseen circumstances led to delays in data collection. Some of these circumstances included school closings or delays due to inclement weather, class field trips, and standardized state testing. T-tests and chi-square analyses were conducted to compare students who did not yet complete post-test measures at the time of analysis to students who did complete post-test measures to determine whether the missingness was predicted by any key observed predictors. Results are displayed in Table 4.3 and reveal that the participants who did not complete the immediate posttest in the allotted time frame, termed the *attrition sample*, did not significantly differ from those that have completed the immediate posttest during that time frame. Thus, FIML was used for the analyses.

Model fit indices.

The hypothesized models will be presented below. However, it is important to note that the analyses initially planned necessitated expansion and modification until models were derived that showed acceptable fit of the covariance structure of the data collected. A series of standard model fit indices available through MPlus was used to test the fit of the covariance matrix in models conducted. These include the model chi-square, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and the Standardized Root Mean Square Residual (SRMR), which will each be briefly described below. These fit indices are used collectively to determine whether changes need to be made in the hypothesized paths. Kline (2011) provides details on fit indices that will be described below.

Model chi-square.

The model chi-square tests whether there is an exact fit between the covariance matrix in the data and the hypothesized model. It assumes multivariate normality, so data were screened prior to testing. In cases where the model chi-square is significant, this indicates that there is misfit somewhere in the model and further examination is needed. In the event that the model chi-square is not significant, then the model tested is consistent with the covariance data, though the model could be misspecified in a variety of ways.

Table 4.2. Pre-intervention comparisons by condition

	Problem-Solving Control (n = 51)		Correct Example Control (n = 51)		Correct Example Error Reflection (n = 53)		Incorrect Example Error Reflection (n = 52)		Difference test
Continuous variables	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	sig. *F*-test
Procedural score	15.60	14.29	18.13	18.52	18.89	18.40	15.64	14.15	*ns*
Conceptual score	44.26	18.14	43.91	17.62	46.89	16.11	48.66	17.20	*ns*
Sense of belonging to math	4.50	1.10	4.77	0.94	4.53	1.07	4.36	1.16	*ns*
Perceived functionality of errors	3.67	0.81	3.54	0.68	3.55	1.01	3.58	0.81	*ns*
Categorical variables	%		%		%		%		sig. χ^2
Low SES	3.9%		1.9%		4.3%		3.9%		*ns*
URM	7.7%		6.3%		8.2%		8.2%		*ns*
Female	10.6%		14.0%		12.1%		13.5%		*ns*

* $p < .05$. ** $p < .01$.

Table 4.3.

Missing data analysis comparing participants by completion on pre-intervention measures

	Participants who completed post-test measures			Attrition sample			Diff Test
Continuous Variable	Mean	*SD*	*N*	Mean	*SD*	*n*	t-test
Conceptual Score	48.15	17.26	108	43.48	17.02	95	*ns*
Procedural Score	18.98	16.73	108	14.88	15.84	95	*ns*
Sense of Belonging to Math	4.52	1.11	107	4.56	1.04	97	*ns*
Perceived Functionality of Errors	3.65	.77	105	3.51	.88	92	*ns*
Categorical variables	%		*N*	%		*n*	χ^2
Low SES	10%		110	18%		97	*ns*
URM	25%		110	36%		97	*ns*
Female	48%		110	53%		97	*ns*

* $p < .05$. ** $p < .01$.

RMSEA.

The RMSEA is a badness of fit indicator in which the value of zero would indicate the best fit. Thus, the closer the value is to zero, the better the fit. A common rule of thumb is to aim for a value of less than .05 for the statistical test of RMSEA.

CFI.

The CFI is used to compare models tested to the baseline model. This is a fit index ranging from zero to one that determines the relative improvement of the hypothesized models to baseline model. Because the CFI can have easily inflated values, it is advisable to test at a threshold of .95 before assuming good fit.

SRMR.

The SRMR is another badness of fit indicator. It measures the average correlation residuals, or the difference between the predicted and observed correlations. For acceptable model fit, this should be close to zero. Values greater than .08 are often seen as an indication of poor fit. The residual matrix should also be examined in consideration to the SRMR to determine whether there is a particular problem area in one part of the model that is influencing the average of the entire model.

Indirect effects.

Indirect effects are tested by testing a model with and without the path in question and comparing the chi square. When indirect effects are of interest in a given model, it is suggested to utilize bootstrapped standard errors (MacKinnon, Lockwood, & Williams, 2004). As mediated effects, or indirect effects, are a product term of regression coefficients, this may violate the assumption of a normal distribution. Bootstrapping is a method of resampling the data with replacement to create a large number of samples. These bootstrapped samples are used to estimate standard errors for non-normal data and when dealing with missing data.

Design effect adjusted standard errors.

Although participants were nested in classrooms, participants were randomly assigned within classroom. Thus, nesting was initially expected to be minimal. However, the intra-class correlations (ICC) were calculated on post-test data and revealed that post-test conceptual and procedural scores were dependent upon cluster or classroom that participants were enrolled in. The ICC for post-test procedural scores by classroom was .298, and it was .371 for post-test conceptual scores. The ICC for post-intervention sense of belonging to math was .201. The ICCs were used to compute the design effect. The design effect is the increase in sampling error due to complex sampling procedures such as clustering. A majority of statistical tests assume that simple random sampling was utilized during data collection. Even with random assignment within classroom, the current classroom study still has a level of clustering incorporated into its sampling design. As ICCs compare the between variance component with the within variance component in a dependent variable, high ICCs indicate score dependence of participants within clusters. Applying standard statistical formulas to complex samples can lead to the underestimation of sampling variance. As standard errors are often in the denominator of single-level statistics, this can lead to a large test statistic and a higher rate of Type I error, or the discovery of a false positive.

Although multilevel modeling (MLM) is the predominant method for analyzing clustered data, certain logistical reasons sometimes prevent the use of MLM such as lack of an adequate number of clusters. Also, as there are alternative methods for analyzing clustered data, the choice to use MLM is dependent upon the researchers' focus or study purpose. Alternative methods can be used if the researcher is not particularly interested in effects at the cluster level (e.g., contextual factors) and rather wishes to focus on factors at Level 1 while still accounting for the noise that results from the clustered design (see Huang, 2014). At the time of analysis the current study only retains eight clusters or classes, a relatively small number of clusters to enable multi-level modeling. Due to the high level of clustering but the small number of clusters within the current data set, design effect adjusted standard errors will be used.

Design effects.

Design effects compare the variance of a statistic in a complex sample with clustering to the variance of a simple random sample taking into account the average cluster size. The design effect does not influence estimates but can potentially underestimate standard errors. Adjusting standard errors by considering the design effect (DEFF) allows one to consider both the clustering as measured by the ICC (ρ) and average cluster size ($\bar{n}_j$) (Huang, 2014; Hahs-Vaughn, 2005). The formula to calculate DEFF = $1 + \rho(\bar{n}_j - 1)$. A common rule of thumb is that design effects greater than 2 require that clustering be accounted for in statistical models. To utilize design effect adjusted standard errors to account for the clustering in a dataset, researchers can use DEFF to first compute an adjusted sample size (N_{adj}). This adjusted sample size is calculated by dividing the original sample size by DEFF. The formula for $N_{adj} = \frac{N}{(1+[ICC\{n_j-1\}])}$. This adjusted sample size is then used to create the design effect adjusted standard errors. The design effect adjusted standard errors are computed by dividing the product of the original standard error and the square root of the original sample size by the square root of the adjusted sample size. This is displayed in the formula of $SE_{adj} = \frac{SE\sqrt{N}}{\sqrt{N_{adj}}}$. Adjusted standard errors are then used to calculate new test statistics. The adjusted sample size was used when determining the critical value for the t-statistics for the model coefficients. The original results prior to considering adjusted standard errors will be presented alongside the more conservative results accounting for clustering with design effect adjusted standard errors.

Predicting Post-test Procedural and Conceptual Scores

Design effect adjusted standard errors for learning.

As previously stated, due to the high ICCs for post-test procedural (ICC = .298) and conceptual (ICC = .371) scores, design effect adjusted standard errors will be utilized in calculating an adjusted sample size and computing more conservative t-statistics that account for clustering in the data. The average cluster size at post-test was 13.75 resulting in a design effect of DEFF = $= 1 + .298(13.75 - 1) = 4.80$ for post-test procedural scores and DEFF = $1 + .371(13.75 - 1) = 5.73$ for post-test conceptual scores. This indicates that the variance for post-test procedural scores in a complex sampling design is estimated to be almost five times greater than if the current study utilized simple random sampling. Similarly, the variance for post-test conceptual scores in a complex sampling design is estimated to be almost six times greater than if the current study utilized simple random sampling. DEFF can be averaged across several variables when predicting multiple outcomes simultaneously as is commonly done in structural equation modeling (Lee, Forthofer, & Lorimor, 1998). As the current study uses models that simultaneously predict post-test conceptual and procedural scores, the original sample size was divided by the averaged DEFF = 5.26. Thus, $N_{adj} = \frac{110}{(5.26)} = 20.91$.

Testing main effects of condition on conceptual and procedural knowledge.

To assess whether the error reflection interventions led to algebra learning, path models were tested utilizing full information maximum likelihood (ML) estimation within MPlus. Path analyses were conducted using MPlus Version 7.3 (Muthén & Muthén, 1998 - 2014) to test the hypothesized model in Figure 4.1. Centered post-test conceptual scores and post-test procedural scores were regressed onto the dummy code for the two error-reflection conditions and the correct control condition, controlling for pre-test conceptual scores and pre-test procedural scores. A zero on these three variables represents a student in the Problem-Solving Control condition. As previously stated, it was expected that being in one of the two error reflection conditions or in the correct example control condition would lead to higher post-test conceptual and procedural scores than being in the problem-solving control group. However, it was also expected that the two error reflection conditions would have a larger effect than the correct example control condition. Further, it was hypothesized that both pre-test conceptual and procedural scores would predict both post-test conceptual and procedural scores. The hypothesized regression model is presented in Figure 4.1.

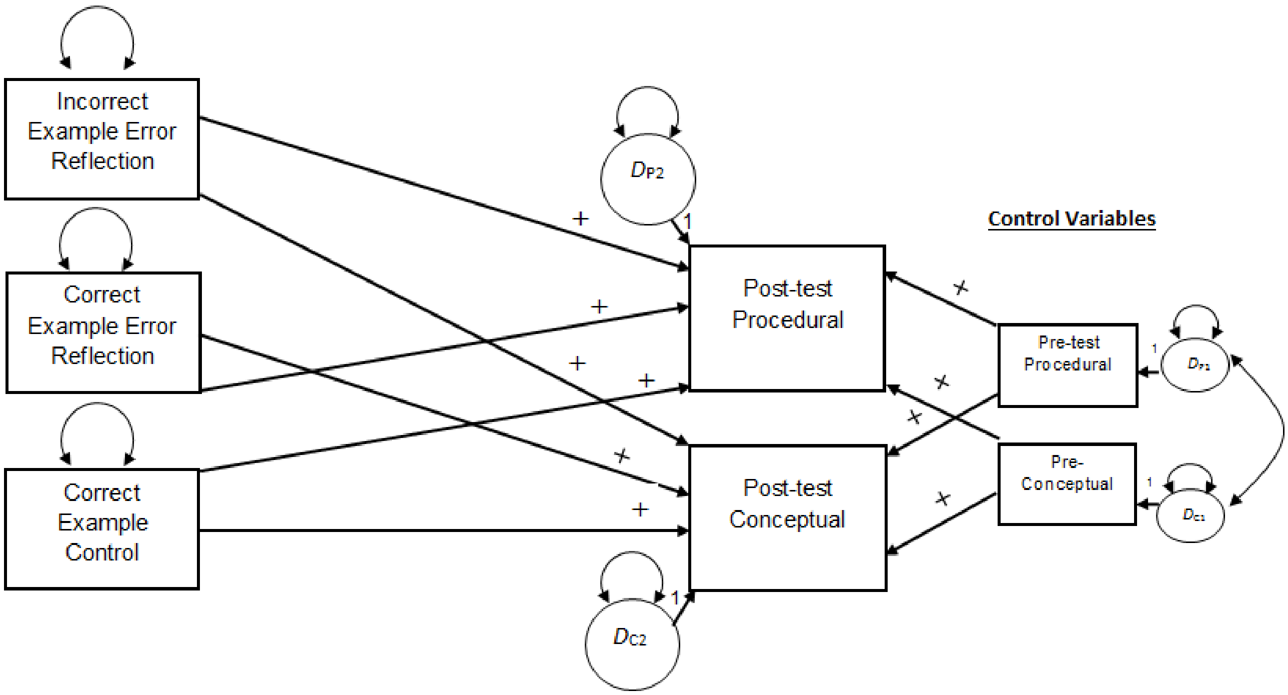


Figure 4.1. Hypothesized main effects model predicting post-test conceptual and procedural score

Preliminary analyses revealed that although pre-test conceptual scores did not predict post-test procedural scores, pre-test procedural scores did indeed predict post-test conceptual scores. Thus, a new model was fitted that included all of the hypothesized paths in Figure 4.1, with the direct path from pre-test conceptual score to post-test procedural score removed in order to improve model fit. Both subscales were controlled for at pre-test when predicting post-test procedural scores but only pre-test conceptual scores were controlled for when predicting post-test conceptual scores.

Table 4.4.

Model fit indices are presented in Table 4.4. Standardized and unstandardized coefficients before and after applying design effect adjusted standard errors are presented in Table 4.5.

Fit indices for hypothesized main effects model predicting learning

Index	Main Effects Model (Fig. 4.2)
χ^2_M	.777
df_M	1
P	.378
RMSEA (90% CI)	.000 (.000 - .243)
^{p}close-fitH_0	.439
CFI	1.00
SRMR	.014
AIC	1876.009
χ^2_B	105.927
df_B	11

Table 4.5. *Maximum likelihood estimates of direct effects in main effects model of post-test procedural and conceptual scores*

	Unadjusted Main Effects Model			Design Effect Adjusted Model		
Parameter	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic	Exact *p*
<u>Direct effects</u>						
Pre-Procedural → Post-Procedural	.744**	.127	.486**	.291	2.55*	.022
Incorrect Example Error Reflection → Post-Procedural	12.052*	5.942	.200*	13.629	.88	.390
Correct Example Error Reflection → Post-Procedural	6.646	5.890	.112	13.509	.49	.630
Correct Example Control → Post-Procedural	2.895	5.781	.050	13.259	.22	.830
Pre-Conceptual → Post-Conceptual	.045	.081	.040	.186	.24	.812
Pre-Procedural → Post-Conceptual	.477**	.101	.410**	.232	2.06	.057
Incorrect Example Error Reflection → Post-Conceptual	9.974*	4.682	.217*	10.739	.93	.368
Correct Example Error Reflection → Post-Conceptual	9.813*	4.624	.216*	10.606	.93	.369
Correct Example Control → Post-Conceptual	9.447*	4.537	.213*	10.406	.91	.378
Post-procedural ↔ Post-Conceptual	213.812**	41.347	.575**	94.834	2.25*	.040
<u>Disturbance variances</u>						
Post-Procedural	473.797**	64.475	.730**	147.880	3.20	.006
Post-Conceptual	291.805**	39.772	.776**	91.221	3.20	.006

* $p < .05$, ** $p < .01$

Prior to adjusting for design effect, this regression model suggested some effects of condition. Prior to adjusting the sample size and standard errors, results indicated that being in the Incorrect Example Error Reflection condition resulted in higher post-test procedural and conceptual scores, after controlling for pre-test procedural scores. Being in the Correct Example Error Reflection condition and the Correct Example control group resulted in higher post-test conceptual scores but not post-test procedural scores, after controlling for pre-test conceptual and procedural scores. The positive effect of the two error reflection conditions and the Correct Example control condition were relatively similar. Incorrect Example Error Reflection Pre-test procedural scores were predictive of both post-test procedural and post-test conceptual scores.

Pre-test conceptual scores did not predict post-test conceptual scores. However, after adjusting the sample size and standard errors in this “main effects” model, the only significant predictor of post-procedural score was procedural score at pre-test. Implications and limitations will be discussed in Chapter 5.

Testing a moderating effect of prior knowledge on the effect of condition on learning.

As prior research suggests that prior knowledge may moderate the effect of instructional interventions on learning, another model was fitted to test whether students' prior knowledge moderates the effect of condition. To determine whether there are differential effects of promoting error reflection through self-explanation prompts alone versus using incorrect worked examples paired with self-explanation prompts, the dummy-coded variables for specific conditions were used to create six interaction terms, each a combination of pre-test Procedural scores or pre-test Conceptual scores, and the three dummy codes for Incorrect Example Error Reflection, Correct Example Error Reflection, and Correct Example Control.

It was hypothesized that promoting error reflection through the use of written self-explanation prompts focusing on errors paired with *correct* worked examples may be more beneficial for students with low prior knowledge than promoting error reflection through the use of *incorrect* worked examples paired with self-explanation prompts focusing on errors.

Table 4.6.

Fit indices for model with pre-procedural by condition interaction on post-procedural

Index	
χ^2_M	2.179
df_M	4
P	.703
RMSEA (90% CI)	.000 (.000 - .109)
^{p}close-fitH_0	.794
CFI	1.00
SRMR	.015
AIC	1879.195
χ^2_B	110.143
df_B	17

However, no hypotheses were made about the particular type of prior knowledge (i.e., procedural or conceptual) that would moderate the effect of condition. Thus separate models were fitted to assess whether prior procedural knowledge or prior conceptual knowledge moderated the effect of condition on either procedural or conceptual post-test scores. These were tested in separate models to reduce potential issues of multicollinearity and to consider restrictions of sample size. First, to assess whether prior procedural knowledge moderates the effect of error reflection on post-test procedural knowledge, three interaction terms were added to the model displayed in Table 4.5 with direct paths to post-test procedural scores. These three terms included the cross products of pre-test procedural knowledge with the dummy-coded variables of

Incorrect Error Reflection, Correct Error Reflection, and Correct Example control. Although the model had good fit (see Table 4.6), the three interaction terms did not predict post-test procedural scores. Therefore, pre-test procedural knowledge did not moderate the effect of condition on post-test procedural knowledge. These results are presented in Table 4.7.

Next, to assess whether prior conceptual knowledge moderates the effect of error reflection on post-test conceptual knowledge, three interaction terms were added to the model displayed in Table 4.5 with direct paths to post-test conceptual scores. These three terms included the cross products of pre-test conceptual knowledge with the dummy-coded variables of Incorrect Error Reflection, Correct Error Reflection, and Correct Example control. The model had good fit and fit indices are presented in Table 4.8. Model estimates are presented in Table 4.9 both before and after applying design effect adjusted standard errors.

Table 4.7. *Maximum likelihood estimates of moderating effects of prior procedural knowledge on the effect of condition on post-test procedural scores*

	Interaction Model			Design Effect Adjusted Model		
Parameter	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic$_{adj}$	Exact *p*
Direct effects						
Pre-Procedural → Post-Procedural	.804**	.255	.524**	.585	1.37	.194
Incorrect Example Error Reflection → Post-Procedural	12.694*	5.957	.209*	13.663	.93	.371
Correct Example Error Reflection → Post-Procedural	6.943	5.947	.116	13.640	.51	.620
Correct Example Control → Post-Procedural	3.186	5.779	.055	13.255	.24	.814
Pre-Conceptual → Post-Conceptual	.041	.080	.037	.183	.22	.827
Pre-Procedural → Post-Conceptual	.478**	.101	.410	.232	2.06	.061
Incorrect Example Error Reflection → Post-Conceptual	9.994*	4.683	.217*	10.741	.93	.370
Correct Example Error Reflection → Post-Conceptual	9.818*	4.625	.216*	10.608	.93	.373
Correct Example Control → Post-Conceptual	9.443*	4.538	.213*	10.408	.91	.382
Post-procedural ↔ Post-Conceptual	217.121**	41.559	.586**	95.320*	2.28*	.042
Pre-Procedural X Incorrect Example Error Reflection → Post-Procedural	-.439	.371	-.105	.851	-.52	.615
Pre-Procedural X Correct Example Error Reflection → Post-Procedural	-.092	.294	-.038	.674	-.14	.894
Pre-Procedural X Correct Example Control → Post-Procedural	.118	.308	.042	.706	.17	.870
Disturbance variances						
Post-Procedural	470.352**	64.200	.719**	147.250	3.19**	.008
Post-Conceptual	291.920**	39.798	.776**	91.281	3.20**	.008

* $p < .05$, ** $p < .01$

Table 4.8. *Fit indices for model with pre-conceptual by condition interaction on post-procedural*

Index	
χ^2_M	3.225
df_M	4
P	.521
RMSEA (90% CI)	.000 (.000 - .132)
^{p}close-fitH_0	.644
CFI	1.00
SRMR	.018
AIC	1875.881
χ^2_B	114.503
df_B	17

Prior to applying design effect adjusted standard errors, a significant interaction was found between pre-test conceptual knowledge and being in the Incorrect Example Error Reflection condition on post-test Conceptual scores. There were no significant interactions between prior conceptual knowledge and the Correct Example Error Reflection or the Correct Example Control condition. Further exploration through graphing the relevant path coefficient suggests that students who struggled at pre-test on the conceptual assessment items may have benefitted most from being in the Incorrect Example Error Reflection condition. However, the design effect adjusted model revealed no significant moderation of the effect of condition by prior knowledge on post-test conceptual scores.

Table 4.9. *Maximum likelihood estimates of moderating effects of prior conceptual knowledge on the effect of condition on post-test conceptual Scores*

	Interaction Model			Design Effect Adjusted Model		
Parameter	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic$_{adj}$	Exact *p*
Direct effects						
Pre-Procedural → Post-Procedural	.744**	.127	.486**	.291	2.55*	.025
Incorrect Example Error Reflection → Post-Procedural	12.052*	5.942	.199*	13.629	0.88	.394
Correct Example Error Reflection → Post-Procedural	6.646	5.890	.112	13.509	.49	.632
Correct Example Control → Post-Procedural	2.895	5.781	.050	13.259	.22	.831
Pre-Conceptual → Post-Conceptual	.253	.160	.226	.367	.69	.504
Pre-Procedural → Post-Conceptual	.477**	.099	.412	.227	2.10	.057
Incorrect Example Error Reflection → Post-Conceptual	11.713*	4.616	.257*	10.587	1.11	.290
Correct Example Error Reflection → Post-Conceptual	9.441*	4.520	.210*	10.367	.91	.380
Correct Example Control → Post-Conceptual	9.626*	4.414	.219*	10.124	.95	.360
Post-procedural ↔ Post-Conceptual	207.827**	40.422	.575**	92.712	2.24*	.045
Pre-Conceptual X Incorrect Example Error Reflection → Post-Conceptual	-.463*	.215	-.232*	.493	-.94	.366
Pre-Conceptual X Correct Example Error Reflection → Post-Conceptual	-.008	.242	-.003	.555	-.01	.989
Pre-Conceptual X Correct Example Control → Post-Conceptual	-.241	.219	-.112	.502	-.48	.640
Disturbance variances						
Post-Procedural	473.801**	64.476	.730**	147.883	3.20**	.008
Post-Conceptual	275.706**	37.768	.743**	86.625	3.18**	.008

* *p* < .05, ** *p* < .01

Testing mediation of the effect of condition on learning by perceived functionality of errors.

To assess whether the influence of the error reflection intervention on post-test procedural and conceptual scores is mediated by changes in students' perceptions of the functionality of errors for learning, a series of path models was tested. The hypothesized relationships are displayed in Figure 4.2. As previously described, it was expected that students in the two error reflection conditions would have higher perceived functionality of errors at post-intervention which would lead to higher post-test procedural scores. Pre-perceived functionality of errors and post-perceived functionality of errors were added to the main effects model displayed in Table 4.5. Direct paths were added from each of the dummy-coded condition variables to post-perceived functionality of errors, controlling for pre-perceived functionality of errors. A direct path was added from post-perceived functionality of errors to post-procedural scores.

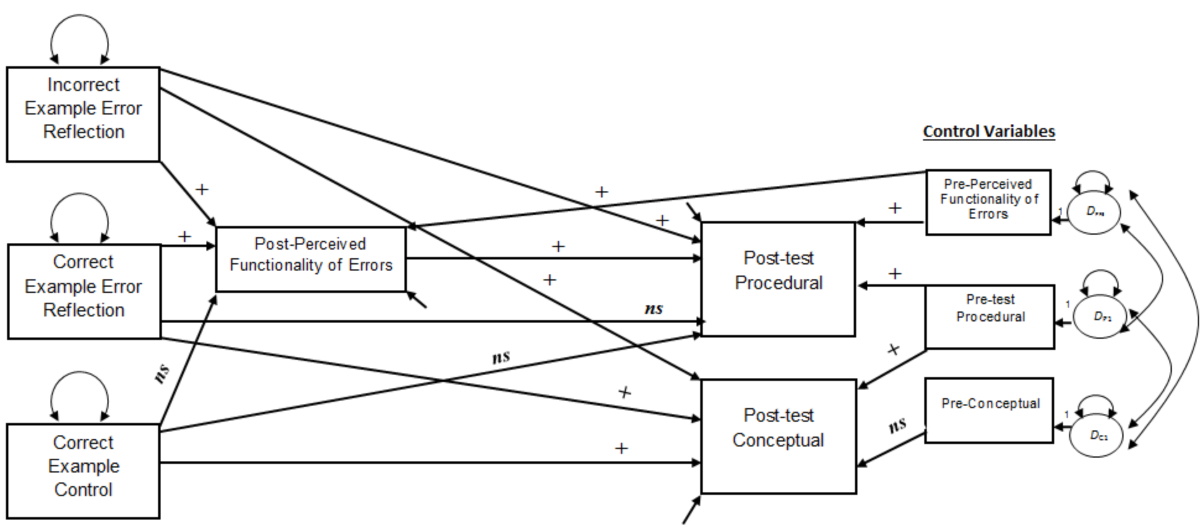


Figure 4.2. Hypothesized mediating effect of perceived functionality of errors on post-test procedural scores

MacKinnon (2008) suggests that indirect effects be assessed using bootstrapping. Fit indices are presented in Table 4.10.

Table 4.10. *Fit indices for mediation model*

Index	
χ^2_M	2.808
df_M	4
P	.590
RMSEA (90% CI)	.000 (.000 - .105)
^{p}close-fitH_0	.740
CFI	1.00
SRMR	.021
AIC	2049.255
χ^2_B	196.315
df_B	21

Although the model had good fit, neither post- nor pre-perceived functionality of errors predicted post-test procedural scores. Estimates are presented in Table 4.11. There was no effect of condition on post-perceived functionality of errors. The sole predictor of post-perceived functionality of errors was pre-perceived functionality of errors. Post-intervention perceived functionality of errors did not mediate the effect of condition on post-test procedural scores. There was also no significant change in students' perceived functionality of errors from pre-intervention to post-intervention.

Table 4.11. *Maximum likelihood estimates testing mediating effect of perceived functionality of errors on post-test procedural scores*

	Mediation Model			Design Effect Adjusted Model		
Parameter	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic$_{adj}$	Exact *p*
Direct effects						
Pre-Procedural → Post-Procedural	.711**	.139	.464**	.319*	2.23	.043
Incorrect Example Error Reflection → Post-Procedural	13.085*	6.524	.228*	14.964	.87	.397
Correct Example Error Reflection → Post-Procedural	5.781	5.856	.095	13.431	.43	.673
Correct Example Control → Post-Procedural	4.114	5.941	..071	13.626	.30	.767
Pre-Perceived Functionality of Errors → Post-Procedural	2.295	3.037	.070	6.966	.33	.747
Post-Perceived Functionality of Errors → Post-Procedural	1.937	3.275	.067	7.512	.26	.800
Pre-Conceptual → Post-Conceptual	.047	.088	.041	.202	.23	.819
Pre-Procedural → Post-Conceptual	.502**	.098	.419**	.225*	2.23	.042
Incorrect Example Error Reflection → Post-Conceptual	10.224	5.319	.228	12.200	.84	.416
Correct Example Error Reflection → Post-Conceptual	9.845*	4.956	.208*	11.367	.87	.401
Correct Example Control → Post-Conceptual	8.812	5.158	.193	11.830	.74	.469
Pre-Perceived Functionality of Errors → Post-Perceived Functionality of Errors	.751**	.074	.666	.170**	4.42	.001
Pre-Procedural → Post-Perceived Functionality of Errors	.003	.004	.051	.009	.33	.749
Incorrect Example Error Reflection → Post-Perceived Functionality of Errors	-.221	.158	-.112	.362	-.61	.552
Correct Example Error Reflection → Post-Perceived Functionality of Errors	.073	.155	.035	.356	.21	.840
Correct Example Control → Post-Perceived Functionality of Errors	-.079	.177	-.039	.406	-.19	.849
Post-procedural ↔ Post-Conceptual	225.661**	41.386	.605**	94.923*	2.38	.032
Indirect effects through Post-Functionality of Errors						
Incorrect Example Error Reflection → Post-Procedural	-.428	.954	-.007	2.188	-.20	.848
Correct Example Error Reflection → Post-Procedural	.141	.627	.002	1.438	.10	.923
Correct Example Control → Post-Procedural	-.153	.738	-.003	1.693	-.09	.929
Disturbance variances						
Post-Procedural	468.064**	51.692	.731	118.561**	3.95	.001
Post-Conceptual	297.407**	40.093	.760	91.958**	3.23	.006
Post-Perceived Functionality of Errors	.399**	.054	.527	.124**	3.22	.006

* $p < .05$, ** $p < .01$

Although the previously mentioned results demonstrate no relationship between condition and perceived functionality of errors, a parallel model was conducted for consistency. This was done by moving the direct paths from pre- and post-perceived functionality of errors that pointed to post-test procedural score as displayed in Figure 4.2 to post-test conceptual scores instead. This model also had good fit and fit indices are presented in Table 4.12 below.

Table 4.12. *Fit indices for mediation model*

Index	
χ^2_M	5.205
df_M	4
P	.267
RMSEA (90% CI)	.045 (.000 - .137)
^{p}close-fitH_0	.444
CFI	.993
SRMR	.039
AIC	2051.652
χ^2_B	196.315
df_B	21

However, as with the prior model, neither post- nor pre-perceived functionality of errors predicted post-test conceptual scores. This demonstrates that post-intervention perceived functionality of errors did not mediate the effect of condition on post-test conceptual scores. Estimates are presented in Table 4.13 below.

Exploring individual differences in the effect of condition on learning.

The last research question posed was consideration of individual differences. Of particular interest was whether the relationships found between condition and post-test conceptual and procedural scores differ by students' URM status. In a prior study (Barbieri & Booth, under review) the influence of worked examples on learning varied by URM status.

Table 4.13. *Maximum likelihood estimates testing mediating effect of perceived functionality of errors on post-test conceptual*

	Mediation Model			Design Effect Adjusted Model		
Parameter	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic$_{adj}$	Exact *p*
Direct effects						
Pre-Procedural → Post-Procedural	.719**	.146	.468**	.335	2.15*	.050
Incorrect Example Error Reflection → Post-Procedural	13.019*	6.471	.226*	14.842	.88	.395
Correct Example Error Reflection → Post-Procedural	5.443	5.884	.090	13.496	.40	.693
Correct Example Control → Post-Procedural	3.403	5.942	.058	13.629	.25	.806
Pre-Perceived Functionality of Errors → Post-Conceptual	-2.827	2.336	-.111	5.358	-.53	.606
Post-Perceived Functionality of Errors → Post-Conceptual	1.715	2.336	.076	5.358	.32	.754
Pre-Conceptual → Post-Conceptual	.050	.090	.043	.206	.24	.812
Pre-Procedural → Post-Conceptual	.497**	.099	.416**	.227	2.19*	.046
Incorrect Example Error Reflection → Post-Conceptual	10.656	5.347	.238	12.264	.87	.400
Correct Example Error Reflection → Post-Conceptual	9.387	5.040	.199	11.560	.81	.430
Correct Example Control → Post-Conceptual	8.665	5.118	.191	11.739	.74	.473
Pre-Perceived Functionality of Errors → Post-Perceived Functionality of Errors	.755**	.071	.670**	.163	4.64**	.0004
Pre-Conceptual → Post-Perceived Functionality of Errors	.001	.003	.010	.007	.15	.887
Incorrect Example Error Reflection → Post-Perceived Functionality of Errors	-.220	.162	-.112	.372	-.59	.563
Correct Example Error Reflection → Post-Perceived Functionality of Errors	.081	.154	.039	.353	.23	.822
Correct Example Control → Post-Perceived Functionality of Errors	-.055	.176	-.028	.404	-.14	.894
Post-procedural ↔ Post-Conceptual	229.457**	42.836	.608**	98.249	2.34*	.035
Indirect effects through Post-Functionality of Errors						
Incorrect Example Error Reflection → Post-Conceptual	-.378	.693	-.008	1.589	-.24	.815
Correct Example Error Reflection → Post-Conceptual	.139	.48	.003	1.101	.13	.901
Correct Example Control → Post-Conceptual	-.095	.523	-.002	1.200	-.08	.938
Disturbance variances						
Post-Procedural	482.450**	51.809	.750	118.830	4.06**	.001
Post-Conceptual	295.281**	41.974	.759	96.272	3.07**	.008
Post-Perceived Functionality of Errors	.401**	.055	.530	0.126	3.18**	.007

* $p < .05$, ** $p < .01$

The findings demonstrated that both correct and incorrect worked examples bolstered URM students' procedural knowledge at post-test compared to a problem solving control group. However, due to the small sample of URM students, it was unclear whether this was indeed a difference based on URM status in particular or whether the finding could be better explained by another difference such as prior knowledge. The current study was planned to attempt to replicate and expand upon this finding.

Participants of the current study are students within a school district that has a marked achievement gap between White and non-Asian minority students. Thus, URM students participating in the current study were hypothesized to begin the study with lower prior knowledge than non-URM students. Therefore, it was expected that any differences found in the influence of the error reflection interventions on post-test conceptual or procedural scores by URM status may be explained by differences in prior knowledge. To test this, differences in pre-test conceptual and procedural scores by URM must first be examined.

A one-way ANOVA revealed significant differences between URM and non-URM students at pre-test on the procedural measure ($F[1,201] = 9.458$, $p = .002$, $MSE = 258.44$). URM students (n = 62) had significantly lower pre-test procedural scores ($M = 11.83$, $SD = 14.05$) than non-URM students (n = 141; $M = 19.36$, $SD = 16.88$). However, no significant differences were found between URM and non-URM students at pre-test on the conceptual measure ($F[1,201] = .192$, $p = .662$, $MSE = 299.14$). The mean for URM students' pre-test conceptual score ($M = 45.16$, $SD = 16.29$) was comparable to that of the non-URM students ($M = 46.32$, $SD = 17.71$).

To determine whether effect of condition on learning varies for URM students, a multiple group analysis should be used. This involves running the same model on URM students and non-URM students separately and then comparing weights of the different variables in the model. If the influence of the error reflection interventions on post-interventions were found to vary by URM status, we planned to consider whether differences in effect of the error reflection interventions by URM status can be mainly accounted for by differences in prior knowledge or prior sense of belonging to math rather than race itself. As previously stated, differences in prior knowledge were expected due to the marked achievement gap found in the particular school district in which the study will be conducted. Further, URM students were expected to have a low sense of belonging to math in general prior to the intervention, due to their underrepresentation in the domain of math. Students' SES, represented by whether they qualify for the free- or reduced lunch program (FRLP), was also to be explored as a contributing factor. However, as the rate of FRLP in the participating school district is generally low for both URM and non-URM students, this is not expected to be a major contributing factor.

The study's sample was initially expected to have a sufficient proportion of URM students. The school district reports that African American students alone make up 50% of the student body. However, at pre-test, only 30.43% of the sample was made up of URM students (63 URM, 144 Non-URM). At post-test, that number dropped to 25.93% (28 URM, 110 Non-URM). Although the URM students who took the post-test were evenly distributed across the four conditions, the URM students within each cell are too few to warrant statistical analyses. Therefore, the planned analyses could not be conducted. However, descriptive statistics for URM students in each condition are

Table 4.14. URM students' post-test measures by condition

	Problem-Solving Control (n = 7)		Correct Example Control (n = 8)		Correct Example Error Reflection (n = 7)		Incorrect Example Error Reflection (n = 6)	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Procedural score	33.33	21.77	26.67	17.46	48.57	20.63	42.22	20.94
Conceptual score	49.69	21.13	59.24	21.79	77.64	11.89	66.67	27.86

provided in Table 4.14 for illustrative purposes. Future work on a larger sample with a higher proportion of URM students is needed to answer this research question.

Predicting Post-Intervention Sense of Belonging to Mathematics

Design effect adjusted standard errors for sense of belonging to math.

As previously stated, due to the high ICCs for post-intervention sense of belonging to math (ICC = .201), design effect adjusted standard errors will be utilized in calculating an adjusted sample size and computing more conservative t-statistics that account for clustering in the data. A slightly greater number of participants completed this post-intervention survey measure than the post-test. The average cluster size during the post-intervention measurement of sense of belonging to mathematics was 14.63 resulting in a design effect of DEFF = $= 1 + .201(14.63 - 1) = 3.74$. This indicates that the variance for post-intervention sense of belonging to math in a complex sampling design is estimated to be almost four times greater than if the current study utilized simple random sampling. The formula for $N_{adj} = \frac{N}{(1+[ICC\{n_j-1\}])}$, with the calculations for DEFF in the denominator, so the adjusted sample size when predicting post-intervention sense of belonging is $N_{adj} = \frac{117}{(3.74)} = 31.28$.

Testing main effects of condition on sense of belonging to math.

To assess whether the error reflection interventions lead to increases in sense of belonging to math, path models were tested utilizing full information maximum likelihood (ML) estimation within MPlus much like those that assessed learning. The hypothesized model in Figure 4.3 was tested. Centered post-intervention sense of belonging to mathematics was regressed onto the dummy code for the two error-reflection conditions and the correct control variable, controlling for pre-intervention sense of belonging to mathematics, pre-test conceptual scores and pre-test procedural scores. A zero on these three variables represents a student in the Problem-Solving Control condition. As previously stated, it was expected that being in one of the two error reflection conditions would predict high post-intervention sense of belonging to math. Higher pre-test scores were also expected to predict higher post-intervention sense of belonging to mathematics.

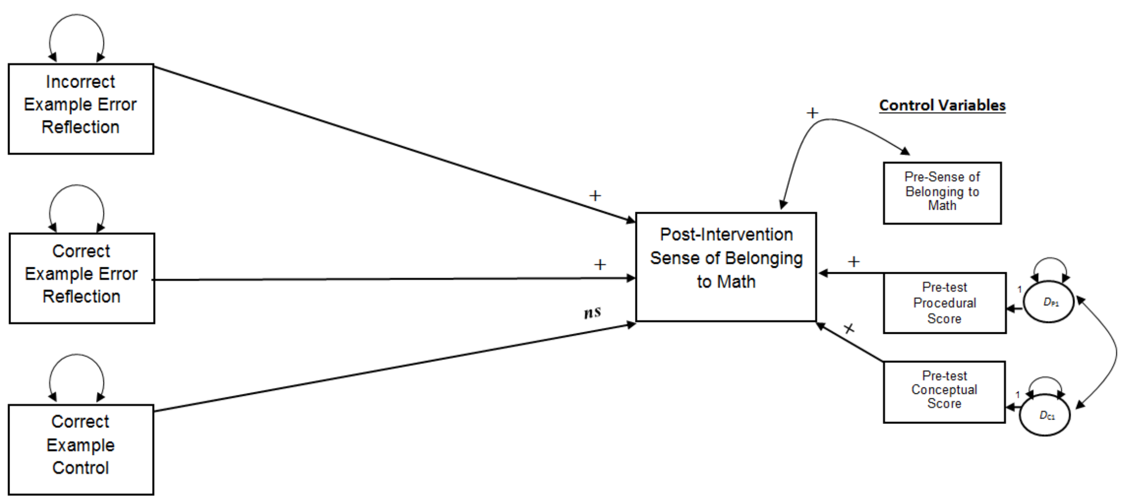


Figure 4.3. Hypothesized main effects model predicting post-intervention sense of belonging to math

The fit indices for the model displayed in Figure 4.3 were poor, as indicated by a significant model chi-square and high RMSEA. These indices are displayed in Table 4.15.

Table 4.15.

Fit indices for hypothesized main effects model

Index	Figure 4.3
χ^2_M	25.186
df_M	5
P	< .001
RMSEA (90% CI)	.141 (.090 - .198)
^{p}close-fitH_0	.003
CFI	.750
SRMR	.071
AIC	899.357
χ^2_B	91.587
df_B	11

Model misfit was likely due to the lack of predictive power of pre-test procedural and conceptual scores for post-intervention sense of belonging. However, modification indices revealed that pre-test procedural scores were related to pre-intervention sense of belonging. Thus, the model was adjusted accordingly until fit was adequate. Figure 4.4 displays the main effects model tested. Fit indices for the revised model are presented in Table 4.16.

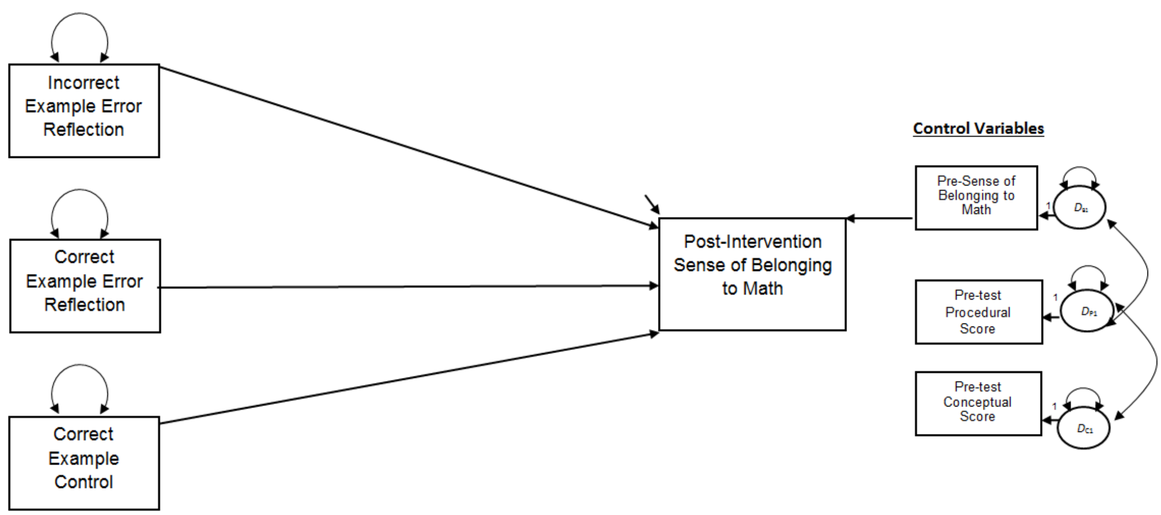


*Figure 4.4.*Main effects model tested for post-intervention sense of belonging to math

Table 4.16.

Fit indices for revised main effect model

Index	Figure 4.4
χ^2_M	15.161
df_M	12
P	.233
RMSEA (90% CI)	.036 (.000 - .084)
^{p}close-fitH_0	.632
CFI	.962
SRMR	.060
AIC	4351.744
χ^2_B	97.321
df_B	15

Although the model had adequate fit, there was no effect of condition on post-intervention sense of belonging to mathematics. Estimated before and after implementing design effect adjusted standard errors are presented in Table 4.16. Results demonstrate that the only significant predictor of post-intervention sense of belonging to mathematics is pre-intervention sense of belonging to mathematics. Pre-test procedural scores positively correlated with pre-intervention sense of belonging to mathematics.

Table 4.17. *Maximum likelihood estimates of direct effects in Models of Post-intervention Sense of Belonging to Math*

	Unadjusted Main Effects Model			Design Effect Adjusted Model		
Parameter	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic	Exact *p*
<u>Direct effects</u>						
Pre-Belonging → Post- Belonging	.618**	.074	.598**	.143	4.32	.0002
Incorrect Example Error Reflection → Post-Belonging	-.008	.224	-.003	.433	-.02	.985
Correct Example Error Reflection → Post-Belonging	.267	.245	.105	.474	.56	.578
Correct Example Control → Post-Belonging	-.090	.232	-.035	.449	-.20	.843
Pre- Belonging ↔ Pre-Procedural Score	5.409**	1.279	.310**	2.474	2.19	.039
Pre-Procedural ↔ Pre-Conceptual	47.057*	19.095	.167*	36.930	1.27	.215
<u>Disturbance variances</u>						
Post-Sense of Belonging to Math	.769**	.101	.628**	0.195	3.94	.001

* *p* < .05, ** *p* < .01

Testing mediation of the effect of condition on sense of belonging to math by perceived functionality of errors.

To assess whether the influence of the error reflection intervention on sense of belonging to mathematics is mediated by students' perceptions of the functionality of errors for learning, the hypothesized model in Figure 4.5 was tested. It was initially hypothesized that the error reflection interventions would have indirect effects on post-intervention sense of belonging to mathematics through increases in post-intervention perceptions of functionality of errors for learning.

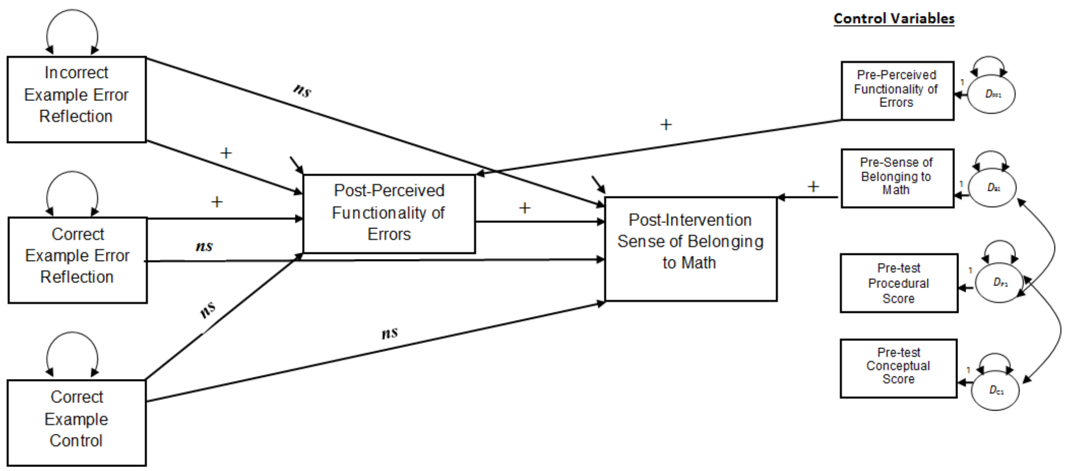


Figure 4.5. Hypothesized mediating effect of perceived functionality of errors on post-intervention sense of belonging to math

The fit indices for the model displayed in Figure 4.5 were poor, as indicated by a significant model chi-square and high RMSEA. Fit indices are presented in Table 4.17. Modification indices revealed that pre-intervention perceived functionality of errors was related to pre-intervention sense of belonging. Thus, a correlation was added between the disturbance variances of pre-intervention sense of belonging and pre-intervention perceived functionality of errors. Table 4.18 also displays fit indices of the revised mediation model run with bootstrapping.

Table 4.18.

Fit indices for hypothesized and revised mediation models

Index	Hypothesized Model	Revised Model
χ^2_M	33.975	18.884
df_M	17	18
P	.009	.399
RMSEA (90% CI)	.071 (.035 - .106)	.015 (.000 - .065)
^{p}close-fitH_0	.147	.841
CFI	.890	.995
SRMR	.077	.048
AIC	4430.618	5081.746
χ^2_B	175.910	182.208
df_B	21	21

Table 4.19 displays model estimates before and after implementing the design effect adjusted standard errors to account for clustering in post-intervention sense of belonging to mathematics.

After implementing design effect adjusted standard errors, the results suggest that perceived functionality of errors does not mediate the effect of condition on post-intervention sense of belonging to math. Further, the only significant predictor of post-intervention perceived functionality of errors was pre-intervention perceived functionality of errors. However, students' perceptions of the functionality of errors after the intervention did indeed predict post-intervention sense of belonging to math. This seems to be unrelated to the experimental manipulation. Implications are discussed in Chapter 5.

Table 4.19. *Maximum likelihood estimates of direct effects in Models of Post-intervention Sense of Belonging to Math*

Parameter	Unadjusted Main Effects Model			Design Effect Adjusted Model		
	Unstandardized	*SE*	Standardized	SE_{adj}	t-statistic	Exact *p*
<u>Direct effects</u>						
Pre-Belonging → Post- Belonging	.547**	.087	.540	.168	3.25	.004
Incorrect Example Error Reflection → Post-Belonging	.097	.180	.039	.348	.28	.783
Correct Example Error Reflection → Post-Belonging	.152	.240	.061	.464	.33	.746
Correct Example Control → Post-Belonging	-.066	.230	-.026	.445	-.15	.883
Pre-Belonging ↔ Pre-Procedural Score	5.381**	1.153	.308	2.230	2.41*	.025
Pre-Procedural ↔ Pre-Conceptual	47.128*	19.372	.168	37.466	1.26	.222
Pre-Functionality of Errors → Post-Belonging	-.238	.142	-.180	.275	-.87	.396
Post-Functionality of Errors → Post-Belonging	.545**	.119	.441	.230	2.37*	.027
Pre-Functionality of Errors → Post-Functionality of Errors	.718**	.076	.673	.147	4.88*	.0001
Incorrect Example Error Reflection → Post-Functionality of Errors	-.190	.156	-.094	.302	-.63	.535
Correct Example Error Reflection → Post-Functionality of Errors	.116	.156	.058	.302	.38	.704
Correct Example Control → Post-Functionality of Errors	-.054	.170	-.027	.329	-.16	.871
Pre-Procedural Score → Post-Functionality of Errors	.003	.004	.061	.008	.39	.702
Pre-Procedural ↔ Pre-Functionality of Errors	1.096	.928	.082	1.795	.61	.548
Pre-Belonging ↔ Pre-Functionality of Errors	.253**	.064	.287	.124	2.04	.053
<u>Indirect effects through Post-Functionality of Errors</u>						
Incorrect Example Error Reflection → Post-Belonging	-.103	.087	-.041	.168	-.61	.547
Correct Example Error Reflection → Post-Belonging	.063	.087	.025	.168	.37	.712
Correct Example Control → Post-Belonging	-.030	.094	-.012	.182	-.17	.870
<u>Disturbance variances</u>						
Post-Sense of Belonging to Math	.632**	.138	.536**	.267	2.37*	.027
Post-Functionality of Errors	.402**	.054	.521**	.104	3.85**	.001

* $p < .05$, ** $p < .01$

Exploring individual differences in the effect of condition on sense of belonging to math.

The last research question posed was consideration of individual differences. Of particular interest was whether the relationships found between condition and post-intervention sense of belonging differ by students' URM status. In an earlier study (Barbieri & Booth, under review), the influence of worked examples on sense of belonging to math varied by URM status. The findings demonstrated that correct and incorrect worked examples bolstered URM students' sense of belonging in comparison to a problem solving control group. The current study was proposed in an attempt to replicate and expand upon this finding.

URM students were expected to have a lowered sense of belonging prior to the intervention based upon their underrepresentation in the field of mathematics. As the district that participated in the current study has a marked achievement gap between White and non-Asian minority students, URM students participating in the proposed study were expected to begin the intervention with lower prior knowledge than non-URM students. As previously stated, URM students did have lower conceptual knowledge at pre-test than non-URM students. Due to the expectation that prior knowledge would predict sense of belonging, it was previously hypothesized that any differences found in the influence of the error reflection interventions on sense of belonging to mathematics by URM status may be explained by differences in prior knowledge and pre-intervention sense of belonging to mathematics. To test this, differences in pre-intervention sense of belonging by URM status must first be examined.

A one-way ANOVA revealed significant differences between URM and non-URM students prior to the intervention on sense of belonging to mathematics (*F*[1,202] = 5.189, *p* = .024, *MSE* = 1.129). URM students (n = 62) had significantly lower sense of belonging to math (*M* = 4.28, *SD* = 1.10) than non-URM students (n = 142; *M* = 4.65, *SD* = 1.04). To determine whether effect of condition on sense of belonging to math varies for URM students, a multiple group analysis should be used. This involves running the same model on URM students and non-URM students separately and then comparing weights of the different variables in the model. If the influence of the error reflection interventions on post-intervention sense of belonging to math was found to vary by URM status, we planned to consider whether differences in effect of the error reflection interventions by URM status can be mainly accounted for by differences in prior knowledge or prior sense of belonging to math rather than URM status itself. However, as with planned multiple group analyses for the post-test procedural and conceptual scores by URM status, the number of URM students in the analysis sample is substantially lower than originally expected. Thus, analyses by URM status could not be conducted. However, descriptive statistics for URM students in each condition are provided below in Table 4.20 for illustrative purposes Future research should consider these potential relationships with a larger and more diverse sample. Implications of these and other results are discussed in the upcoming Chapter 5.

Table 4.20. URM students' post-intervention survey measures by condition

	Problem-Solving Control (n = 8)		Correct Example Control (n = 9)		Correct Error Reflection (n = 8)		Incorrect Error Reflection (n = 10)	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Sense of belonging to math	3.68	.89	3.91	1.86	4.18	1.54	3.98	.70
Perceived functionality of errors	3.33	.73	3.40	1.14	4.16	.41	3.50	1.00

CHAPTER 5

DISCUSSION

Overview

This chapter is divided into five major segments: Summary, Discussion, Summary of Limitations and Recommendations for Future Research, Conclusion, and Educational Implications. The Summary presents an overview of the current study's purpose, methodology and findings. The Discussion section provides a more detailed interpretation of the findings, including both expected and unexpected results, limitations, and suggestions for future research. The third segment summarizes limitations of the study methodology and claims that can be made with the current data, as well as recommendations for future directions. The chapter closes with general conclusions about the study results and implications for both theory and practice.

Summary

Overview of the problem.

Algebra I is considered a gatekeeper course for higher level math courses as well as select science courses (Matthews & Farmer, 2008) and eventually college admittance (Schneider, Swanson, & Riegle-Crumb, 1998). However, adolescents in the US struggle with both mathematics achievement (Kelly et al., 2013) and motivation for mathematics (Wang & Pomerantz, 2009; Wang, 2013). Further, there is a well-documented achievement gap between underrepresented minority students and their White and Asian counterparts in mathematics that has not narrowed since 2008 (NCES, 2013). Despite increased acknowledgement of these issues, the US continues to perform below average in international mathematics assessments and both ethnic and income achievement gaps persist. While it is likely that many influential factors on student achievement and motivation are outside of the classroom environment, such as the important influence of families and peer groups (Reeves, 2012), the current study sought to explore a more malleable factor of achievement and motivation: instruction.

As achievement and motivation are closely linked, focusing on simultaneously increasing student learning and improving student motivation may be key to promoting lasting change in mathematics achievement. The current study focused on an understudied motivation construct that may be particularly relevant to students who are underrepresented in the mathematics domain: sense of belonging to mathematics (Good, Rattan, & Dweck, 2012). Although sense of school belonging in general has been thoroughly studied (Anderman, 2003; Goodenow, 1993), sense of belonging to the academic domain of mathematics in particular is relatively understudied and is a potentially fruitful area of exploration when concerned with students who struggle with mathematics in particular.

In a prior study, a classroom intervention that utilized either correct or incorrect worked examples paired with written self-explanation prompts demonstrated benefits for underrepresented minority students' sense of belonging to mathematics in comparison to a control group (Barbieri & Booth, under review). In addition, students with low prior knowledge at the start of the study demonstrated particular benefits from being in the incorrect worked examples condition. One possible causal mechanism of these benefits is the reflection upon errors within the incorrect worked examples condition. It is possible that reflecting upon errors could refine learners' knowledge of concepts and procedures. This reflection may also alter their perceptions upon the usage of errors for learning, which has been found to be a predictive factor of adaptive reactions to struggles in mathematics problem-solving (Steuer, Rosentritt-Brunn, & Dresel, 2013). However, the incorrect examples condition in this particular study also prompted students to reflect upon correct concepts. Further, students' perceptions of the functionality of errors were not measured. Thus, the current study sought to isolate the effects of promoting error reflection in particular from the well-studied effects of worked examples and self-explanation prompts that focus on reflecting on correct concepts and procedures.

Review of purpose and hypotheses.

The main purpose of the current study was to determine the influence of an error reflection intervention on middle school students' algebra learning and sense of belonging to mathematics. Specifically, the current study assessed an educational intervention that utilized various forms of worked examples and self-explanation prompts in ways that were designed to promote reflection on errors and hence deeper learning of algebraic content as well as greater feelings of belonging to the mathematics domain. Promoting reflection on errors was hypothesized to promote learning by enabling refinement of both concepts and procedures being reflected upon. Promoting error reflection was also expected to increase students' sense of belonging to the math domain by altering students' perspectives of errors' usefulness as a learning tool. In particular, it was hypothesized that error reflection would lead to a normalization of errors which should in turn lead to more inclusive feelings related to normally exclusive field of mathematics. Although listed and tested separately, nested within this hypothesis was the expectation that participants' perceptions of the functionality of errors would be altered by the error reflection intervention. Specifically, it was surmised that the error reflection intervention would serve to normalize errors as part of the learning process and that this would be reflected in increases of their perceptions of the functionality of errors for learning. This change in perception could then potentially allow students to see mathematics as a more inclusive domain in which any learner may belong regardless of prior struggles, leading to increases in sense of belonging. This altered perception of errors could also encourage students to engage deeply with the learning material and lead to the learning increases previously suggested.

Review of the methodology.

The current study employed an in-vivo classroom experiment to answer the research questions posed. Middle school students (N = 207) across eleven Algebra I classrooms were randomly assigned within classroom to one of four experimental worksheet conditions. The experimental manipulation was within the worksheet conditions students were assigned to. The intervention consisted of four worksheets on solving quadratic equations completed over the course of several weeks that vary on whether students are asked to reflect on correct or incorrect concepts and procedures. There were two control conditions which are referred to as the *Problem-Solving Control* condition and the *Correct Examples Control* condition, and two experimental error reflection conditions which are referred to as the *Incorrect Examples Error Reflection* condition and the *Correct Examples Error Reflection* condition. The Problem-Solving Control worksheets simply required students to practice solving problems with no worked examples or self-explanation prompts. The remaining three conditions each had worked example-problem pairs and written self-explanation prompts throughout.

The examples across these conditions all displayed a fictitious student solving the same problems. The worked examples varied by the correctness of the problem and whether writing prompts asked for self-explanation of an error or a correct concept or procedure. In the Correct Examples Control condition, students studied a correctly solved problem and were asked to reflect upon and provide a written explanation of a correct concept or procedure within the problem. In the Correct Examples Error Reflection condition, students studied a correctly solved problem and were asked to reflect upon and provide a written explanation of a potential error that a student may make in a problem such as this. In the Incorrect Example Error Reflection condition, students studied an incorrectly solved problem and were asked to reflect upon and provide a written explanation of an incorrect concept or procedure displayed within the problem. The practice problems paired with each worked example were the same across these three conditions. These manipulations were made so that the effect of promoting error reflection itself could be determined without the effects being necessarily tied to studying incorrect worked examples. If the effects from the aforementioned study (i.e., Barbieri & Booth, under review) are simply due to the benefits of worked examples and self-explanation prompts, students in the three worked examples conditions should have equally outperformed the problem-solving control group. However, if it is indeed error reflection that is the key factor in promoting deeper learning of the content, than the two error reflection conditions should outperform the Correct Examples Control group and the Problem-Solving Control group. Lastly, if error reflection is indeed a useful learning strategy, there may still be particular design features that are more effective at promoting error reflection than others. Comparing the two error reflection intervention conditions could allow one to determine whether viewing the error itself is necessary to promote useful error reflection or if considering a hypothetical error while simultaneously studying a correct procedure may be more fruitful in refining algebraic knowledge.

Two survey measures were administered pre- and post-intervention that measured students' sense of belonging to mathematics (i.e., Survey A) and students' perceived functionality of errors (i.e., Survey B) along with a several other motivation constructs. These surveys are presented in Appendix B and C. The other motivation constructs measured included constructs within Eccles' and colleagues' (Wigfield & Eccles, 2002) expectancy-value framework (i.e., math self-concept, perceptions of math importance, interest in math) and Dweck's entity view of intelligence (Dweck, 1999). Although not the focus of the current study, these motivation measures were taken to test alternative hypotheses in the event that there were changes in students' sense of belonging to mathematics or perceived functionality of errors.

The learning measure included a test that measured students' conceptual and procedural knowledge of solving quadratic equations and is displayed in Appendix A. This measure was used as a pre-test, an immediate post-test and as a delayed post-test. The order of data collection is presented in Chapter 3 and Table 3.1. Students began the study by taking survey A followed by the pre-test the next day. Students took survey B after the pre-test and before beginning the intervention. Once students completed all four worksheets, they took survey B again, followed by survey A, and then took the immediate post-test. Students were scheduled to complete the delayed post-test two weeks after completing the immediate post-test. Participants worked at their own pace so the exact timing of each of the measures varied across participants. However, the order of the administration of the measures remained constant across all participants.

Despite agreeing to complete all components of the study, none of the teachers administered the delayed post-tests during the allotted time. Discussion with the math coordinator about the matter revealed that this was due to a combination of their unpredictable testing schedule, school closings and delays due to inclement weather, and class trips delaying data collection. This then led teachers to decide to drop the collection of the delayed post-tests in favor of covering regular instructional materials that needed to be made up as well. Therefore, analysis on delayed post-tests is not possible and the findings presented only reflect study results for measures completed during the allotted timeframe (i.e., up to and including the immediate post-test). However, as no particular hypotheses were made regarding retention at delayed post-test, the data collected were deemed acceptable for answering the posed research questions. All analyses were conducted using MPlus Version 7.3 (Muthén & Muthén, 1998 - 2014) using full information maximum likelihood (FIML) estimation.

Major findings.

A brief overview of the major findings relevant to each research question is first presented here. The discussion segment that follows provides a more detailed interpretation of the findings. When considering the findings, it is important to consider that the data collected in the current study had considerable clustering effects making the planned single-level analyses difficult to conduct. After making modifications to the standard errors in each of the models tested according to the design effect, many of the effects of condition were eradicated. The following is a brief overview of the general answers to the research questions based on results utilizing design effect adjusted standard errors.

Research Question 1.

1. Does promoting error reflection within a worked examples intervention improve middle grade students' algebra learning? Are there differential effects of promoting error reflection with the error displayed in comparison to the error simply mentioned?

   a. If so, does prior knowledge moderate the effect of condition?

The first research question posed concerned whether the error reflection interventions improved middle grade students' algebra learning. Prior to applying design effect adjusted standard errors to the models, the Incorrect Examples Error Reflection condition did seem to lead to greater learning benefits on the procedural and conceptual measures. The Correct Example Error Reflection condition also suggested some learning benefits for conceptual knowledge. However, after applying design effect adjusted standard errors to account for clustering in the data, participants in each of the conditions demonstrated comparable learning gains from pre- to post-test on both the procedural and conceptual items. The effect of the error reflection intervention was no longer statistically significant.

Another question posed was whether prior knowledge moderated the effect of condition on learning. Analyses prior to adjusted standard errors suggested that prior conceptual knowledge moderated the effect of being in the Incorrect Example Error Reflection condition on post-test conceptual scores. However, after adjusting standard errors, this was no longer statistically significant. Thus, we should conclude that neither pre-test procedural nor pre-test conceptual scores moderated the effect of condition on learning.

Research Question 2.

2. Does promoting error reflection within a worked examples intervention improve middle grade students' sense of belonging to mathematics?

The second research question posed concerned whether the error reflections improved middle grade students' sense of belonging to mathematics. Sense of belonging to mathematics was high at pre-intervention across all conditions and remained stable from pre- to post-intervention. The error reflection conditions did not lead to improved sense of belonging to mathematics.

Research Question 3.

3. Is the influence of error reflection on the outcome variables mediated by changes in students' perception of the functionality of errors?

The third research question posed concerned whether the influence of the error reflection interventions on the outcome variables was mediated by changes in students' perceptions of the functionality of errors from pre- to post-intervention. Perceived functionality of errors for learning was high at pre-intervention across all conditions and remained stable from pre- to post-intervention. The error reflection conditions did not lead to improved perceived functionality of errors for learning. Further, pre- and post-intervention perceived functionality of errors was unrelated to post-test conceptual and procedural scores, even prior to administering design effect adjusted standard errors. Pre- and post-intervention perceived functionality of errors was also unrelated to post-test sense of belonging to mathematics, after applying design effect adjusted standard errors. There were no indirect effects of condition on the outcome variables through perceived functionality of errors, indicating that perceived functionality of errors did not mediate the effect of condition on learning or post-intervention sense of belonging.

Research Question 4.

4. Do the relationships between the error reflection intervention and students' learning and sense of belonging to math differ by URM status? If so, can prior knowledge and prior sense of belonging explain these differences?

The fourth and final research question posed concerned whether there were differential effects of condition on learning and sense of belonging to math for URM students. Although the sample was anticipated to include an adequate sample of URM students, this did not hold true during data collection. Of the final analysis sample, ten or fewer students within each condition completed the post-test and post-intervention survey measures. Thus, analyses could not be conducted by URM status and this research question is still open for exploration in future work. These major findings will be discussed in more depth in relation to the existing literature in the section that follows.

Discussion

The main purpose of the current study was to determine the influence of an error reflection intervention on students' algebra learning and sense of belonging to mathematics. This section will provide a discussion of the results tied to each research question posed in the current study. Connections to related literature will be made and unexpected findings will be discussed.

Research question 1.

Prior to applying design effect adjusted standard errors to the models, the Incorrect Examples Error Reflection condition did seem to lead to learning benefits on the procedural and conceptual measures. The Correct Example Error Reflection condition also suggested some learning benefits for conceptual knowledge. However, after taking a more conservative approach to analyzing the data and considering the design effect due to the clustering of the data, error reflection conditions did not seem to have particular learning benefits. All four conditions showed equal learning gains on procedural and conceptual items. There was no difference in learning by condition. While the purpose of the current study was not to test the worked examples effect but rather the effect of error reflection as promoted through worked examples, the finding that the Problem-Solving Control group learned equally as much as the three worked examples conditions is inconsistent with much prior research.

Much research on worked examples has demonstrated learning benefits from worked examples in a variety of domains (Booth et al., in press; Caroll, 1994; Pol et al., 2009; Ward & Sweller, 1990) that suggests benefits of correct worked examples in general as well as benefits of a combination of correct and incorrect worked examples (Durkin & Rittle-Johnson, 2012; Grosse & Renkl, 2007). However, some prior research has suggested that worked examples are not effective across the board. For example, not surprisingly, Renkl (1997) found that students must actively engage with worked examples for learning to occur. Without active engagement, students may not focus on the relevant features within the example and fail to grasp the concepts or procedures that the examples were designed to present. It is quite possible that students were not actively engaged in any of the worked examples conditions, including those focused on error reflection. However, this possible explanation for lack of findings is not a particularly useful one, as the worked examples, and the error reflection examples in particular, were designed in ways that they should have been inherently more engaging than the problem-solving control worksheets. One of the hypothesized arguments as to why error reflection should be more effective than either reflecting upon correct concepts or procedures or problem-solving alone is that error reflection should promote a deeper level of processing than either of the two control conditions. A possible explanation to this finding is that each of the conditions was equally engaging for students. However, there is no way to test this hypothesis with the current data. Future research may attempt to address this issue by asking participants to report on their engagement with the materials or even by measuring and comparing effort expenditure between the conditions.

A possible explanation as to why the Incorrect Example Error Reflection condition did not outperform the other conditions is that benefits from studying incorrect worked examples are more likely to be seen at a delay, as suggested by Adams and colleagues (2014). These researchers suggest that the deeper processing of material that stems from working with incorrect worked examples leads to deeper learning over time rather than immediately. They term this as a result of desirable difficulties. This hypothesis could be tested if delayed post-tests were completed. However, this potential explanation does not explain why the two correct example conditions did not outperform the Problem-Solving Control group. Also, the *desirable difficulties* framework does not clearly explain the causal mechanism of effects only seen at a delay or how learning occurs exactly between immediate and delayed post-tests. Further, prior work has seen benefits of incorrect worked examples at an immediate post-test (e.g., Heemsoth & Heinze, 2014).

Certain learner characteristics often determine the effectiveness of educational interventions. Students' prior knowledge is an important factor to consider when designing and implementing instructional interventions. When initially considering whether prior knowledge moderated the effect of condition on learning, a significant interaction did suggest that prior conceptual knowledge moderated the effect of being in the Incorrect Example Error Reflection condition on post-test conceptual scores. However, after adjusting standard errors, this moderation effect was no longer statistically significant. It was not deemed appropriate to interpret the initial moderation found. Thus, neither pre-test procedural nor pre-test conceptual scores moderated the effect of condition on learning when using the more conservative approach for analyses. These findings are inconsistent with prior research on worked examples. However, no prior work to date addresses the moderating effects of prior knowledge on the influence of error reflection on learning in particular. Kalyuga and colleagues (2001) found that correct worked examples were more beneficial for students with low prior knowledge. This was also found to be true in our own work (Barbieri & Booth, under review), in which students with low prior knowledge of solving systems of equations showed particular learning gains when studying incorrect worked examples. Heemsoth and Heinze (2014) found greater learning benefits of incorrect rather than correct worked examples for students with high prior knowledge. They suggest that students benefit more from incorrect worked examples once they have developed a foundational knowledge of the to-be-learned material with correct worked examples.

The null results relating the effects of condition on learning were surprising. Great effort was made to design worked examples according to Atkinson and colleagues' (2000) suggestions on structuring examples in ways that should be beneficial to students' learning. Features were highly integrated within the examples and steps taken by fictitious students were presented visually separate within the examples to avoid the split-attention effect and increase clarity. Atkinson and colleagues also suggested that steps within the examples be labeled. This was not done in the current study and could have potentially improved results. Another potential reason for lack of effect of condition could be that some classrooms did not fully integrate worksheets into the appropriate lessons. In some cases, teachers had their students work on worksheets prior to learning the content presented within the worksheet. In other cases, teachers used worksheets well after learning the content presented within the worksheet. Thus, the implementation of the intervention could have been suboptimal. More careful timing and integration of worksheets into the appropriate lesson could reveal the true effects of the error reflection intervention on learning.

Yet another possible explanation is that adjusting the standard errors to protect against Type I error led to committing a Type II error. In other words, utilizing the design effect adjusted standard errors to interpret the current data may be too conservative of an approach and may be hiding an effect of condition that actually does exist. There is reason to believe that this is possible. Prior work in this area suggests that there should be some effect of condition, if not as a main effect then as an effect moderated by prior knowledge. Prior research conducted by the authors' lab group has suffered from similar implementation issues and used similar design features as utilized in the current study's worksheets and still found effects of condition (Booth et al., in press). As the current study utilized random within-class assignment, and controlled for prior knowledge when predicting outcome measures, it is possible that the clustering seen at post-test could potentially be ignored when concerned mainly with effects of condition. Clustering was also seen at pre-test. Therefore, it is possible that even if the effect of condition varied across the eleven participating classes, this effect might be evenly distributed across the four conditions and not particularly relevant to consider when interpreting effects of condition on learning. This may especially be true for the current study as there were no differences at pre-test across the four conditions. However, due to the limited number of clusters, two-level modeling was not possible for the current study. Future research should explore the effects of error reflection on learning with a greater number of classrooms to be able to determine the effect of condition even after accounting for differences at the classroom level.

Research question 2.

Prior research demonstrates that sense of belonging to the domain of mathematics is important for math learning and motivation (Barbieri & Booth, under review; Good et al., 2012). However, the error reflection conditions did not lead to improved sense of belonging to mathematics at post-test even before adjusting standard errors with the design effect in mind.

Sense of belonging to mathematics was high at pre-intervention across all conditions and remained stable from pre- to post-intervention. Thus, modeling change was not possible. One possible explanation for the lack of change in sense of belonging to mathematics is that the implementation issues mentioned previously did not allow students to fully engage in the intervention as planned. Another likely explanation is that student motivation develops over a long period of development and is particularly difficult to alter in such a short time frame. Students' sense of belonging to mathematics may have already developed at an earlier age and studying this construct at an earlier age may be more informative. However, it is important to note that sense of belonging to mathematics was relatively high in this sample, with a mean of M = 4.54 at pre-test out of a possible range of 1.00 - 6.00. Therefore, it is also possible that studying this construct in late adolescence may capture a lower sense of belonging that may be more informative and relevant to student mathematics achievement. Although we know that sense of school belonging declines during adolescence (Anderman, 2003), the development of sense of belonging to the domain of mathematics has not been studied. Thus, future research should assess sense of belonging to mathematics across the adolescent years to understand its development. This may also inform decisions on potential interventions specific to feelings of belonging in the domain of mathematics.

Good and colleagues (2012) found that a stereotype threat manipulation *decreased* female college students' sense of belonging to the math domain whereas an intervention focused on promoting an incremental view of intelligence allowed students to maintain their already high sense of belonging to mathematics. These students, as those in the current study, began the study with a high sense of belonging. More work is needed to understand how this sense could be *increased,* particularly for those students with low feelings of belonging, and if this increase translates to learning gains.

The error reflection interventions were hypothesized to alter sense of belonging to the mathematics domain by promoting the idea that mathematics is a more inclusive domain. However, this idea was not directly promoted and students' perceptions of the mathematics community in general were not measured. One possible explanation as to why there was no change in sense of belonging to mathematics is that the error reflection intervention did not actually promote the idea of mathematics being an inclusive domain. This holds the assumption that middle school students view mathematics as an exclusive domain to begin with. In order to test whether reflecting on errors can lead to this alteration in perceptions of the math domain either short term or over a long period of time, future research should not only measure sense of belonging over time but also ask participants to provide their views on mathematics as a domain, including whether they believe mathematics to be an exclusive or elitist domain or whether they see it as a more inclusive and open domain at this young age.

Still, there may be other more suitable interventions for altering students' sense of belonging to mathematics than the error reflection interventions tested here. It is still possible that making errors a normal part of the learning process in mathematics classrooms may have both learning and motivation benefits. However, assessing whether existing interventions for sense of school belonging may also be effective for sense of belonging to math in particular may be a more productive effort.

For example, Walton and Cohen (2007) designed an intervention that required students to read fictitious reports of upperclassmen reflecting on their struggles in the first year of being in a computer science program. This intervention aimed to increase awareness that all early college students struggle and worry about their academic fit, regardless of race. Researchers found benefits of this intervention for Black college students but not white students. This is just one example of a potentially useful intervention that can be applied and tested for altering students' sense of belonging to mathematics. Interventions may need to be more clear, obvious, and rigorous than the subtle error reflection intervention tested here, when attempting to alter students' achievement motivation and perceptions of themselves that have presumably been developed over a course of several years and salient success and failure experiences in the domain of mathematics.

Research question 3.

Students' perceptions of their classrooms' error climate has been found to influence how they react to struggles in mathematical problem-solving (Steuer, et al., 2013). Students' perceptions of how functional errors are for their own learning is a component of students' *perceived error climate*. It was hypothesized that changes in students' perceived functionality of errors would mediate the effect of condition on both learning and sense of belonging to mathematics. However, as with students' sense of belonging to mathematics, students' perceived functionality of errors did not vary between pre- and post-intervention. Pre-intervention perceived functionality of errors for the sample was relatively high with a mean of M = 3.59 out of a possible range of 1.00 – 5.00. This remained stable at post-intervention. The perceived functionality of errors scale asks students to report on whether and how errors are treated as learning tools within the classroom. There are several possibilities as to why the error reflection conditions did not lead to perceptions of greater functionality of errors. The first is similar to the explanation provided for the lack of change in students' sense of belonging to mathematics. Perceptions of errors may be developed over many years of experience both within and outside of the classroom and may require a more rigorous and long-lasting intervention than the one utilized in the present study.

In several classrooms, teachers did not make attempts to fully integrate the worksheets into their regular teaching. Therefore, another possible explanation for the lack of change in this measure is that the interventions did not actually alter how teachers and classmates handled errors in the classroom outside of use on the worksheets. However, it can be argued that students' *perception* of the functionality of errors is what is key here and not teachers' and students' actual use of errors for learning in the classroom. If the error reflection interventions led students to view errors as a useful learning tool, they should have reported higher functionality of errors on the self-report measure. Still, another possibility is that when reporting on this measure, students did not consider the worksheets as part of their normal classroom practices and instead reported on practices outside of the worksheets. Lastly, administering the intervention over a longer period of time may have lead to more change in students' perceptions of how functional errors are for their own learning in mathematics.

It is interesting and surprising to note that students' perceived functionality of errors was unrelated to students' procedural and conceptual scores but was related to post-intervention sense of belonging prior to administering adjusted standard errors. The only prior work conducted with this construct in mind focused not on overall learning measures as outcome variables but on strategy use within difficult assessment situations

(Steuer et al., 2013). It may be possible that perceived functionality of errors is more predictive of strategy usage, persistence, and achievement motivation than on learning in general which can be affected by a number of factors besides strategy use. Yet another possibility is that utilizing the full *perceived error climate* scale as opposed to simply the subscale of *perceived functionality of errors* may provide more informative findings related to how students view errors in mathematics classrooms both before and after an error reflection intervention.

Research question 4.

In a prior study conducted by the lab group of the author, a correct and incorrect worked examples intervention paired with written self-explanation prompts demonstrated benefits for underrepresented minority students' sense of belonging to mathematics in comparison to a problem-solving control group (Barbieri & Booth, under review). Further, students with low prior knowledge demonstrated particular learning gains from incorrect worked examples. Due to the restricted number of URM students within the sample of the aforementioned study as well as the confounding factor that the few URM students within this sample did indeed have low prior knowledge, further exploration is needed to interpret these findings. Although the sample was anticipated to include an adequate sample of URM students upon recruitment based upon the school districts' published demographics data, this did not prove to be the case once data collection began. As previously explained, ten or fewer URM students were within each condition at post-intervention. Therefore, the results from the prior study still need to be explicated. One possible explanation for the fewer than expected URM students in the sample is that the current study utilized Algebra I students who were mainly in the eighth grade. As

Algebra I is commonly taken in the eighth or ninth grade, sampling Algebra I students at the high school level as well may provide a more representative sample of Algebra I students in general and may also result in a greater number of URM students within the sample.

Although the proposed research question could not be answered, descriptive statistics revealed that, as hypothesized, URM students within this sample did indeed have a lowered sense of belonging to mathematics at the start of the intervention. This is a small but unique contribution to the field as URM students' sense of belonging to mathematics in particular has not been previously explored. Good and colleagues (2012) suggest that stigmatized individuals in the domain of mathematics may have a lowered sense of belonging to math but did not assess this in URM students. Although beyond the scope of the current study, this is a potential area of exploration for future research. Assessing the development of URM students' sense of belonging to mathematics, how it differs from non-URM students' development of sense of belonging to mathematics, and whether it is a function of URM status in particular or can be better explained by factors such as prior knowledge and socioeconomic status are all questions that can and should be answered in future research.

URM students also demonstrated lower pre-test procedural scores than non-URM students. However, pre-test conceptual scores did not differ between URM and non-URM students. This is a particularly perplexing finding. Upon further inspection of the data, one practical explanation of this finding is that URM students were predominantly enrolled in the same classrooms. Across the four teachers participating in the study, 76.2% of the URM students in the sample were enrolled in two of the four teachers' classrooms that had low procedural knowledge at pre-test overall. One possibility is that these two particular teachers did not focus as much on developing procedural skills as much as the teachers of the two classes that had higher procedural scores at pre-test. If this is indeed the case, it may be that the differences between classrooms at pre-test and the disproportionate number of URM students within the lower performing classrooms at pre-test may be driving this effect as opposed to the finding being relevant to URM students in particular. A multi-level model on a larger data set controlling for proportion URM at the classroom level may help to elucidate these findings. However, there are three reasons why this does not explain the findings of the prior study on which the current study was based, in which differential effects of condition were found by URM status: 1) students were assigned to condition within classroom, 2) URM students were equally distributed across conditions, and 3) no differences were found at pre-test by condition. Therefore, further exploration is needed to better understand the effects of error reflection on learning within consideration to individual differences.

Summary of Limitations and Recommendations for Future Research

The prior discussion revealed several limitations of the current study which will be summarized here along with recommendations for future directions. The greatest challenge of interpreting the results from the current study stems from complications in studying nested data. Outcome measures differed across the classes that participated in the study. However, due to the few participating clusters, this did not allow for multi-level modeling. Not accounting for the non-independence in the data can lead to smaller, somewhat biased standard errors that may not be accurate which may also lead the researcher to committing a Type I error. However, it is important to note that the participants in the current study were randomly assigned within classroom to one of the four conditions. Participants did not vary systematically on any of the observed variables at pre-test by condition. Some classes or clusters as a whole began the study with lower pre-test scores than others. However, all models controlled for pre-intervention measures of learning and sense of belonging to mathematics. Therefore, one taking a more liberal approach to interpreting the effects of condition might argue that the experimental design of the study and the ability to control for pre-test scores gives one little reason to believe that the participants varied systematically across conditions. To err on the side of caution, the author of the current study opted for a conservative approach to account for the clustering in the data. Design effect adjusted standard errors were applied as recommended by Huang (2015) and conclusions made are based upon these adjusted standard errors. Future in-vivo classroom research exploring the effects of an error reflection intervention should be sure to sample an adequate number of classes to be able to utilize multi-level modeling in the interpretation of the data.

Another limitation of the current study is related to sampling. One of the research questions posed asked whether there were differential effects of condition on learning and sense of belonging to math for URM students. This question was posed in order to clarify findings from a prior study that suggested a particular benefit of studying incorrect worked examples for URM students. Unfortunately, although an adequate number of URM students were expected after discussion with the school district, this did not hold true during data collection. This prevented the planned analyses from being conducted for this particular question. Future studies to address this question will need to attempt an oversampling of participants from schools with a high proportion of URM students.

The current study attempted to create an intervention to modify middle grade students' sense of belonging to mathematics. However, little is known about this particular construct. Although sense of school and social belonging has been thoroughly studied, sense of belonging to the domain of mathematics is a relatively new and understudied construct. Thus, research is needed on the development of this construct in general as well as whether and how this development differs for different groups of students (e.g., underrepresented groups, students at varying achievement levels, etc.). Better understanding how sense of belonging to mathematics develops may allow one to design a more effective intervention to help promote stronger feelings of belonging in this domain.

There are a few limitations in the design of the current study that make interpretation difficult. One assumption that was made but not tested was that students' views of the academic domain of mathematics would relate to their sense of belonging to mathematics and their perceived functionality of errors. In particular, it was hypothesized that participating in one of the error reflection intervention would alter students' perceptions of the functionality of errors which would in turn lead them to viewing math as a more inclusive domain in which anyone could belong. However, as previously noted, the current study did not attempt to measure students' perceptions of the mathematics domain. Measuring this along with students' sense of belonging to mathematics and their perceived functionality of errors would be more informative and allow a fuller answer to the research questions posed.

Lastly, after adjusting standard errors, the effects of condition were no longer significant. It is unclear why this might be the case. One possibility is that students were not fully engaged in the error reflection interventions or not utilizing them in the intended way. However, the current data collection procedures do not allow for exploring this possibility. Future studies to clarify these findings would benefit from including some measure of engagement as well as collection of process data. Think-alouds may be particularly useful in understanding how learners work with the designed materials.

Conclusions

The findings of the current study suggest that reflecting on and explaining errors in algebra worked examples is no more or less effective at promoting learning of conceptual and procedural algebraic knowledge than traditional problem solving alone or working with traditional correct worked examples and written self-explanation prompts. Each of these instructional interventions seems to result in student learning over the course of a unit of study. In addition, students' prior knowledge does not seem to moderate the effect of the different forms of instructional interventions on learning.

Further, no changes were seen in students' sense of belonging to mathematics before and after the intervention, across all conditions. It is possible that the error reflection intervention may not have been rigorous enough to lead to substantial changes in sense of belonging to mathematics. It is also possible that a more socially-focused intervention may be more effective at altering sense of belonging to mathematics than a cognitive one.

The current study was the first to address the relationship between students' perceived functionality of errors and learning. Students' perceptions of the functionality of errors were high at pre-intervention and remained stable across all conditions and throughout the intervention. However, these perceptions were unrelated to students' algebra knowledge and sense of belonging to math either at pre- or post-test. Not surprisingly then, perceived functionality of errors did not mediate the effect of condition on these outcome measures.

Although the current study was proposed to have a particular focus on potential differences in the effects of condition on learning and sense of belonging to math by URM status, the current sample size did not allow for these analyses to be conducted. The main finding in relation to URM students based on possible analyses is that URM students do indeed have a lowered sense of belonging to mathematics than non-URM students.

Educational implications.

The current study sought ways to increase learning opportunities for Algebra I students that may simultaneously foster improvements in students' motivation. The current study is one of the first to test methods for altering middle school students' sense of belonging to the mathematics domain as well as students' perceptions of the functionality of errors for learning. The current study used what is known about cognitive science as a basis to design an even more effective intervention than the learning tools frequently used in middle school math classrooms. The findings suggest that what has worked consistently in a laboratory setting may not as easily translate to the classroom. Although students in the current study demonstrated considerable learning gains regardless of condition, there was still much room for improvement. Further research and modifications may be needed in order to integrate and fully utilize findings from cognitive science to create learning tools for use in the classroom.

Though various types of achievement motivation orientations have been shown to decline during adolescence, it seems as though middle grade algebra students have a relatively strong sense of belonging to mathematics. Even though URM students seemed to report a lowered sense of belonging to mathematics than non-URM students, this was still relatively high. Researchers and educators alike may desire to seek ways to maintain this high sense of belonging, as it has been previously found to predict math learning and intent to pursue more difficult math courses. It is also vital to further explore URM students' sense of belonging in particular, as this may have important implications for practice in relation to the well-noted achievement gap in the US. Although math achievement and math motivation are closely linked, implementing separate classroom interventions for learning and motivation may be more effective albeit time consuming on behalf of classroom teachers. Still, future research may reveal a more comprehensive approach that could potentially increase learning and motivation simultaneously.

The two different error reflection interventions designed and tested within the current study are at least as effective as traditional problem-solving or working with correct worked examples, making it a viable option for classroom instruction. Future research should explore potential improvements for classroom use.

REFERENCES

Adams, D.M, McLaren, B.M., Durkin, K., Mayer, R.E., Rittle-Johnson, B., Isotani, S., & van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior, 36*, 401-411.

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

Anderman, L.H. & Anderman, E.M. (1999). Social predictors of changes in students' achievement goal orientations. *Contemporary Educational Psychology, 24*, 21-37.

Anderman, L.H. (2003). Academic and social perceptions as predictors of change in middle school students' sense of school belonging. *The Journal of Experimental Education, 72*(1), 5-22.

Aronson, J., Fried, C., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113-125.

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

Baddeley, A.D. & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York: Academic Press.

Barbieri, C. & Booth, J.L. (under review). Compensating for competence expectancy: Contributions of incorrect worked examples.

Baumeister, R.F. & Leary, M.R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497-529.

Berthold, K., Eysink, T.H.S., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science, 37*, 345-363.

Blackwell, L.S., Trzesniewski, K.H. & Dweck, & C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246-263.

Booth, J.L. (2011). Why can't students get the concept of math? *Perspectives on Language and Literacy, Spring, 2011*, 31-35.

Booth, J.L., Barbieri, C., Eyer, F., & Pare-Blagoev, E.J. (2014). Persistent and pernicious errors in algebraic problem solving. *Journal of Problem Solving*, *7*(1), 10-23.

Booth, J.L., Lange, K.E., Koedinger, K.R., & Newton, K.J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning & Instruction, 25*, 24-34.

Booth, J.L., Oyer, M.H., Paré-Blagoev, E.J., Elliot, A., Barbieri, C., Augustine, A.A., & Koedinger, K.R. (in press). Learning algebra by example in real-world classrooms. *Journal of Research on Educational Effectiveness.*

Borasi, R. (1994). Capitalizing on errors as "springboards for inquiry": A teaching experiment. *Journal for Research in Mathematics Education*, *25,* 166-208.

Byrne, B.M. (1996). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B.A. Bracken (Ed.), *Handbook of self-concept* (pp. 287-316). New York: Wiley.

Byrnes, J.P. (2003). Factors predictive of mathematics achievement in White, Black, and Hispanic 12th graders. *Journal of Educational Psychology, 95* (2), 316-326.

Brooks-Gunn, J. & Markman, L.B. (2005). The contribution of parenting to ethnic and racial gaps in school readiness. *Future Child, 15*(1), 139-168.

Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S.L., McLoyd, V., & Pianta, R. (2011). Examining the Black-White achievement gap among low-income children using the NICHD study of early child care and youth development. *Child Development*, *82*(5), 1404-1420.

Carroll, W.M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology, 86*(3), 360-367.

Chen, X. (2009). Students who study Science, Technology, Engineering, and Mathematics (STEM) in postsecondary education (NCES 2009-161). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Cherepinsky, V. (2011). Self-reflective grading: Getting students to learn from their mistakes. *PRIMUS, 21*(3), 294-301.

Chi, M.T.H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology, (*pp. 161-238). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

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

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

Clifford, M.H., Kim, A., & McDonald, B.A. (1988). Responses to failure as influenced by task attribution, outcome attribution and failure tolerance. *Journal of Experimental Education, 57*(1), 19-37.

Cohen, G.L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*, 1307-1310.

Cohen, G.L., Garcia, J., Purdie-Vaugns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*, 400-403.

Darling-Hammond, L. (2004). The color line in American education: Race, resources, and student achievement. *W.E.B. DuBois Review: Social Science Research on Race, 1(2),* 213-246.

Darling-Hammond, L. (2006). Securing the right to learn: Policy and practice for powerful teaching and learning. *Educational Researcher, 35*(7), 13-24.

Dickens, W.T. (2005). Genetic differences in school readiness. *The Future of Children, 15,* 55-69.

Duncan, G.J. & Magnuson, K.A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, 15, 35-54.

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

Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality and development.* Philadelphia: Taylor and Francis/Psychology Press.

Eccles, J., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon & N. Eisenberger (Eds.), *Handbook of Child Psychology* (5th ed., Vol 3, pp. 1017-1095). New York: John Wiley.

Falco, L.D., Summers, J.J., & Bauman, S. (2011). Encouring mathematics participation through improved self-efficacy: a school counseling outcomes study. *Educational Research and Evaluation: An International Journal on Theory and Practice, 16*(6), 529-549.

Fine, M. (1991). *Framing dropouts.* Albany, NY: SUNY Press.

Fryer, R.G. & Levitt, S.D. (2004). Understanding the black-white test score gap in the first two years of school. The Review of Economics and Statistcs, 86, 447-464.

García Coll, C., Crnic, K., Lamberty, G., Wasik, B.H., Jenkins, R., Vázquez Garcia, H., & McAdoo, H.P. (1996). Toward an integrative theoretical model for the study of developmental competencies in minority children. *Child Development, 67*, 1891-1914.

Good, C., Rattan, A. & Dweck, C.S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700-717.

Goodenow, C. & Grady, K.E. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *Journal of Experimental Education, 62*, 60-71.

Goodenow, C. (1992). Strengthening the links between educational psychology and the study of social contexts. *Educational Psychologist, 27*, 177-196.

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *The Journal of Early Adolescence, 13*(1), 21-43.

Goska, R. E., & Ackerman, P. L. (1996). An aptitude-treatment interaction approach to transfer within training. *Journal of Educational Psychology, 88,* 249-259.

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

Gutman, L.M. & Midgely, C. (2000). The role of protective factors in supporting the academic achievement of poor African-American students during the middle school transition. *Journal of Youth and Adolescence, 29,* 223-248.

Hahs-Vaughn, D.L. (2005). A primer for using and understanding weights with national datasets. *The Journal of Experimental Education, 73*, 221-238.

Huang, F.L. (2014). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education, 0*(0), 1-22.

Heemsoth, T. & Heinze, A. (2014). The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instructional Science, 42*(4), 639-657.

Hemphill, F. C., and Vanneman, A. (2010). *Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress* (NCES 2011-459). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Henderson, C. & Harper, K.A. (2009). Quiz corrections: Improving learning by encouraging students to reflect on their mistakes. *Physics Teacher, 47*(9), 581-586.

Herrnstein, R.J. & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

Kalyuga, S. (2011). Cognitive Load Theory: How many types of load does it really need? *Educational Psychology Review, 23*, 1-19.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23-31.

Kalyuga, S., Chandler, P. & Sweller, J. (2001). Learner experience and efficiency of instructional guidance. *Educational Psychology, 21*(1), 5-23.

Kelly, D., Xie, H., Nord, C.W., Jenkins, F., Chan, J.Y., and Kastberg, D. (2013). Performance of U.S. 15-Year-Old Students in Mathematics, Science, and Reading Literacy in an International Context: First Look at PISA 2012 (NCES 2014-024). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academies Press.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd Ed). *New York: Guilford Press*.

Kozol, J. (2005). *The shame of the nation: The restoration of apartheid schooling in America.* New York: Crown Publishing.

Landivar, L.C. *(*2013)*. Disparities in STEM employment by sex, race, and Hispanic origin*. Retrieved from United States Census Bureau website: http://www.census.gov/prod/2013pubs/acs-24.pdf

Lee, E., Forthofer, R., & Lorimor, R. (1989). *Analyzing complex survey data.* Newbury Park, CA: Sage.

Leventhal, T. & Brooks-Gunn, J. (2000). The neighborhoods they live in: Effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, *126*, 309-337.

MacKinnon, D.P., Lockwood, C.M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39,* 99-128.

Matthews, M.S. & Farmer, J.L. (2008). Factors affecting the Algebra I achievement of academically talented learners. *Journal of Advanced Academics, 19*(3): 472-501.

Mau, W.C. (2003). Factors that influence persistence in science and engineering career aspirations. *The Career Development Quarterly, 51*, 234-243.

Meyer, D. & Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review, 18*(4), 377-390.

Middleton J.A. & Spanias, P.A. (1999). Motivation for achievement in mathematics: Findings, generalizations, and criticisms of the research. *Journal for Research in Mathematics Education, 30*(1), 65-88.

Morrison, G.M., Cosden, M.A., O'Farrel, S.L., Campos, E. (2003). Changes in Latino students' perceptions of school belonging over time: Impact of language proficiency, self-perceptions, and teacher evaluations. *California School Psychologist, 8*, 87-98.

Muthén, L.K., & Muthén, B.O. (1998-2014). *Mplus User's Guide* (7th Ed). Los Angeles, CA: Muthén & Muthén.

National Center for Education Statistics (2013). *The nation's report card: A first look: 2013 mathematics and reading* (NCES 2014-451). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics.* Washington, DC: Authors.

National Science Foundation, National Center for Science and Engineering Statistics. (2014). *Doctorate Recipients from U.S. Universities: 2012*. Special Report NSF 14-305. Arlington, VA. Available at http://www.nsf.gov/statistics/ sed/2012/.

Oakes, J. (2004). Investigating the claims in *Williams v. State of California:* An unconstitutional denial of education's basic tools? *Teachers College Record, 106*(10), 1889-1906.

Ogbu, J.U. (2003). Black-White differences in achievement: The importance of wealth. *Sociology of Education, 76*(4), 281-304.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review, 103*(2), 241-262.

Oser, F., & Spychiger, M. (2005). Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur [Learning is painful. On the theory of negative knowledge and the practice of error culture]. Weinheim, Germany: Beltz.

Osterman, K.F. (2000). Students' need for belonging in the school community. *Review of Educational Research, 70*(3), 323-367.

Pajares, F. & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology*, *24*, 124-139.

Pickett, C.L., Gardner, W.L. & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality and Social Psychology Bulletin*, *30*, 1095-1107.

Pol, H.J., Harskamp, E.G., Suhre, C.J.M. & Goedhart, M.J. (2009). How indirect supportive digital help during and after solving physics problems can improve problem-solving abilities. *Computers & Education, 53*, 34-50.

Potts, R. & Shanks, D.R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*, 644-667.

Rattan, A., Good, C., & Dweck, C.S. (2012). "It's ok - Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*, 731-737.

Reardon, S.F. (May 2013). The widening income achievement gap. *Educational Leadership*, 10-16.

Reeves, E.B. (2012). The effects of opportunity to learn, family socioeconomic status, and friends on the rural math achievement gap in high school. *American Behavioral Scientist, 56(7)*, 887-907.

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

Renkl, A., Atkinson, R.K., Maier, U.H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education, 70*, 293-315.

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

Rittle-Johnson, B. & 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. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1-15.

Rodkin, P.C.; Ryan, A.M. (2012). Child and adolescent peer relations in educational context. In Harris, K.R., Graham, S., Urdan, T., Graham, S., Royer, J.M., Zeidner, M. (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors. APA handbooks in psychology*, (pp. 363-389). Washington, DC, US: American Psychological Association.

Roy, M. & Chi, M.T.H. (2005). Self-explanation in a multi-media context. In R. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning* (pp. 271-286). New York: Cambridge.

Ryan, R.M. & Deci, E.L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E.L. Deci & R.M. Ryan (Eds.), *Handbook of self-determination theory research* (pp. 3-33). Rochester, NY: University of Rochester Press.

Santagata, R. (2004). “Are you joking or are you sleeping?” Cultural beliefs and practices in Italian and U.S. teachers’ mistake-handling strategies. *Linguistics and Education*, 15, 141-164.

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

Schleppenbach, M., Flevares, L. M., Sims, L. M., & Perry, M. (2007). Teachers' responses to student mistakes in Chinese and US mathematics classrooms. *Elementary School Journal, 108*(2), 131-147.

Schneider, B., Swanson, C.B., & Riegle-Crumb, C. (1998). Opportunities for learning: Course sequences and positional advantages. *Social Psychology of Education, 2*(1), 25-53.

Seifert, T.L. (1995). Academic goals and emotions: A test of two models. *Journal of Psychology: Interdisciplinary and Applied, 129*, 543-552.

Siegler, R. S. (2002). Microgenetic studies of self-explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31-58). New York: Cambridge University Press.

Siegler, R.S. & Chen, Z. (2008). Differentiation and integration: guiding principles for analyzing cognitive change. *Developmental Science, 11*(4), 433 – 453.

Skinner, B.F. (1961). Why we need teaching machines. Harvard Educational Review, 31, 377-398.

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.

Star, J.R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*, 404-411.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Steele, C.M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629.

Steuer, G., Rosentritt-Brunn, G., Dresel, M. (2013). Dealing with errors in mathematics classrooms: Structure and relevance of perceived error climate. *Contemporary Educational Psychology, 38*, 196-210.

Stevens, T., Hamman, D., Olivárez Jr., A. (2007). Hispanic students' perception of White teachers' mastery goal orientation influences sense of school belonging. *Journal of Latinos and Education, 6*(1), 55-70.

Stevenson, H. & Stigler, J. (1992). *The Learning Gap.* New York: Summit Books.

Stevenson, H.W., Chen, C. & Uttal, D.H. (1990). Beliefs and achievement: A study of black, white, and Hispanic children. *Child Development, 61*, 508-523.

Stigler, J.W. & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.

Stigler, J.W. & Perry, M. (1988). Mathematics learning in Japanese, Chinese, and American classrooms. In G.B. Saxe & M. Gearhart (Eds.), *Children's mathematics: New directions for child development*, No. 41 (pp. 27 – 54.) San Francisco: Jossey-Bass.

Stipek, D. (2002). Good instruction is motivating. In A. Wigfield & J.S. Eccles (Eds.), *Development of achievement motivation. A volume in the educational psychology series.* San Diego, CA: Academic Press.

Sweller, J. & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59-89.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*, 165-169.

Sweller, J. (2012). Human cognitive architecture: Why some instructional procedures work and others do not. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook; Volume 1: Theories, Constructs, and Critical Issues.* Washington DC: American Psychological Association.

Tjosvold, D., Yu, Z. & Hui, C. (2004). Team learning from mistakes: The contribution of cooperative goals and problem-solving. *Journal of Management Studies, 41*, 1223-1245.

Tulis, M. & Ainley, M. (2011). Interest, enjoyment and pride after failure experiences? Predictors of students' state-emotions after success and failure during learning in mathematics. *Educational Psyc*http://www.census.gov/topics.html*hology*, *31*(7), 779-807.

Tulis, M. (2013). Error management behavior in classrooms: Teachers' responses to student mistakes. *Teaching and Teacher Education*, *33*, 56-68.

United States Census Bureau/American FactFinder, "DP05: ACS Demographic and Housing Estimates". 2011 – 2013 American Community Survey, U.S. Census Bureau's American Community Survey Office, 2013. Web. <http://factfinder2.census.gov>.

Van Dyck, C., Frese, M., Baer, M. & Sonnentag, S. (2005). Organizational error management culture and its impact on performance: A two-study replication. *Journal of Applied Psychology, 90*, 1228-1240.

van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive Load Theory: Advances in research on worked examples, animations, and cognitive load measurement. *Educational Psychology Review*, *22*, 375 - 378.

Vanneman, A., Hamilton, L., Baldwin Anderson, J., and Rahman, T. (2009). Achievement Gaps: *How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress,* (NCES 2009-455). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Walton, G.M. & Cohen, G.L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*, 82-96.

Walton, G.M. & Cohen, G.L. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science, 331*, 1447-1451.

Wang, Q., & Pomerantz, E. (2009). The motivational landscape of early adolescence in the United States and China: A longitudinal investigation. *Child Development, 80*, 1272-1287.

Wang, T. & Murphy, J. (2004). An examination of coherence in a Chinese mathematics classroom. In L. Fan, N.-Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 107-123). Hackensack, NJ: World Scientific.

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal, 50*(5), 1081-1121.

Ward, M. & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7*, 1-39.

Wentzel, K.R. & Wigfield, A. (Eds.). (2009). *Handbook of motivation at school.* New York, NY: Routledge.

Yerushalmi, E. & Polingher, C. (2006). Guiding students to learn from mistakes. *Physics Education, 41*(6), 532-538.

Zhu, X. & Simon, H.A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4,* 137-166.

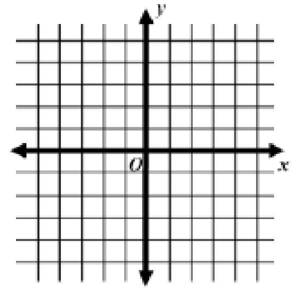
APPENDIX A: QUADRATICS TEST

Quadratics Pretest

Procedural

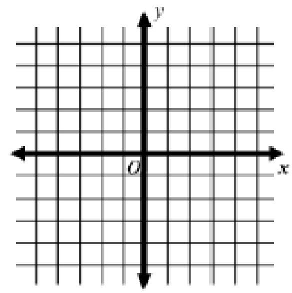
1. Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. Then graph the function.
   $y = x^2 - 2x - 3$

Axis of Symmetry	
Vertex	
Min or Max?	



Procedural

2. The point $(4,3)$ lies on the graph of a quadratic function whose axis of symmetry is $x = 2$. Find another point on the graph. Explain how you found the point.



3. State whether each of the following is true for the quadratic function $y = -x^2 - 2x + 3$.

Conceptual

a.	The axis of symmetry is $x = 1$.	Yes	No
b.	The axis of symmetry is $x = -1$.	Yes	No
c.	The vertex is a minimum.	Yes	No
d.	The vertex is $(-1,4)$.	Yes	No
e.	The vertex is $(-1,0)$.	Yes	No
f.	The vertex is $(1,4)$.	Yes	No
g.	The vertex is a maximum.	Yes	No
h.	The vertex is $(4,-1)$.	Yes	No

Procedural

4. Solve the quadratic equation by factoring. Show all of your work.
   $0 = x^2 - 6x - 27$

Procedural

5. Solve the quadratic by factoring. Show all of your work
   . $0 = x^2 - 20x + 91$

Conceptual

6. State whether the following can be solved by factoring <u>with whole number solutions</u>.

a.	$0 = x^2 + 7x + 12$	Yes	No
b.	$0 = x^2 + 10x + 20$	Yes	No
c.	$0 = x^2 - 6 - 27$	Yes	No
d.	$0 = x^2 + 6x + 10$	Yes	No
e.	$0 = x^2 - 7x$	Yes	No

Procedural

7. Solve the quadratic by using the square root. Show all of your work.
   $(x + 2)^2 = 0$

Procedural

8. Solve the quadratic by using the square root. Show all of your work.
   $(x+25)^2 = 100$

Conceptual

9. State whether the following quadratics will have one, two, or zero solutions.

a.	$(x+4)^2 = 25$	One	Two	No Solution
b.	$(x+25)^2 = 0$	One	Two	No Solution
c.	$(x-3)^2 = 9$	One	Two	No Solution
d.	$(x+6)^2 = -4$	One	Two	No Solution
e.	$(x-5)^2 = -16$	One	Two	No Solution

Procedural

10. Solve the equation using the quadratic formula.
    $2x^2 - 3x + 4 = 0$

Procedural

11. A ball falls off of a ledge. The equation $0 = -2t^2 + 10t + 0$ gives the time t in seconds when the ball hits the ground (height of 0 feet). Use the quadratic equation to find out how many seconds it took for the ball to hit the ground.

APPENDIX A (continued)

Conceptual

12. **For each example (I – V), circle the correct next step (a or b) for solving using the quadratic formula**

I) $2x^2 - 3x + 2 = 0$

Step 1: $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Is Step 2 a or b?

a. $x = \dfrac{-3 \pm \sqrt{(3)^2 - 4(2)(2)}}{2(2)}$

b. $x = \dfrac{3 \pm \sqrt{(-3)^2 - 4(2)(2)}}{2(2)}$

Conceptual

II) $3x^2 - 2x + 5 = 0$

Step 1: $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Step 2: $x = \dfrac{2 \pm \sqrt{(-2)^2 - 4(3)(5)}}{2(3)}$

Is Step 3 a or b?

a. $x = \dfrac{2 \pm \sqrt{4 - 60}}{6}$

b. $x = \dfrac{2 \pm \sqrt{-4 - 60}}{6}$

Conceptual

III) $-x^2 + 4x + 2 = 0$

Step 1: $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Step 2: $x = \dfrac{-4 \pm \sqrt{(4)^2 - 4(-1)(2)}}{2(-1)}$

Is Step 3 a or b?

a. $x = \dfrac{-4 \pm \sqrt{16 + 8}}{-2}$

b. $x = \dfrac{-4 \pm \sqrt{16 - 8}}{-2}$

APPENDIX A (continued)

Conceptual

IV) $2x^2 + 3x + 1 = 0$

Step 1: $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Step 2: $x = \dfrac{-3 \pm \sqrt{(3)^2 - 4(2)(1)}}{2(2)}$

Step 3: $x = \dfrac{-3 \pm \sqrt{9-8}}{4}$

Step 4: $x = \dfrac{-3 \pm \sqrt{1}}{4}$

Is Step 5 a or b?

a. $x = \dfrac{-3+\sqrt{1}}{4}$ or $x = \dfrac{-3-\sqrt{1}}{4}$

b. $x = \dfrac{-3+\sqrt{1}}{4}$

Conceptual

V) $x^2 - 4x + 1 = 0$

Step 1: $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Step 2: $x = \dfrac{4 \pm \sqrt{(-4)^2 - 4(1)(1)}}{2(1)}$

Step 3: $x = \dfrac{4 \pm \sqrt{16-4}}{2}$

Step 4: $x = \dfrac{4 \pm \sqrt{12}}{2}$

Step 5: $x = \dfrac{4+\sqrt{12}}{2}$ or $x = \dfrac{4-\sqrt{12}}{2}$

Is Step 6 a or b?

a. $x = 8$ or $x = -4$

b. $x = 3.73$ or $x = 0.27$

APPENDIX B: SURVEY A

Your Name: ______________________ **Date:** _____________ **Teacher:** ________________ **Section:** ___________

Survey A.1

These questions are about your algebra class and math more generally. Your participation may help us understand and create materials that will help other students learn algebra better, so staying focused is very important. There are two different sections, please read each set of instructions carefully. There are no right or wrong answers.

Section 1

Please circle an answer for each question based on how you are thinking and feeling about algebra and math RIGHT NOW. Some of the questions are very similar. Please read each question carefully and answer honestly about how you are thinking and feeling at the moment.

1. **How good at math are you?**

   1 2 3 4 5 6 7

   Not at all good ... Very good

2. **In general, how useful is what you learn in math?**

   1 2 3 4 5 6 7

   Not at all useful ... Very useful

3. **If you were to list all of the students in your class from worst to best in math, where would you put yourself?**

   1 2 3 4 5 6 7

   One of the worst ... One of the best

4. **Compared to most of your other activities, how much do you like math?**

   1 2 3 4 5 6 7

   Not as much ... A lot more

5. **Compared to most of your other school subjects, how good are you at math?**

   1 2 3 4 5 6 7

   A lot worse ... A lot better

6. **For me being good at math is . . . .**

   1 2 3 4 5 6 7

   Not at all important ... Very important

7. **How well do you expect to do in math this year?**

   1 2 3 4 5 6 7

   Not at all well ... Very well

8. **In general, I find working on math assignments . . . .**

   1 2 3 4 5 6 7

   Very boring ... Very interesting

Continued on the other side →

APPENDIX B (continued)

9. How good would you be at learning something new in math?

1	2	3	4	5	6	7
Not at all good						Very good

10. Compared to most of your other activities, how important is it to you to be good at math?

1	2	3	4	5	6	7
Not as important						A lot more important

11. How much do you like doing math?

1	2	3	4	5	6	7
A little						A lot

Section 2

These questions are about math "communities." What do we mean by a community? We mean people who feel like they belong together as a group. Your *school* community includes all the people in the *school*. A *math* community includes all people who feel that they belong with others who care about *math*. This includes other math students and math teachers, and others outside the school who care about mathematics. Please respond to these items based on how you feel about belonging in a math community *when you are in math class*.

When I am in math class...	**Strongly DISAGREE**					**Strongly AGREE**
1. I feel like I am part of the math community.	1	2	3	4	5	6
2. I feel accepted.	1	2	3	4	5	6
3. I feel comfortable.	1	2	3	4	5	6
4. I trust my teachers to help me learn.	1	2	3	4	5	6
5. I enjoy participating.	1	2	3	4	5	6

Today's Date is: ____________________

APPENDIX C: SURVEY B

Date: ____________ Teacher: ______________ Section: __________ ID Code: ____________
(not your name!)

Survey B.1

These questions are about your algebra class and math more generally. Your participation may help us understand and create materials that will help other students learn algebra better, so staying focused is very important. There are two different sections, please read each set of instructions carefully. There are no right or wrong answers. Some of the questions are very similar. Please read each question carefully and answer honestly about how you are thinking and feeling at the moment.

<u>Section 1</u>

Please circle an answer for each question based on how you are thinking and feeling about your algebra class RIGHT NOW.

1. **In our Math class the mistakes students make are often used to make sure you really understand Math.**

   1 (Strongly Disagree) 2 3 4 5 (Strongly Agree)

2. **In our Math class we learn a lot from assignments that were *not* done correctly.**

   1 (Strongly Disagree) 2 3 4 5 (Strongly Agree)

3. **In our Math class wrong answers on assignments are used to learn something.**

   1 (Strongly Disagree) 2 3 4 5 (Strongly Agree)

4. **In our Math class wrong answers are often a good opportunity to really understand the material.**

   1 (Strongly Disagree) 2 3 4 5 (Strongly Agree)

<u>Section 2</u>

Please circle an answer for each question based on how you are thinking about algebra and math RIGHT NOW.

1. **You have a certain amount of math ability, and you can't really do much to change it.**

   1 (Strongly disagree) 2 3 4 5 6 7 (Strongly agree)

2. **Your math ability is something about you that you can't change very much.**

   1 (Strongly disagree) 2 3 4 5 6 7 (Strongly agree)

3. **To be honest, you can't really change your ability in math.**

   1 (Strongly disagree) 2 3 4 5 6 7 (Strongly agree)

4. **You can learn new things, but you can't really change your basic math ability.**

   1 (Strongly disagree) 2 3 4 5 6 7 (Strongly agree)

Today's Date is: ________________

APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS

ID code: ____________ Date: ____________ Teacher: ____________ Section: ____________
(not your name!)

quadratic formula

SET 1 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

1a. $w^2 + 6w + 8 = 0$

1b. $w^2 + 2w - 8 = 0$

more
quadratic formula

Set 2 Solve each equation using the quadratic formula. SHOW ALL OF YOUR WORK.

2a. $4w^2 - 4w = -1$

2b. $9w^2 + 12w = -4$

APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS (continued)

more
quadratic formula

SET 3 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

3a. $-5 = x^2 + 5x$

3b. $1 + 3x^2 = -5x$

ID code: ______________ Date: __________ Teacher: ______________ Section: __________
(not your name)

solving quadratics by factoring

SET 1 Solve each equation by **factoring.** SHOW ALL OF YOUR WORK.

1a. $x^2 - 3x = 0$

1b. $x^2 + 25x = 0$

APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS
(continued)

more
solving quadratics by factoring

SET 2 Solve each equation by **factoring.** SHOW ALL OF YOUR WORK.

2a. $x^2 + 8x - 48 = 0$

2b. $x^2 + x - 12 = 0$

more
solving quadratics by factoring

SET 3 Solve each equation by **factoring.** SHOW ALL OF YOUR WORK.

3a. $x^2 + 9x + 8 = 0$

3b. $x^2 + 3x - 28 = 0$

ID code: ______________ (not your name) Date: __________ Teacher: ______________ Section: __________

solving quadratics using the square root

SET 1 Solve each equation by using the square root. SHOW ALL OF YOUR WORK.

1a. $(n+2)^2 = 9$

1b. $(n-5)^2 = 100$

APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS (continued)

more
solving quadratics using the square root

SET 2 Solve each equation by using the square root. SHOW ALL OF YOUR WORK.

2a. $(n-6)^2 = 0$

2b. $(n+25)^2 = 0$

APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS (continued)

more

solving quadratics using the square root

SET 3 Solve each equation by <u>using the square root</u>. SHOW ALL OF YOUR WORK.

3a. $(n+3)^2 = -7$

3b. $(n-3)^2 = -4$

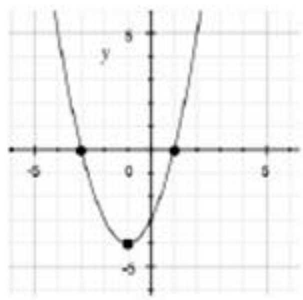
APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS
(continued)

ID code: ______________ (not your name) Date: __________ Teacher: ______________ Section: __________

graphing quadratic functions

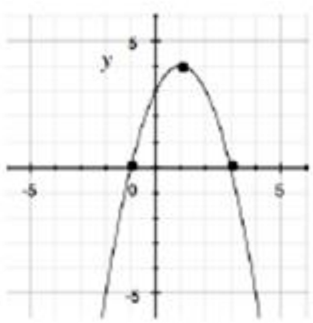
SET 1 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value.

1a.



Axis of Symmetry	
Vertex	
Min or Max?	

1b.



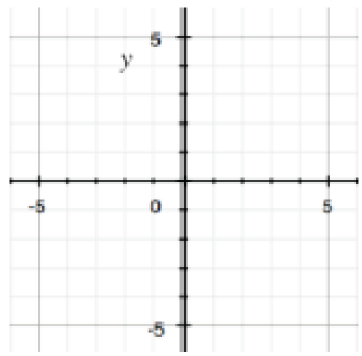
Axis of Symmetry	
Vertex	
Min or Max?	

more
graphing quadratic functions

SET 2 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. **Then graph the function.**

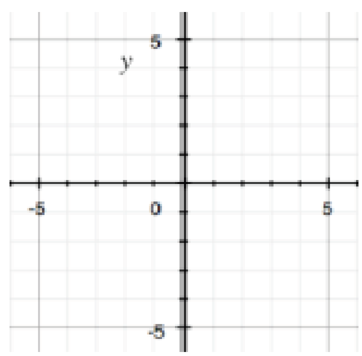
2a. $y = -x^2 + 4x + 1$

Axis of Symmetry	
Vertex	
Min or Max?	



2b. $y = x^2 - 4x + 1$

Axis of Symmetry	
Vertex	
Min or Max?	



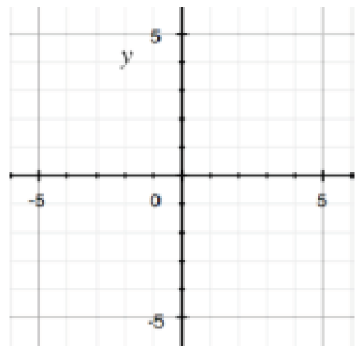
APPENDIX D: PROBLEM-SOLVING CONTROL WORKSHEETS (continued)

more
graphing quadratic functions

SET 3 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. **Then graph the function.**

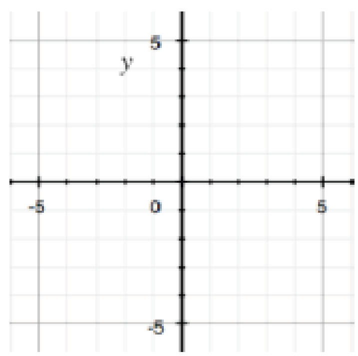
3a. $y = -x^2 - 2x - 1$

Axis of Symmetry	
Vertex	
Min or Max?	

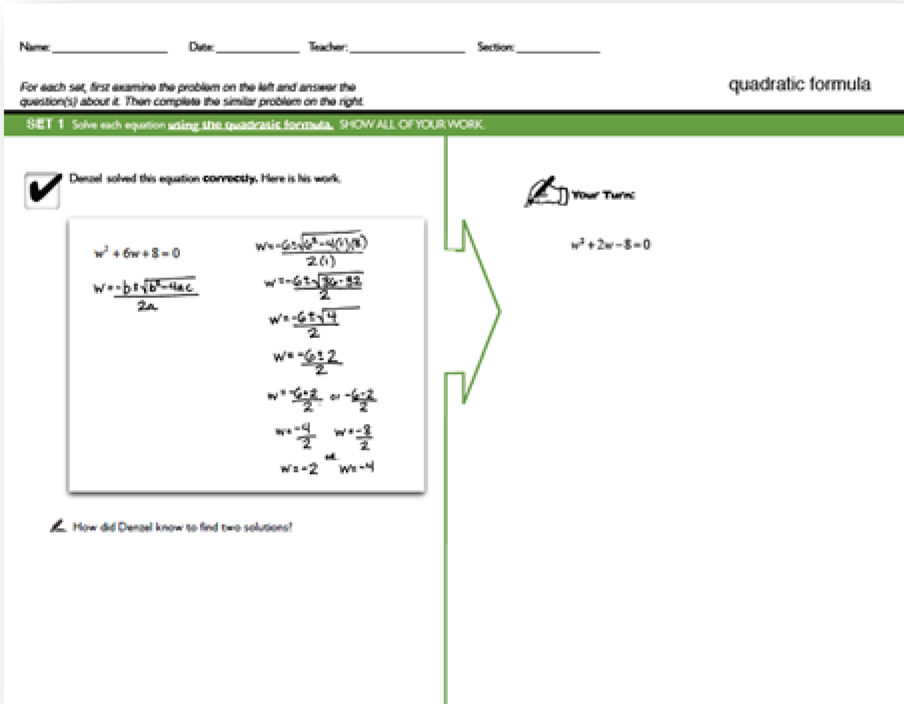


3b. $y = x^2 - 2x - 1$

Axis of Symmetry	
Vertex	
Min or Max?	



APPENDIX E: CORRECT EXAMPLE CONTROL WORKSHEETS



Name: ______________ Date: __________ Teacher: ______________ Section: ___________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

quadratic formula

SET 1 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

Denzel solved this equation **correctly.** Here is his work.

$w^2 + 6w + 8 = 0$

$w = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$w = \frac{-6 \pm \sqrt{6^2 - 4(1)(8)}}{2(1)}$

$w = \frac{-6 \pm \sqrt{36 - 32}}{2}$

$w = \frac{-6 \pm \sqrt{4}}{2}$

$w = \frac{-6 \pm 2}{2}$

$w = \frac{-6+2}{2}$ or $\frac{-6-2}{2}$

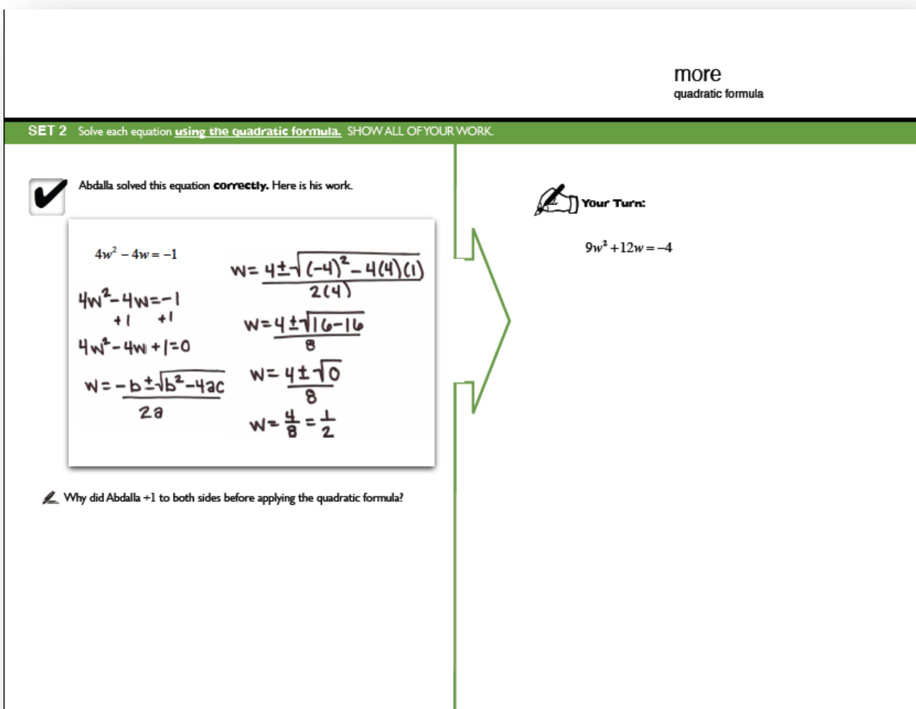
$w = \frac{-4}{2}$ $w = \frac{-8}{2}$

$w = -2$ or $w = -4$

How did Denzel know to find two solutions?

Your Turn:

$w^2 + 2w - 8 = 0$



more
quadratic formula

SET 2 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

Abdalla solved this equation **correctly.** Here is his work.

$4w^2 - 4w = -1$

$4w^2 - 4w = -1$
$+1 \quad +1$

$4w^2 - 4w + 1 = 0$

$w = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$w = \frac{4 \pm \sqrt{(-4)^2 - 4(4)(1)}}{2(4)}$

$w = \frac{4 \pm \sqrt{16 - 16}}{8}$

$w = \frac{4 \pm \sqrt{0}}{8}$

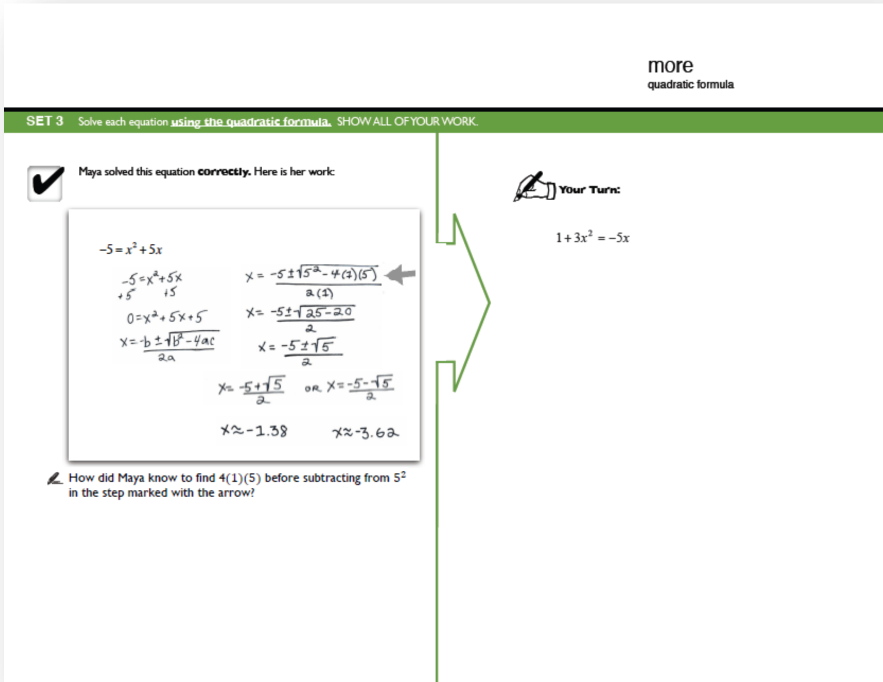
$w = \frac{4}{8} = \frac{1}{2}$

Why did Abdalla +1 to both sides before applying the quadratic formula?

Your Turn:

$9w^2 + 12w = -4$

APPENDIX E: CORRECT EXAMPLE CONTROL WORKSHEETS (continued)



more
quadratic formula

SET 3 Solve each equation using the quadratic formula. SHOW ALL OF YOUR WORK.

Maya solved this equation **correctly.** Here is her work:

$-5 = x^2 + 5x$

$-5 = x^2 + 5x$
$+5 \quad +5$
$0 = x^2 + 5x + 5$
$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$x = \frac{-5 \pm \sqrt{5^2 - 4(1)(5)}}{2(1)}$
$x = \frac{-5 \pm \sqrt{25 - 20}}{2}$
$x = \frac{-5 \pm \sqrt{5}}{2}$

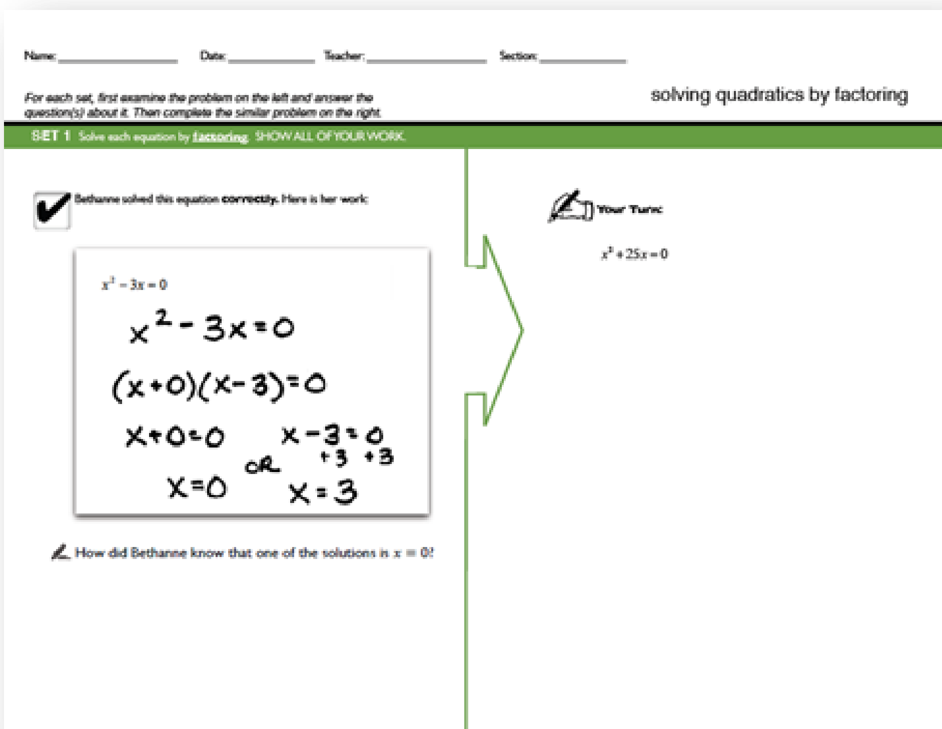
$x = \frac{-5 + \sqrt{5}}{2}$ or $x = \frac{-5 - \sqrt{5}}{2}$

$x \approx -1.38$ $\quad x \approx -3.62$

How did Maya know to find 4(1)(5) before subtracting from 5^2 in the step marked with the arrow?

Your Turn:

$1 + 3x^2 = -5x$



Name: ________ Date: ________ Teacher: ________ Section: ________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

solving quadratics by factoring

SET 1 Solve each equation by factoring. SHOW ALL OF YOUR WORK.

Bethanne solved this equation **correctly.** Here is her work:

$x^2 - 3x = 0$

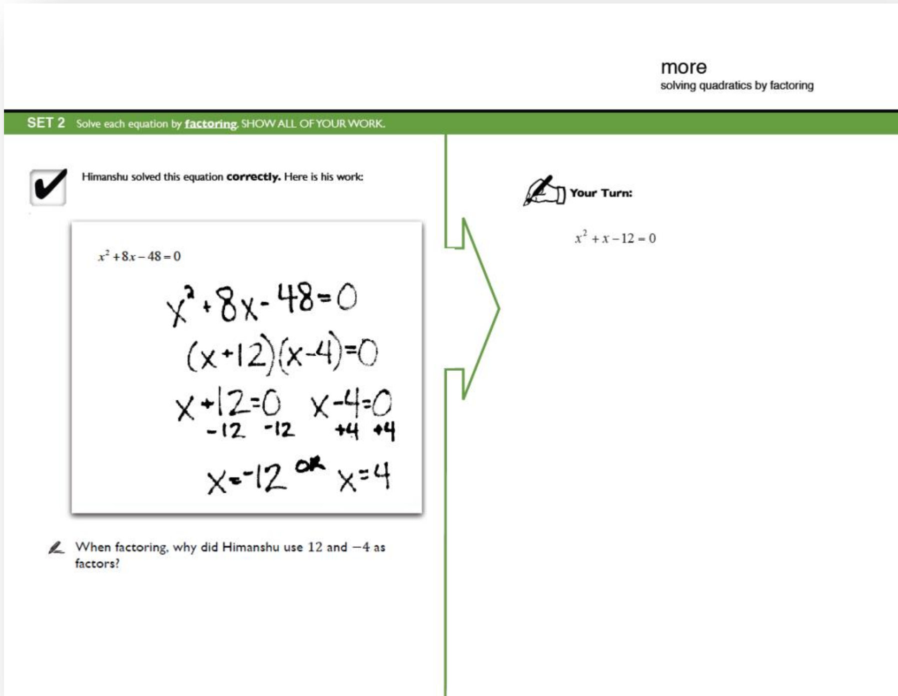
$x^2 - 3x = 0$
$(x + 0)(x - 3) = 0$
$x + 0 = 0$ or $x - 3 = 0$
$\quad +3 \quad +3$
$x = 0 \quad x = 3$

How did Bethanne know that one of the solutions is $x = 0$?

Your Turn:

$x^2 + 25x = 0$

APPENDIX E: CORRECT EXAMPLE CONTROL WORKSHEETS (continued)



more
solving quadratics by factoring

SET 2 Solve each equation by **factoring**. SHOW ALL OF YOUR WORK.

Himanshu solved this equation **correctly**. Here is his work:

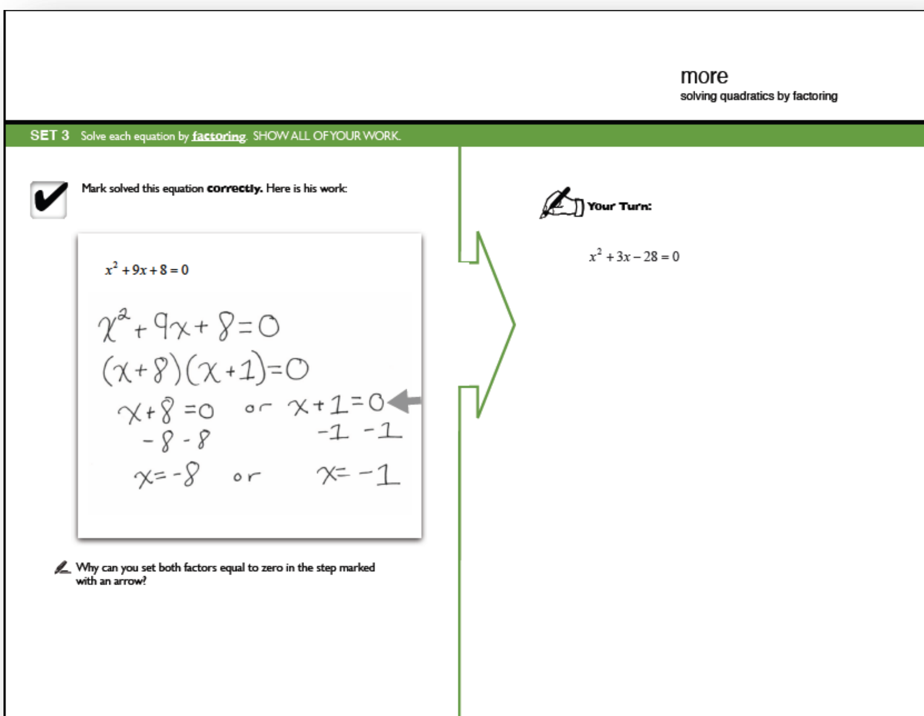
$x^2 + 8x - 48 = 0$

$x^2 + 8x - 48 = 0$
$(x+12)(x-4)=0$
$x+12=0 \quad x-4=0$
$-12 \quad -12 \quad +4 \quad +4$
$x=-12$ OR $x=4$

When factoring, why did Himanshu use 12 and −4 as factors?

Your Turn:

$x^2 + x - 12 = 0$



more
solving quadratics by factoring

SET 3 Solve each equation by **factoring**. SHOW ALL OF YOUR WORK.

Mark solved this equation **correctly**. Here is his work:

$x^2 + 9x + 8 = 0$

$x^2 + 9x + 8 = 0$
$(x+8)(x+1)=0$
$x+8=0$ or $x+1=0$
$-8 \; -8 \quad -1 \; -1$
$x=-8$ or $x=-1$

Why can you set both factors equal to zero in the step marked with an arrow?

Your Turn:

$x^2 + 3x - 28 = 0$

APPENDIX E: CORRECT EXAMPLE CONTROL WORKSHEETS (continued)

Name: ____________ Date: ________ Teacher: ____________ Section: ________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

solving quadratics using the square root

SET 1 Solve each equation by **using the square root**. SHOW ALL OF YOUR WORK.

Osvaldo solved this equation **correctly**. Here is his work:

$(n+2)^2=9$

$$(n+2)^2=9$$
$$\sqrt{(n+2)^2}=\sqrt{9}$$
$$n+2=\pm 3$$
$$n+2=3 \quad n+2=-3$$
$$-2 \;\; -2 \qquad -2 \;\; -2$$
$$n=1 \quad \text{or} \quad n=-5$$

Osvaldo found both answers correctly. What did he do to find the answer $x=-5$?

Your Turn:

$(n-5)^2=100$

more

solving quadratics using the square root

SET 2 Solve each equation by **using the square root**. SHOW ALL OF YOUR WORK.

Jasmine solved this equation **correctly**. Here is her work:

$(n-6)^2=0$

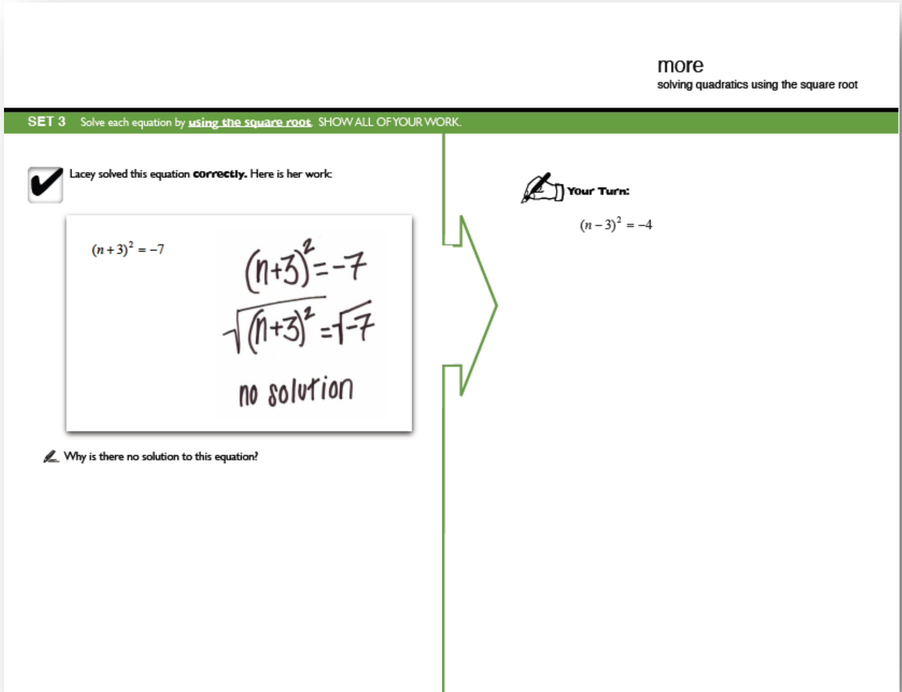
$$(n-6)^2=0$$
$$\sqrt{(n-6)^2}=\sqrt{0}$$
$$n-6=0$$
$$+6 \;\; +6$$
$$n=6$$

Why does this problem have one solution when the problems in the first set have two solutions?

Your Turn:

$(n+25)^2=0$

APPENDIX E: CORRECT EXAMPLE CONTROL WORKSHEETS (continued):



more
solving quadratics using the square root

SET 3 Solve each equation by **using the square root**. SHOW ALL OF YOUR WORK.

Lacey solved this equation **correctly**. Here is her work:

$(n+3)^2 = -7$

$(n+3)^2 = -7$

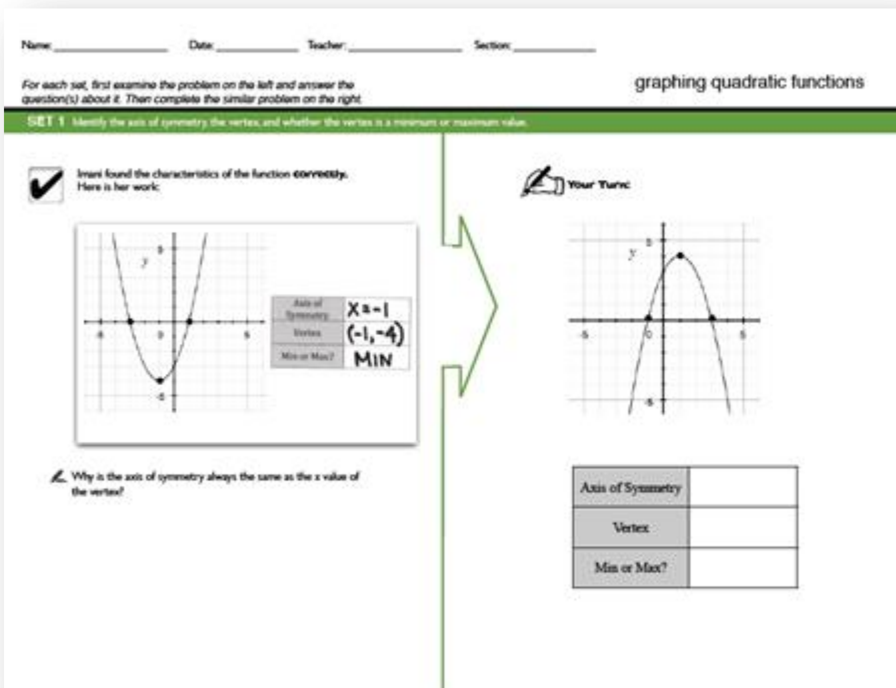
$\sqrt{(n+3)^2} = \sqrt{-7}$

no solution

Why is there no solution to this equation?

Your Turn:

$(n-3)^2 = -4$



Name: ______________ Date: __________ Teacher: ______________ Section: __________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

graphing quadratic functions

SET 1 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value.

Imani found the characteristics of the function **correctly**. Here is her work:

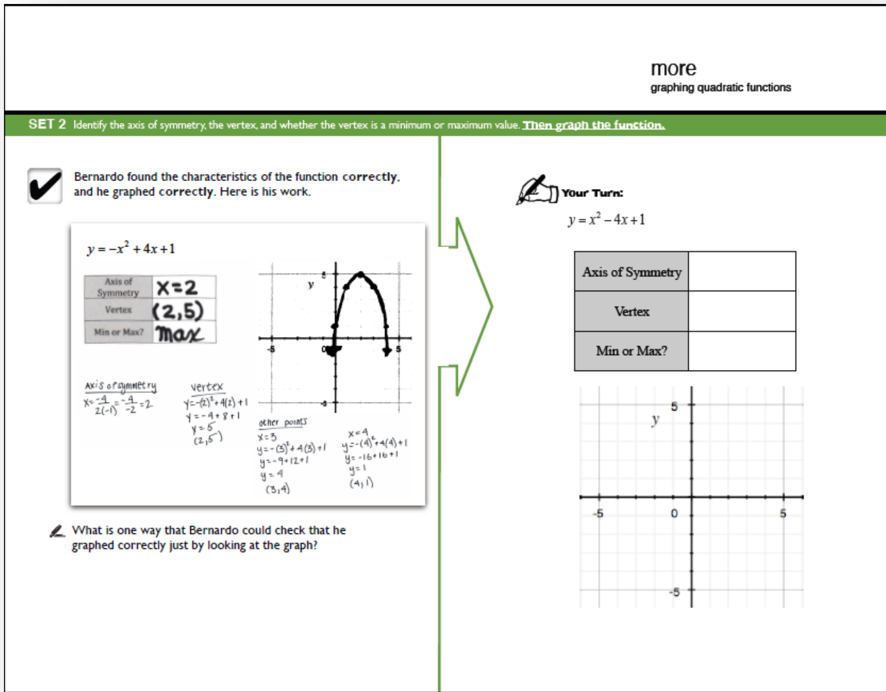
Axis of Symmetry	X = -1
Vertex	(-1, -4)
Min or Max?	MIN

Why is the axis of symmetry always the same as the x value of the vertex?

Your Turn:

Axis of Symmetry	
Vertex	
Min or Max?	

APPENDIX E: CORRECT EXAMPLE CONTROL WORKSHEETS (continued)



more
graphing quadratic functions

SET 2 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. **Then graph the function.**

Bernardo found the characteristics of the function **correctly**, and he graphed **correctly**. Here is his work.

$y = -x^2 + 4x + 1$

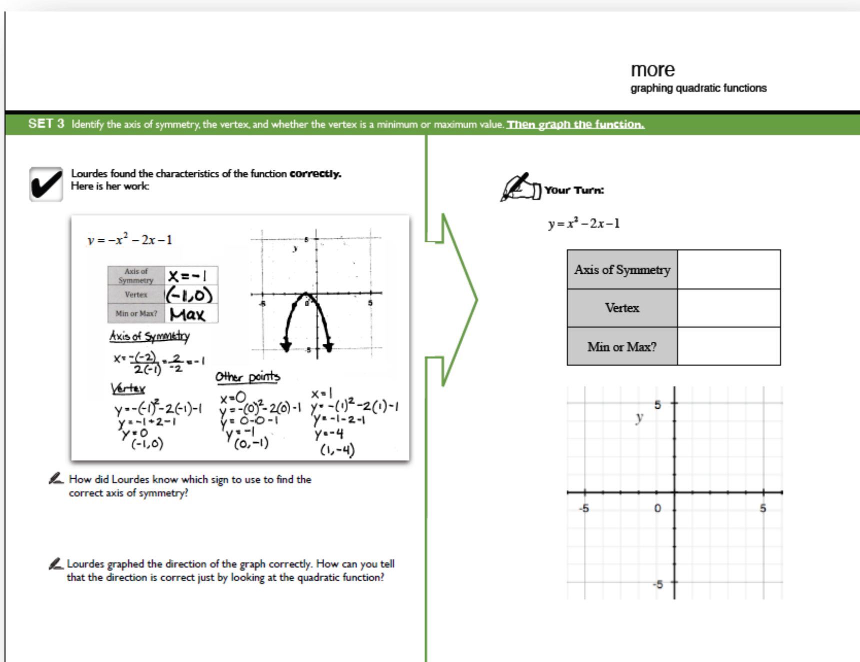
Axis of Symmetry	x=2
Vertex	(2,5)
Min or Max?	max

What is one way that Bernardo could check that he graphed correctly just by looking at the graph?

Your Turn:

$y = x^2 - 4x + 1$

Axis of Symmetry	
Vertex	
Min or Max?	



more
graphing quadratic functions

SET 3 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. **Then graph the function.**

Lourdes found the characteristics of the function **correctly**. Here is her work.

$y = -x^2 - 2x - 1$

Axis of Symmetry	x=-1
Vertex	(-1,0)
Min or Max?	Max

How did Lourdes know which sign to use to find the correct axis of symmetry?

Lourdes graphed the direction of the graph correctly. How can you tell that the direction is correct just by looking at the quadratic function?

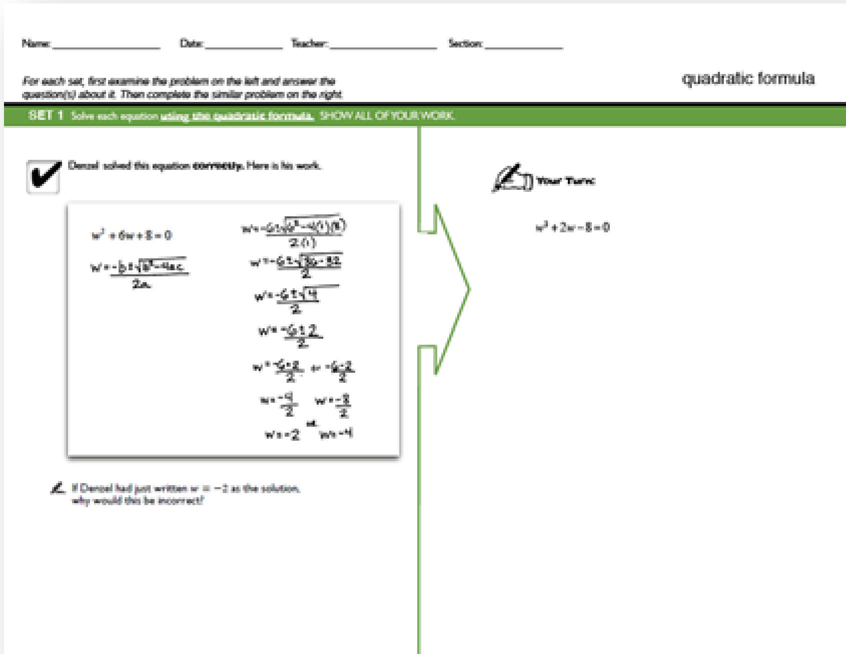
Your Turn:

$y = x^2 - 2x - 1$

Axis of Symmetry	
Vertex	
Min or Max?	

APPENDIX F:

CORRECT EXAMPLE ERROR REFLECTECTION WORKSHEETS



Name: ____________ Date: ____________ Teacher: ____________ Section: ____________

quadratic formula

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

SET 1 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

Denzel solved this equation **correctly.** Here is his work.

$w^2 + 6w + 8 = 0$

$w = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$w = \frac{-6 \pm \sqrt{6^2 - 4(1)(8)}}{2(1)}$

$w = \frac{-6 \pm \sqrt{36 - 32}}{2}$

$w = \frac{-6 \pm \sqrt{4}}{2}$

$w = \frac{-6 \pm 2}{2}$

$w = \frac{-6+2}{2}$ or $w = \frac{-6-2}{2}$

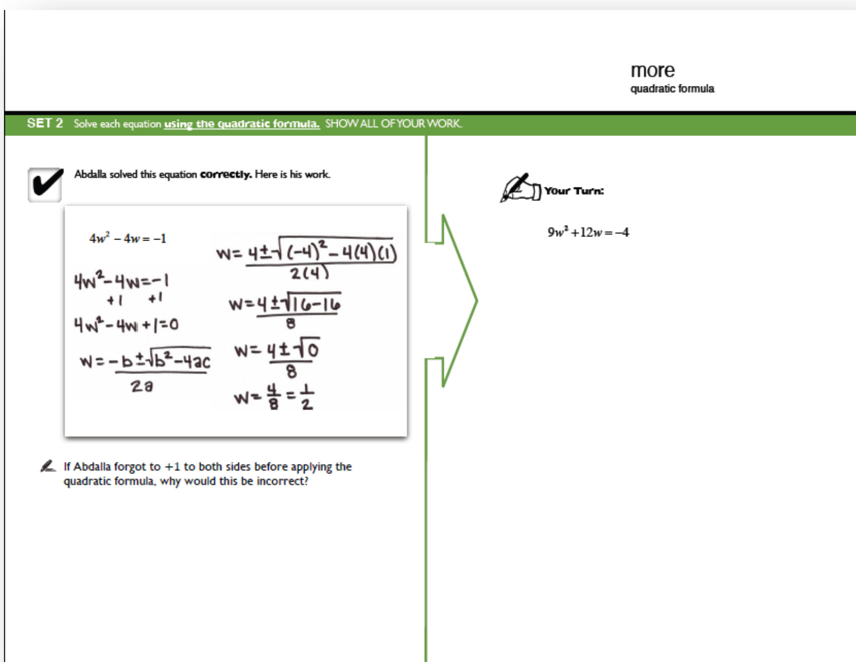
$w = \frac{-4}{2}$ $w = \frac{-8}{2}$

$w = -2$ or $w = -4$

If Denzel had just written $w = -2$ as the solution, why would this be incorrect?

Your Turn:

$w^2 + 2w - 8 = 0$



more
quadratic formula

SET 2 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

Abdalla solved this equation **correctly.** Here is his work.

$4w^2 - 4w = -1$

$4w^2 - 4w = -1$
$+1 \quad +1$

$4w^2 - 4w + 1 = 0$

$w = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$w = \frac{4 \pm \sqrt{(-4)^2 - 4(4)(1)}}{2(4)}$

$w = \frac{4 \pm \sqrt{16 - 16}}{8}$

$w = \frac{4 \pm \sqrt{0}}{8}$

$w = \frac{4}{8} = \frac{1}{2}$

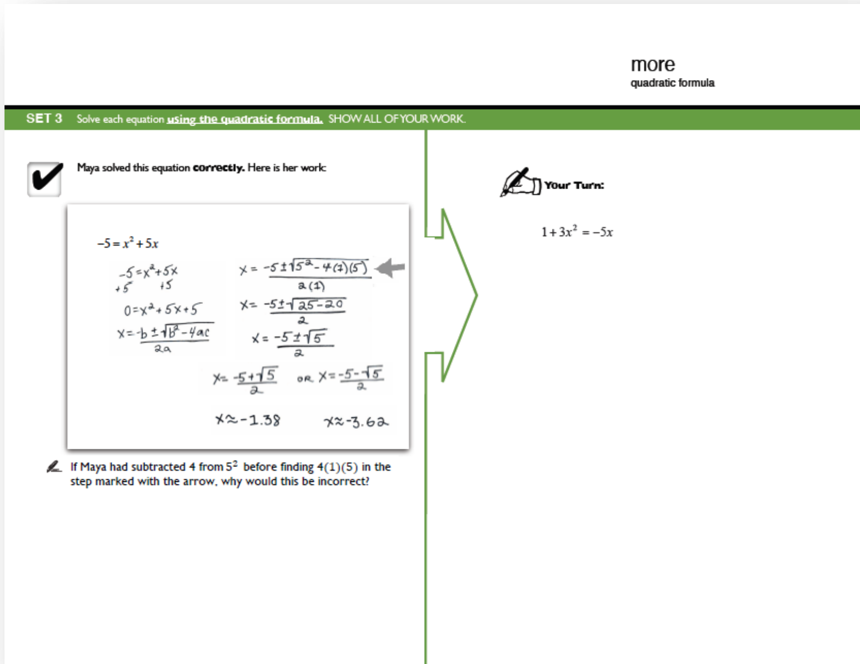
If Abdalla forgot to +1 to both sides before applying the quadratic formula, why would this be incorrect?

Your Turn:

$9w^2 + 12w = -4$

APPENDIX F:

CORRECT EXAMPLE ERROR REFLECTECTION WORKSHEETS (continued)



more
quadratic formula

SET 3 Solve each equation **using the quadratic formula.** SHOW ALL OF YOUR WORK.

Maya solved this equation **correctly.** Here is her work:

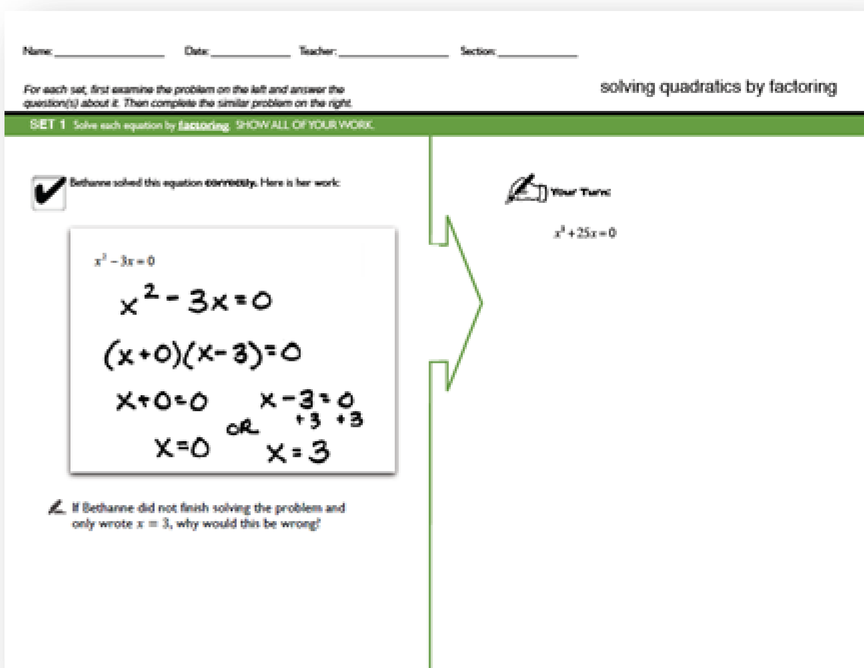
$-5 = x^2 + 5x$

$-5 = x^2 + 5x$
$+5 \quad +5$
$0 = x^2 + 5x + 5$
$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
$x = \frac{-5 \pm \sqrt{5^2 - 4(1)(5)}}{2(1)}$
$x = \frac{-5 \pm \sqrt{25 - 20}}{2}$
$x = \frac{-5 \pm \sqrt{5}}{2}$
$x = \frac{-5 + \sqrt{5}}{2}$ or $x = \frac{-5 - \sqrt{5}}{2}$
$x \approx -1.38 \quad x \approx -3.62$

If Maya had subtracted 4 from 5^2 before finding $4(1)(5)$ in the step marked with the arrow, why would this be incorrect?

Your Turn:

$1 + 3x^2 = -5x$



Name: ____________ Date: ____________ Teacher: ____________ Section: ____________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

solving quadratics by factoring

SET 1 Solve each equation by **factoring.** SHOW ALL OF YOUR WORK.

Bethanne solved this equation **correctly.** Here is her work:

$x^2 - 3x = 0$

$x^2 - 3x = 0$
$(x + 0)(x - 3) = 0$
$x + 0 = 0$ or $x - 3 = 0$
$\quad +3 \quad +3$
$x = 0 \quad x = 3$

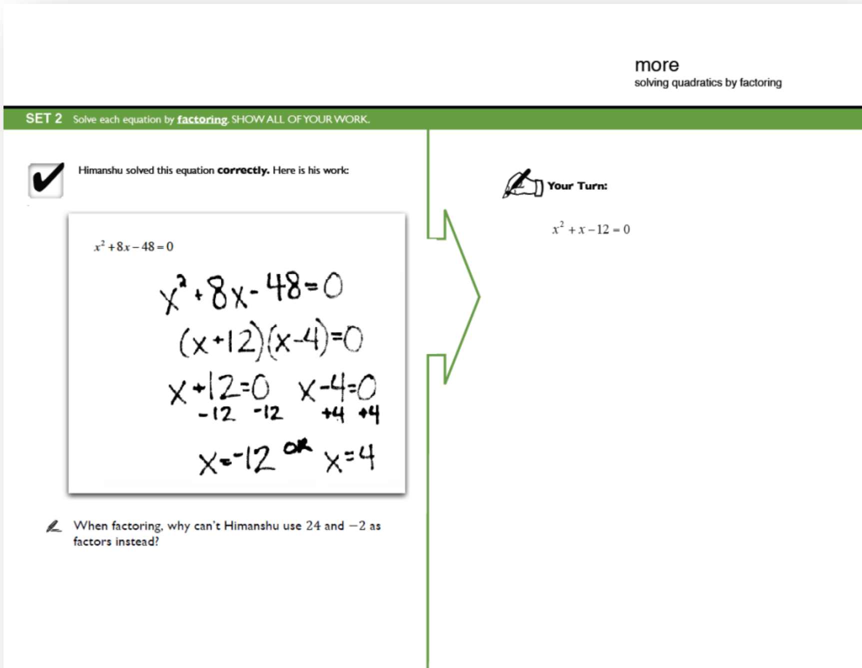
If Bethanne did not finish solving the problem and only wrote $x = 3$, why would this be wrong?

Your Turn:

$x^2 + 25x = 0$

APPENDIX F:

CORRECT EXAMPLE ERROR REFLECTECTION WORKSHEETS (continued)



more
solving quadratics by factoring

SET 2 Solve each equation by **factoring**. SHOW ALL OF YOUR WORK.

Himanshu solved this equation **correctly.** Here is his work:

$x^2 + 8x - 48 = 0$

$x^2 + 8x - 48 = 0$

$(x+12)(x-4) = 0$

$x+12=0 \quad x-4=0$

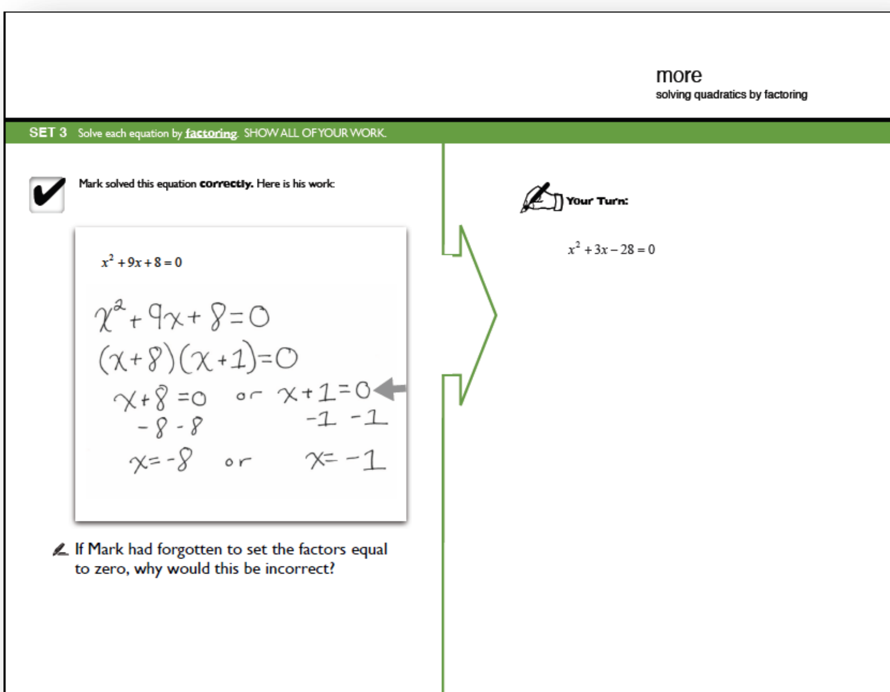
$-12 \;\; -12 \quad +4 \;\; +4$

$x = -12$ or $x = 4$

When factoring, why can't Himanshu use 24 and -2 as factors instead?

Your Turn:

$x^2 + x - 12 = 0$



more
solving quadratics by factoring

SET 3 Solve each equation by **factoring**. SHOW ALL OF YOUR WORK.

Mark solved this equation **correctly.** Here is his work:

$x^2 + 9x + 8 = 0$

$x^2 + 9x + 8 = 0$

$(x+8)(x+1) = 0$

$x+8=0$ or $x+1=0$

$-8 \;\; -8 \quad -1 \;\; -1$

$x = -8$ or $x = -1$

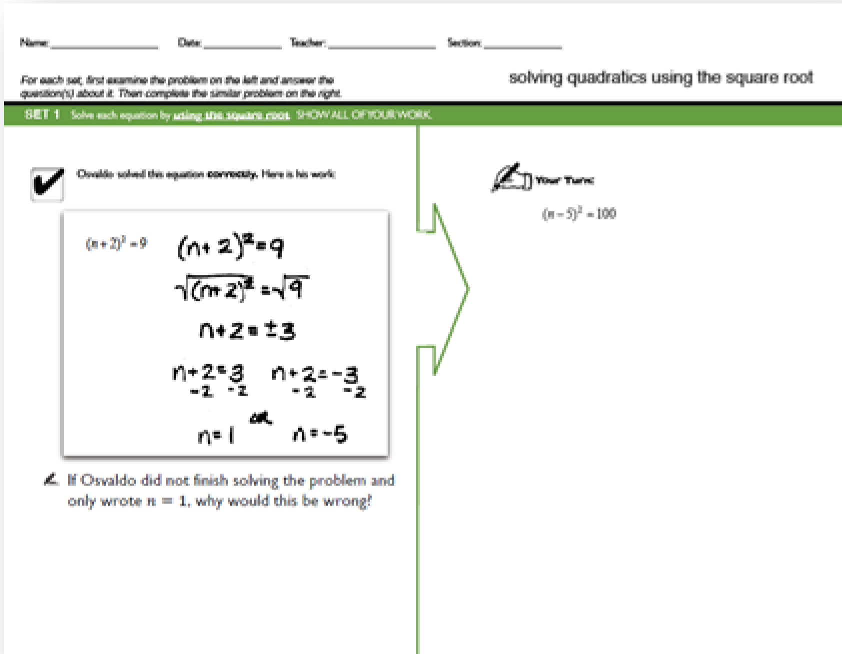
If Mark had forgotten to set the factors equal to zero, why would this be incorrect?

Your Turn:

$x^2 + 3x - 28 = 0$

APPENDIX F:

CORRECT EXAMPLE ERROR REFLECTECTION WORKSHEETS (continued)



Name: ______________ Date: __________ Teacher: ______________ Section: __________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

solving quadratics using the square root

SET 1 Solve each equation by <u>using the square root</u>. SHOW ALL OF YOUR WORK.

Osvaldo solved this equation **correctly.** Here is his work:

$(n+2)^2 = 9$

$(n+2)^2 = 9$

$\sqrt{(n+2)^2} = \sqrt{9}$

$n+2 = \pm 3$

$n+2=3 \quad n+2=-3$

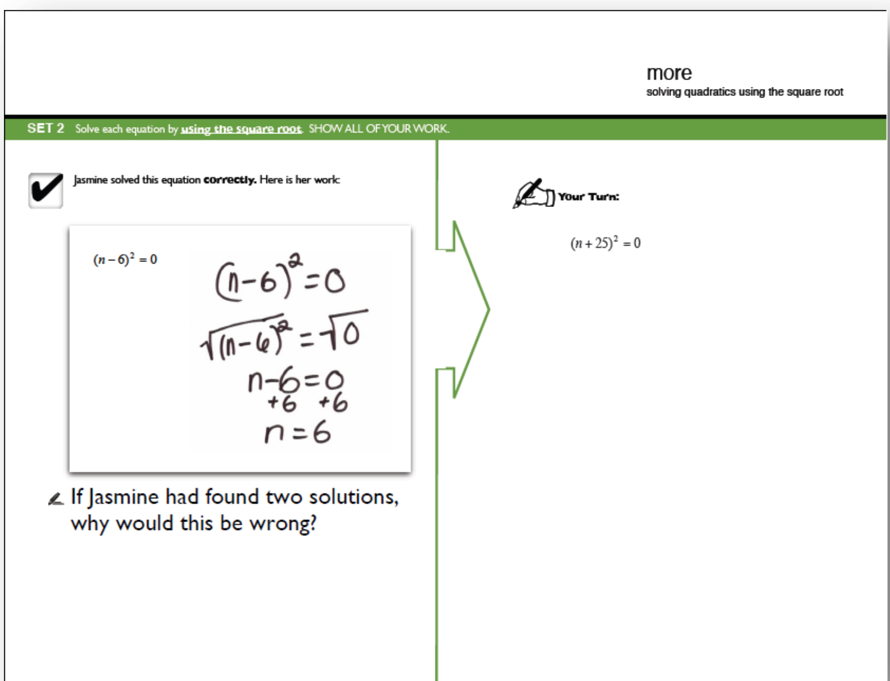
$-2 \quad -2 \qquad -2 \quad -2$

$n=1$ or $n=-5$

If Osvaldo did not finish solving the problem and only wrote $n = 1$, why would this be wrong?

Your Turn:

$(n-5)^2 = 100$



more

solving quadratics using the square root

SET 2 Solve each equation by <u>using the square root</u>. SHOW ALL OF YOUR WORK.

Jasmine solved this equation **correctly.** Here is her work:

$(n-6)^2 = 0$

$(n-6)^2 = 0$

$\sqrt{(n-6)^2} = \sqrt{0}$

$n-6=0$

$+6 \quad +6$

$n=6$

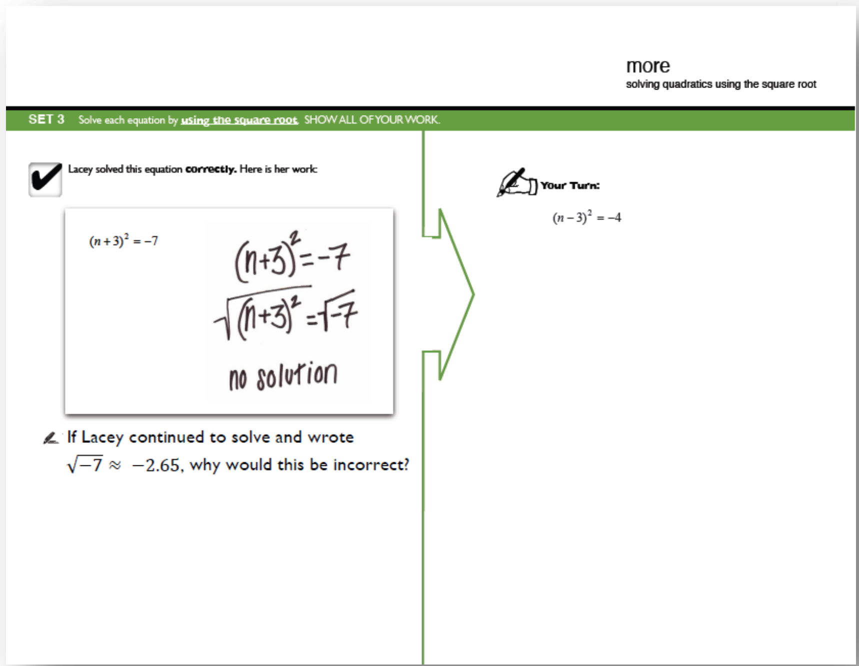
If Jasmine had found two solutions, why would this be wrong?

Your Turn:

$(n+25)^2 = 0$

APPENDIX F:

CORRECT EXAMPLE ERROR REFLECTECTION WORKSHEETS (continued)


more

solving quadratics using the square root

SET 3 Solve each equation by using the square root. SHOW ALL OF YOUR WORK.

Lacey solved this equation **correctly.** Here is her work:

$(n+3)^2 = -7$

$(n+3)^2 = -7$

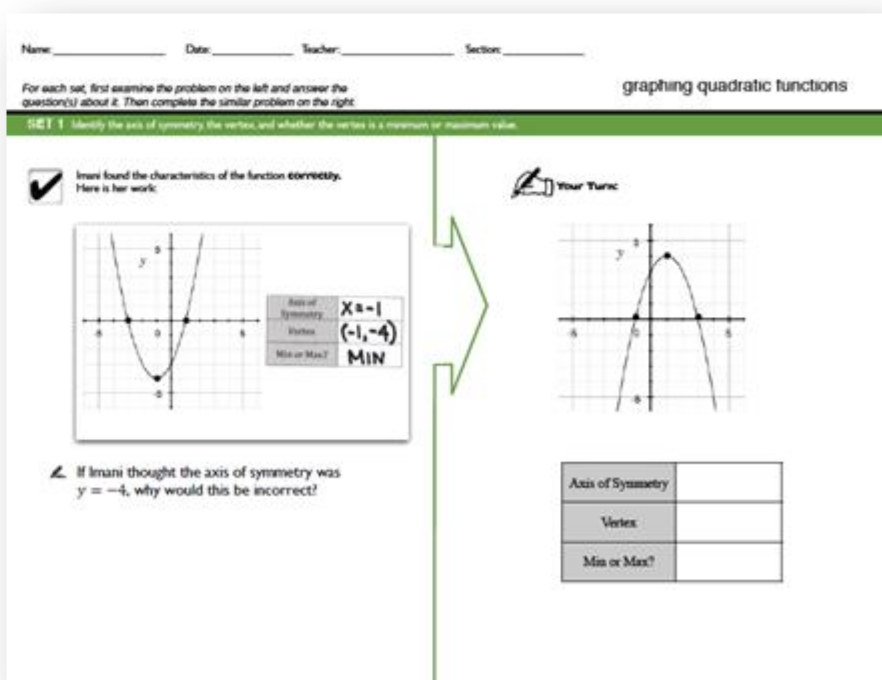
$\sqrt{(n+3)^2} = \sqrt{-7}$

no solution

If Lacey continued to solve and wrote $\sqrt{-7} \approx -2.65$, why would this be incorrect?

Your Turn:

$(n-3)^2 = -4$


Name: ______ Date: ______ Teacher: ______ Section: ______

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

graphing quadratic functions

SET 1

Imani found the characteristics of the function **correctly.** Here is her work:

Axis of Symmetry	X = -1
Vertex	(-1, -4)
Min or Max?	MIN

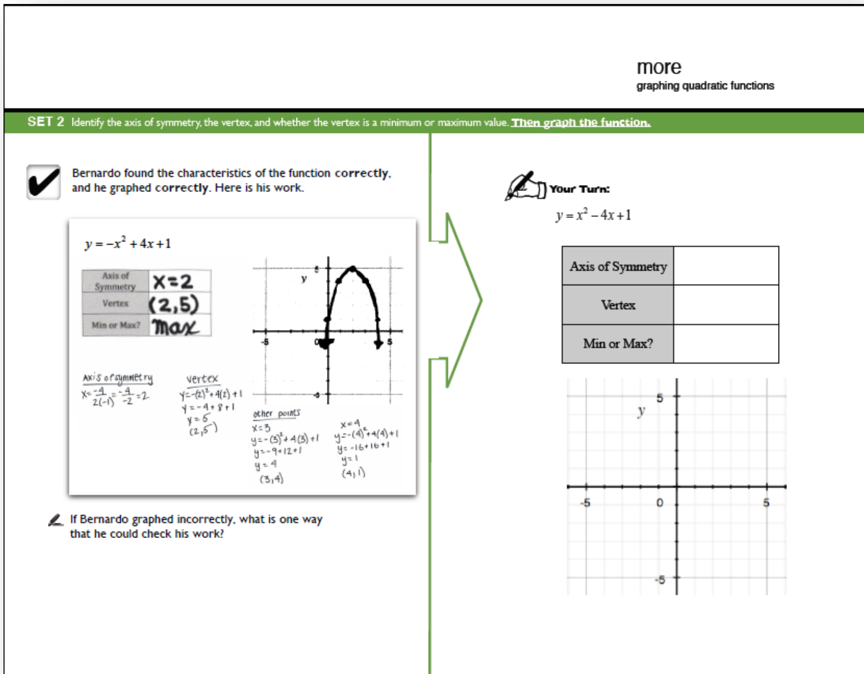
If Imani thought the axis of symmetry was $y = -4$, why would this be incorrect?

Your Turn:

Axis of Symmetry	
Vertex	
Min or Max?	

APPENDIX F:

CORRECT EXAMPLE ERROR REFLECTECTION WORKSHEETS (continued)



more
graphing quadratic functions

SET 2 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. **Then graph the function.**

Bernardo found the characteristics of the function **correctly**, and he graphed **correctly**. Here is his work.

$y = -x^2 + 4x + 1$

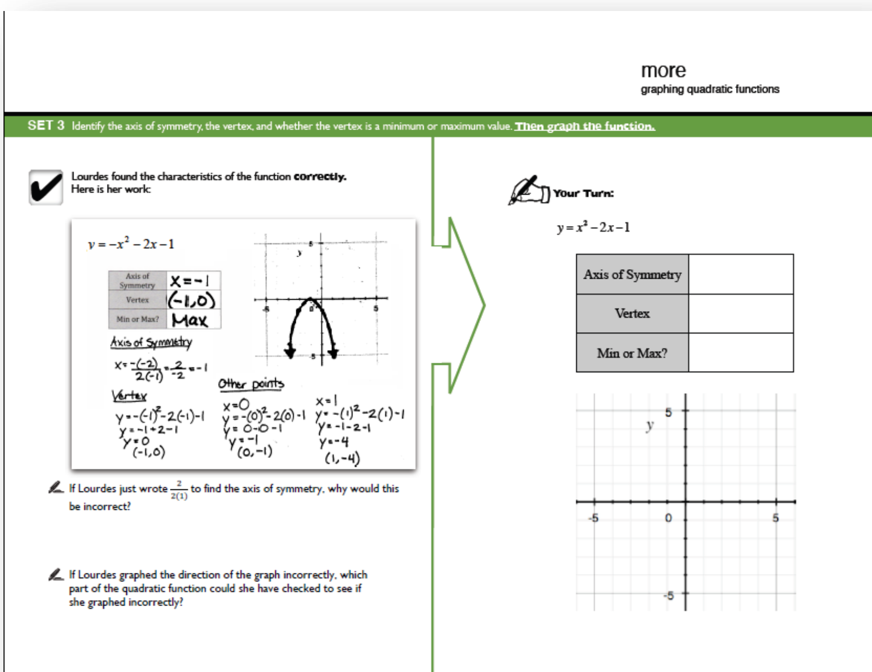
Axis of Symmetry	x=2
Vertex	(2,5)
Min or Max?	max

If Bernardo graphed incorrectly, what is one way that he could check his work?

Your Turn:

$y = x^2 - 4x + 1$

Axis of Symmetry	
Vertex	
Min or Max?	



more
graphing quadratic functions

SET 3 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value. **Then graph the function.**

Lourdes found the characteristics of the function **correctly**. Here is her work.

$y = -x^2 - 2x - 1$

Axis of Symmetry	x=-1
Vertex	(-1,0)
Min or Max?	Max

If Lourdes just wrote $\frac{2}{2(1)}$ to find the axis of symmetry, why would this be incorrect?

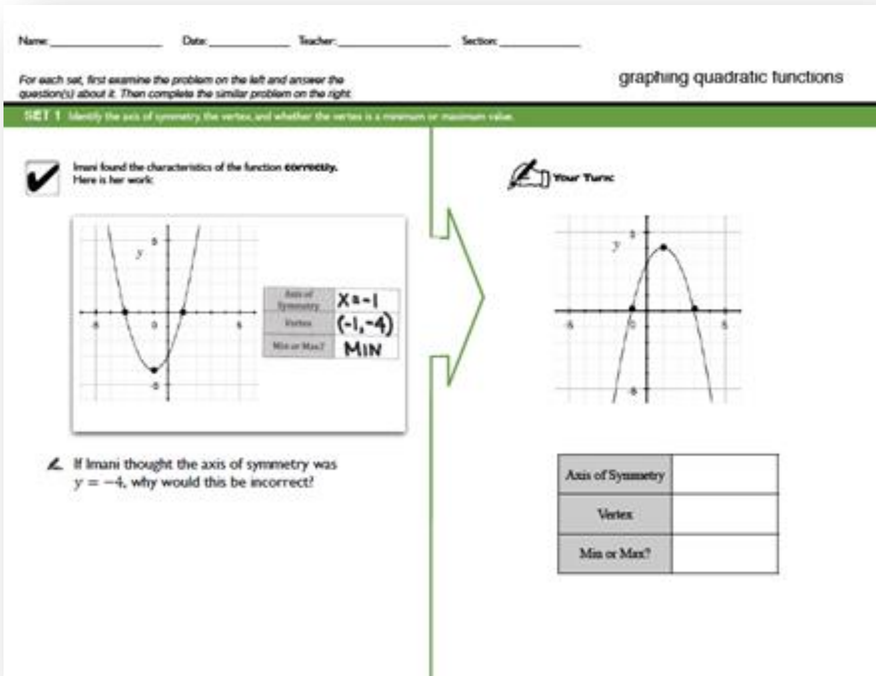
If Lourdes graphed the direction of the graph incorrectly, which part of the quadratic function could she have checked to see if she graphed incorrectly?

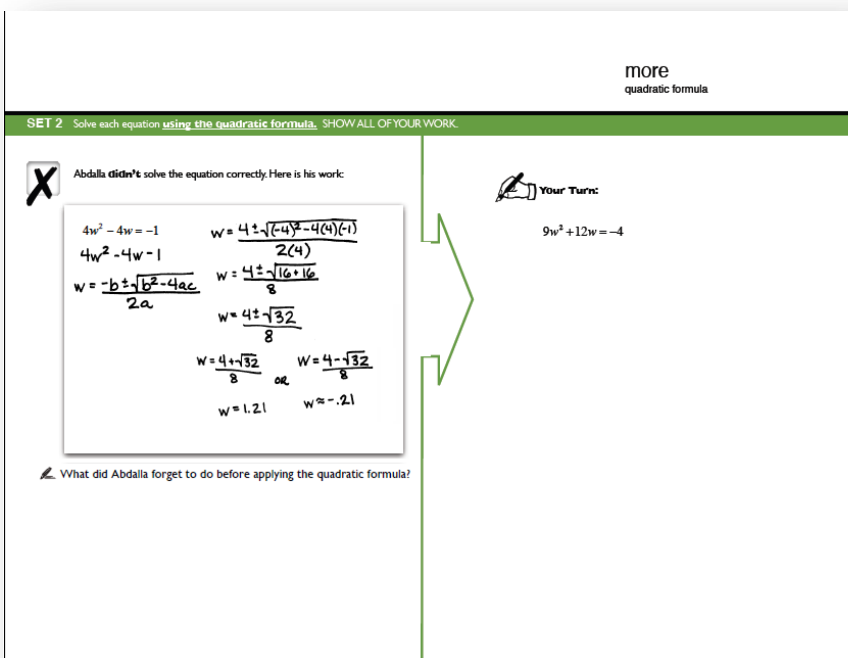
Your Turn:

$y = x^2 - 2x - 1$

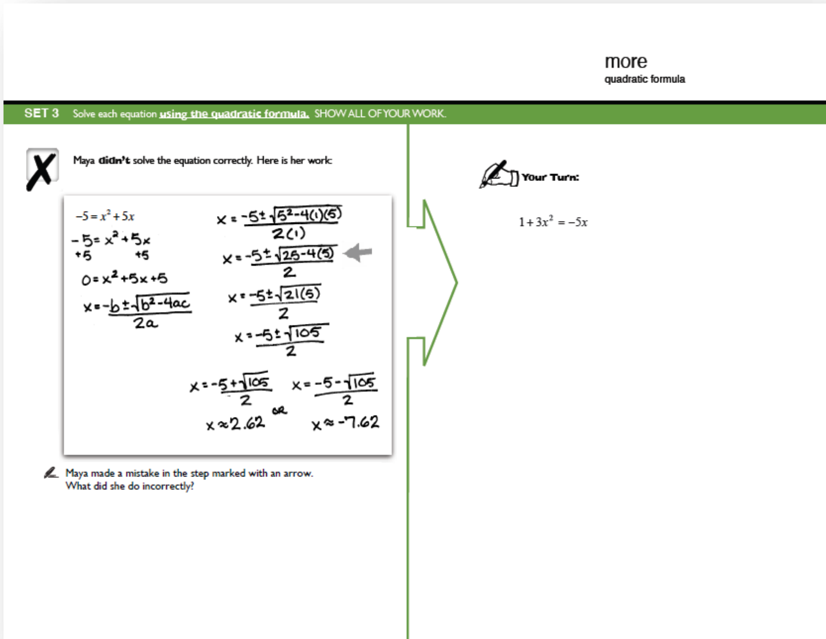
Axis of Symmetry	
Vertex	
Min or Max?	

APPENDIX G: INCORRECT EXAMPLE ERROR REFLECTION WORKSHEETS





APPENDIX G: INCORRECT EXAMPLE ERROR REFLECTION WORKSHEETS (continued)



more
quadratic formula

SET 3 Solve each equation using the quadratic formula. SHOW ALL OF YOUR WORK.

Maya **didn't** solve the equation correctly. Here is her work:

$-5 = x^2 + 5x$

$-5 = x^2 + 5x$
$+5 \qquad +5$
$0 = x^2 + 5x + 5$

$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

$x = \frac{-5 \pm \sqrt{5^2 - 4(1)(5)}}{2(1)}$

$x = \frac{-5 \pm \sqrt{25 - 4(5)}}{2}$ ←

$x = \frac{-5 \pm \sqrt{21(5)}}{2}$

$x = \frac{-5 \pm \sqrt{105}}{2}$

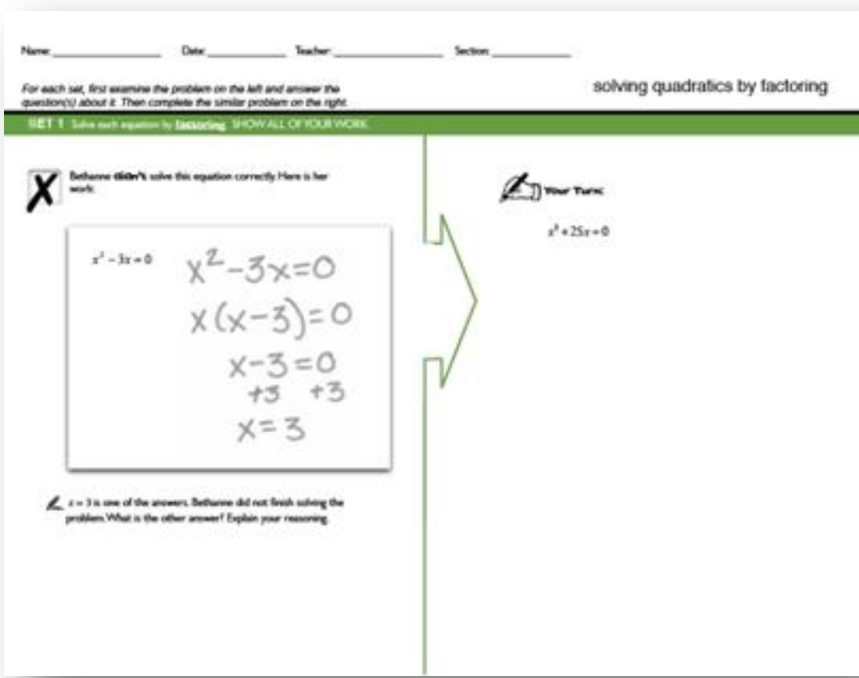
$x = \frac{-5 + \sqrt{105}}{2}$ or $x = \frac{-5 - \sqrt{105}}{2}$

$x \approx 2.62$ or $x \approx -7.62$

Maya made a mistake in the step marked with an arrow. What did she do incorrectly?

Your Turn:

$1 + 3x^2 = -5x$



Name: ________ Date: ______ Teacher: ________ Section: ______

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

solving quadratics by factoring

SET 1 Solve each equation by factoring. SHOW ALL OF YOUR WORK.

Bethanne **didn't** solve this equation correctly. Here is her work:

$x^2 - 3x = 0$

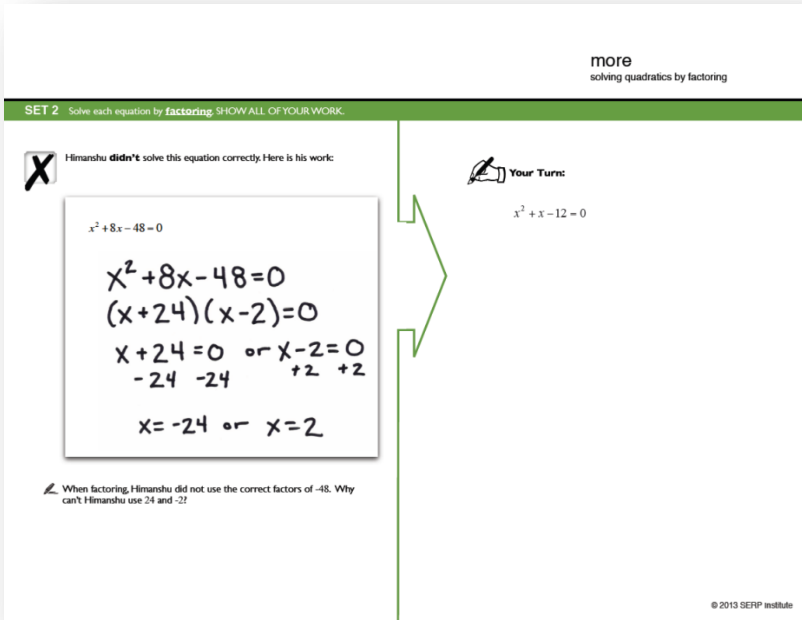
$x^2 - 3x = 0$
$x(x - 3) = 0$
$x - 3 = 0$
$+3 \quad +3$
$x = 3$

$x = 3$ is one of the answers. Bethanne did not finish solving the problem. What is the other answer? Explain your reasoning.

Your Turn:

$x^2 + 25x = 0$

APPENDIX G: INCORRECT EXAMPLE ERROR REFLECTION WORKSHEETS (continued)



more
solving quadratics by factoring

SET 2 Solve each equation by **factoring**. SHOW ALL OF YOUR WORK.

Himanshu **didn't** solve this equation correctly. Here is his work:

$x^2 + 8x - 48 = 0$

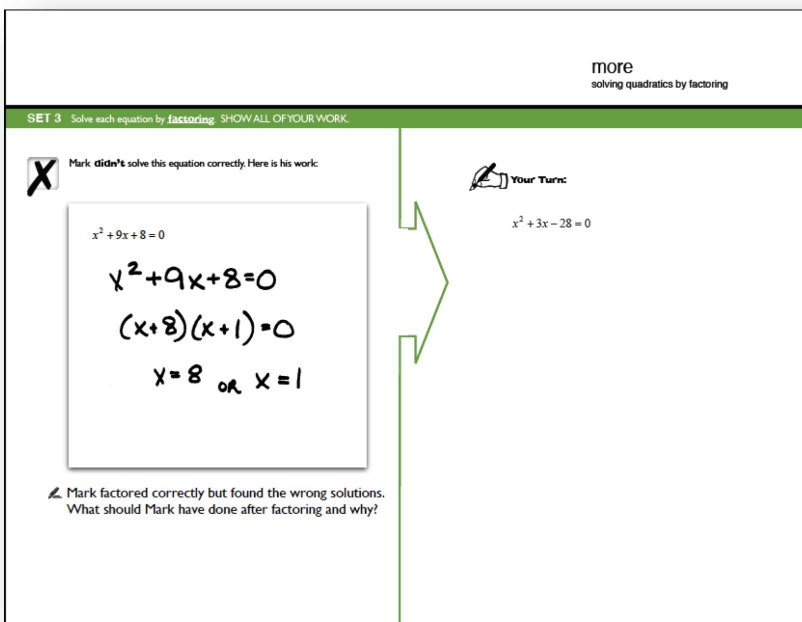
$x^2 + 8x - 48 = 0$
$(x + 24)(x - 2) = 0$
$x + 24 = 0$ or $x - 2 = 0$
$-24 \quad -24 \qquad +2 \quad +2$
$x = -24$ or $x = 2$

When factoring, Himanshu did not use the correct factors of -48. Why can't Himanshu use 24 and -2?

Your Turn:

$x^2 + x - 12 = 0$

© 2013 SERP Institute



more
solving quadratics by factoring

SET 3 Solve each equation by **factoring**. SHOW ALL OF YOUR WORK.

Mark **didn't** solve this equation correctly. Here is his work:

$x^2 + 9x + 8 = 0$

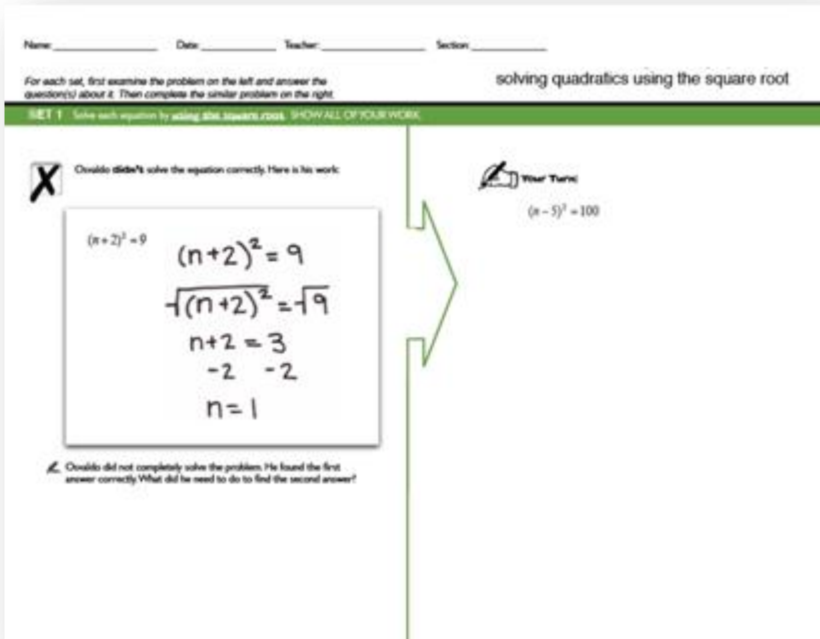
$x^2 + 9x + 8 = 0$
$(x + 8)(x + 1) = 0$
$x = 8$ or $x = 1$

Mark factored correctly but found the wrong solutions. What should Mark have done after factoring and why?

Your Turn:

$x^2 + 3x - 28 = 0$

APPENDIX G: INCORRECT EXAMPLE ERROR REFLECTION WORKSHEETS (continued)


Name: ______________ Date: __________ Teacher: ______________ Section: __________

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

solving quadratics using the square root

SET 1 Solve each equation by using the square root. SHOW ALL OF YOUR WORK.

Osvaldo **didn't** solve the equation correctly. Here is his work:

$(n+2)^2=9$

$(n+2)^2 = 9$

$\sqrt{(n+2)^2} = \sqrt{9}$

$n+2=3$

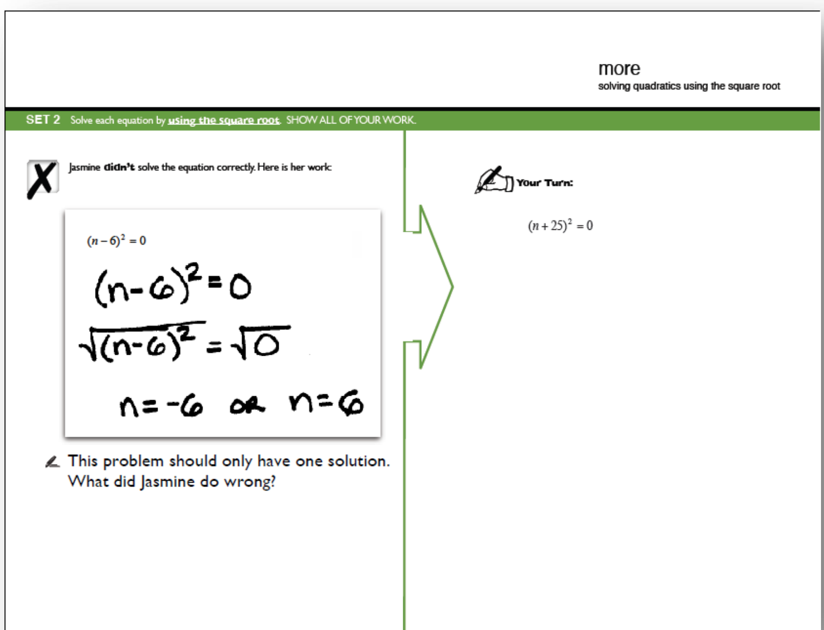
$-2 \quad -2$

$n=1$

Osvaldo did not completely solve the problem. He found the first answer correctly. What did he need to do to find the second answer?

Your Turn:

$(n-5)^2=100$


more

solving quadratics using the square root

SET 2 Solve each equation by using the square root. SHOW ALL OF YOUR WORK.

Jasmine **didn't** solve the equation correctly. Here is her work:

$(n-6)^2=0$

$(n-6)^2=0$

$\sqrt{(n-6)^2} = \sqrt{0}$

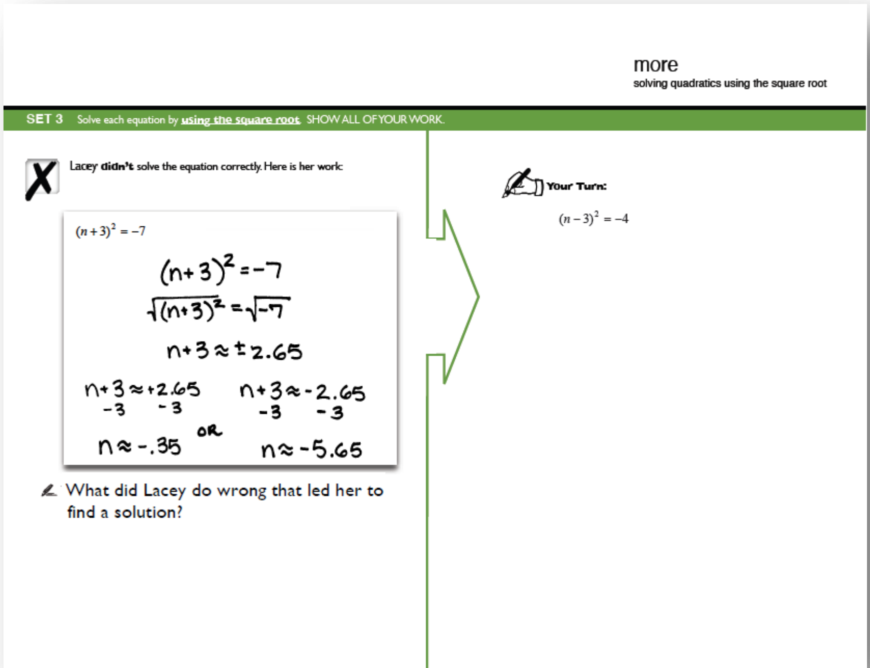
$n=-6$ OR $n=6$

This problem should only have one solution. What did Jasmine do wrong?

Your Turn:

$(n+25)^2=0$

APPENDIX G: INCORRECT EXAMPLE ERROR REFLECTION WORKSHEETS (continued)



more
solving quadratics using the square root

SET 3 Solve each equation by using the square root. SHOW ALL OF YOUR WORK.

Lacey **didn't** solve the equation correctly. Here is her work:

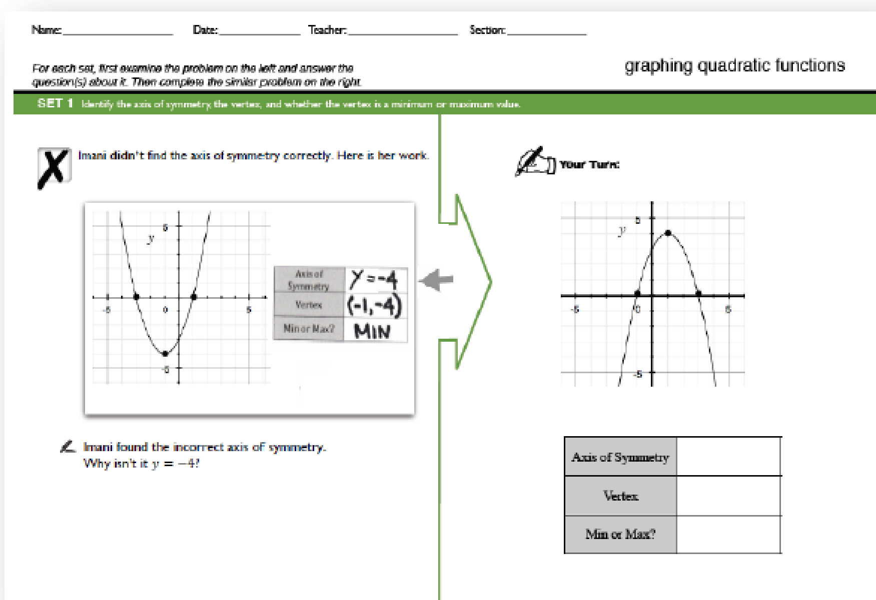
$(n+3)^2 = -7$

$(n+3)^2 = -7$
$\sqrt{(n+3)^2} = \sqrt{-7}$
$n+3 \approx \pm 2.65$
$n+3 \approx +2.65$, -3 -3 OR $n+3 \approx -2.65$, -3 -3
$n \approx -.35$ OR $n \approx -5.65$

What did Lacey do wrong that led her to find a solution?

Your Turn:

$(n-3)^2 = -4$



Name:__________ Date:__________ Teacher:__________ Section:__________

graphing quadratic functions

For each set, first examine the problem on the left and answer the question(s) about it. Then complete the similar problem on the right.

SET 1 Identify the axis of symmetry, the vertex, and whether the vertex is a minimum or maximum value.

Imani didn't find the axis of symmetry correctly. Here is her work.

Axis of Symmetry	$Y=-4$
Vertex	$(-1,-4)$
Min or Max?	MIN

Imani found the incorrect axis of symmetry. Why isn't it $y = -4$?

Your Turn:

Axis of Symmetry	
Vertex	
Min or Max?	

APPENDIX G: INCORRECT EXAMPLE ERROR REFLECTION WORKSHEETS (continued)