THE UNIVERSITY OF CHICAGO

LEARNING BEYOND ACCURACY: EVIDENCE FOR WORKED EXAMPLES AS A SUPPORT FOR STUDENTS' PROPORTIONAL REASONING GAINS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPARATIVE HUMAN DEVELOPMENT

AND

DEPARTMENT OF PSYCHOLOGY

BY

ELAYNE PATRICE VOLLMAN

CHICAGO, ILLINOIS

JUNE 2021

TABLE OF CONTENTS

ACKNOWLEDGMENTS

This dissertation was funded by the Institute of Education Sciences. I thank the school districts, teachers and students who participated in these studies. I also thank my lab and the research assistants at University of Chicago and University of California, Irvine who worked with me during the data collection and coding process for these studies.

There are many people who have been instrumental to my success in graduate school that I must also thank. First, for those who served on my committee. To my advisor Lindsey Richland who supported my various research curiosities as they developed throughout my graduate program. I am grateful for your guidance in developing my skills as a researcher and scholar; I appreciate your "big picture" perspective and conducting research that has real-life implications. Thank you to Susan Goldin-Meadow for bringing me on as your student in Psychology. Your kindness to include me in your lab and support my development accelerated my learning and scholarship in ways that otherwise would have not been possible. Thank you also to Susan Levine for meeting with me various times throughout my program and providing valuable insight and inspiration to me. Steve Raudenbush, thank you for the expertise, mentorship, and the opportunities you have provided me throughout my IES Fellowship. I'd also like to thank Kelly Trezise, for the research expertise and mentorship she provided to me especially in developing my thinking and writing skills.

Finally, thank you to my family and friends who have supported me in my life-long academic journey.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A series of studies examined the cognitive underpinnings of worked examples as a tool for introducing multiple solution strategies for the same concept, here proportional reasoning, and for supporting complex higher order thinking. Worked examples were also tested as a tool to protect against temporarily strained cognitive resources (i.e., high pressure contexts), and resultant student solutions were explored in detail to understand affordances and limits in problem solving based on worked examples. Specifically, across this dissertation latent class analysis was used to assess learning beyond problem accuracy; qualitative frequency codes provided quantitative data about emergent patterns in students' problem-solving including their selection of approach (type of proportional or non-proportional solution strategy), setup, procedures, and final solution accuracy. In Study 1, I examined students' learning of novel solution strategies when provided with instructional packets that included i) fully worked examples or ii) partially worked examples, compared to learning gains in a control condition: iii) problems only. Study 2 built on Study 1 by examining the potential benefits of incorporating worked examples during classroom instruction. Study 3 further built on Studies 1 and 2 by imposing pressure as a means to reduce students' cognitive capacity, thereby allowing us to evaluate whether visual support with worked examples could increase instances of higher order thinking and improve longer-term problem-solving outcomes. Together, findings from these experiments indicate that fully worked examples are a highly implementable, malleable tool, which provide the necessary scaffolding for students to fundamentally change their thinking and promote novel and more sophisticated problem-solving approaches with success. Further, worked examples may buffer against the detriments caused by pressure or lower availability of cognitive functioning resources.

CHAPTER ONE: INTRODUCTION

There is broad consensus that the ability to use mathematics flexibly is increasingly important as mathematics aptitude commonly represents an integral component of cognition and the basis for many 21st century skills. Mathematics knowledge in students' early years is a predictor of high school graduation and earning potential (McCoy et al., 2017; Watts et al., 2014). Despite its importance, students in the United States lag behind international peers as early as preschool and this gap widens over time (Gerofsky, 2015; Mullis et al., 2012; Ritchie & Bates, 2013). Within the US there is also an increasing gap between children who grow up in higher and lower resourced communities (Bachman et al., 2015). This presents a crucial need for understanding and reducing persistent achievement gaps in mathematics within the United States and between the US and its higher achieving peer nations.

High quality instruction is one avenue to reduce these gaps. Yet, in the US there is a lack of normative teaching practices that engage students in making connections and reasoning deeply about mathematics (see Hiebert, 2003; Hiebert et al., 2003; Richland et al., 2007). In the Third International Mathematics and Science Study (TIMSS - Hiebert et al., 2003), US teaching practices were compared with higher achieving countries through live-unedited classroom videos. The key difference between the higher and lower achieving regions was the amount of support teachers provided students for drawing connections during problem-solving. While mathematics teachers in the United States frequently use comparison and relational reasoning as instructional tools to teach concepts and procedures, it is not always in ways that encourage conceptual and procedural gains (Richland et al., 2007).

In high achieving regions like Hong Kong and Japan, teachers seem to be more attentive to the support required to learn complex material and the limitations it places on cognitive resources. Their teaching reflects the use of highly implementable, teaching principles that use visual supports such as worked examples and gestures to draw explicit comparisons, highlight key structural features of problem-solving, and reduce cognitive load (Gonzales et al., 2003; Richland et al., 2007). For example, expert teachers deploy five key teaching practices: 1) used a familiar source representation to compare to the target representation being taught, 2) presented the source representation visually, 3) kept the source visible to learners during comparison with the target, 4) used spatial cues to highlight alignment between the corresponding elements, and 5) used gestures to draw attention to the comparison (Richland et al., 2007). A teacher can achieve the first three practices through the use of visual support via worked examples and practices 4 and 5 by highlighting the relational comparisons through visual cueing (i.e., gesturing) during direct instruction. While these teaching practices were observed in high achieving classrooms and are correlated with mathematics achievement, the direct impact and mechanisms of these teaching practices on problem-solving have not yet been experimentally manipulated and understood in US classrooms. Moreover, no work has examined the features and steps of problem-solving that are enhanced by visual support.

In this thesis, these issues are investigated in the learning context of proportional reasoning. Proportional reasoning is an everyday mathematics context that underpins important structural relationships in mathematics and science as well as in everyday life (Cramer & Post, 1993; Lesh et al., 1988). For example, it is required for economic values, spatial contrasts, temperatures, densities, concentrations, velocities, chemical compositions, and recipes (Karplus et al., 1983; Moore et al., 1991; Siegler & Vago, 1978). It also foregrounds complex domains of mathematics for rational number operations, unit partitioning, and basic algebra (Empson, 1999; Fuson & Abrahamson, 2005). Proportional reasoning has been identified by the National Council of Teachers of Mathematics (1989) as a crucial mathematics skill for students to develop.

Despite the importance of proportional reasoning, students have pervasive difficulties acquiring these complex reasoning skills. There are at least four reasons students experience difficulties. First, proportional reasoning highlights one of the first conceptual shifts in mathematics learning. In less sophisticated mathematics contexts, students engage in additive thinking; proportional reasoning requires students to shift to a relational mindset (Gentner, 1988) which uses multiplicative thinking (Fuson & Abrahamson, 2005). Second, students have difficulty moving from whole number relationships with basic ratio (e.g., 3:7 = ?:14) to middle difficulty non-divisible problems (e.g., 6:14 = ?:35; Fuson & Abrahamson, 2005). This is in part due to the procedural difficulty of long division and decimal multiplication and in part due to the increased demand on working memory that is necessary to manipulate all parts of the problem. Third, proportional reasoning requires attending to and understanding the relationship between units (e.g., A:A -> B:B) and across units (e.g., A:B -> A:B). For example:

> Alex is making a strawberry cake. To make a small cake, the recipe calls for 2 eggs and 6 strawberries. Alex wants to make a big cake so he uses 8 eggs. How many strawberries does Alex need for the big cake?

Students must engage in multiple levels of relational reasoning between the ingredients within the cake and across the cakes. The ability to understand the multiplicative relationship between quantities and phenomena as systems of structured relationships that can be aligned, compared, and mapped together is both conceptually challenging and cognitively taxing (Cramer & Post, 1993; Boyer et al., 2008; Richland & Simms, 2015). Finally, these learning variations are affected by individual differences in students' cognitive functions such as working memory (Ackerman & Humphreys, 1990; Schneider et al., 1996).

Proportional reasoning is a complex and highly challenging domain to master and an even more complex domain to instruct. This is in part due to the fact that classrooms are not homogenous; teachers must develop lessons that fit the needs of 20 or more students with varying levels of cognitive capacity, prior knowledge, feelings of pressure, and mathematics anxiety. Developing effective and implementable instructional practices for teaching can be optimized if pedagogical practices take into account a cognitive science framework and what is known about how children think and learn (Wood et al., 2001). Visual supports, like the ones used in high achieving contexts as described above, may be related to achievement since they provide the necessary scaffolding to enhance relational comparison and reduce cognitive load; however, the exact mechanisms that underpin the links between effective practices and strong math outcomes are not clear. Examination of specific visual supports and their impact on problem-solving skills is necessary to elucidate how visual support impacts learning. Specifically, it is critical for research to understand how worked examples help children change their thinking from elementary and incorrect strategies (i.e., additive thinking) to more sophisticated and accurate strategy solutions (e.g., equivalent fraction and unit ratio strategy).

Worked examples combine many features of high quality visual supports that were observed in the TIMMS study (Hiebert et al., 2003; Richland et al., 2007). Worked examples are visual supports that are presented during problem-solving as a means to scaffold learning by providing an opportunity for students to map expert solution strategies to near-transfer contexts (Sweller & Cooper, 1985). Importantly, worked examples are posited to reduce cognitive load allowing students to focus on the structural components of problem-solving (Tuovinen & Sweller, 1999).

There are different forms of worked examples, including fully worked examples and partially worked examples. Fully worked examples include a written worked example that includes all aspects of the problem soliton - strategy approaches, setups, procedures, and solutions, and have been consistently shown to boost problem-solving skills compared to problem-solving alone (Sweller & Cooper, 1985; Zhu & Simon, 1987). Partial worked examples describes an approach in which problem setup and some of the solution is provided but the reasoner must complete a key step to the solution (Carroll, 1994). These can be used on their own or with a technique termed fading in which a series of worked examples are presented and the problem steps are gradually eliminated so that the student solves increasing amounts of the procedures independently. Partial and faded worked examples have been found to lead to higher transfer performance and enhanced learning compared to problem-solving alone (Moreno et al., 2006; Salden et al., 2010). Partial and faded worked examples have also been used as a way to transition students from studying fully worked examples to solving problems independently (see Renkl & Atkinson, 2003; Sweller et al., 2003). Criticism of the worked example literature (Koedinger & Aleven, 2007; McLaren, et al., 2006) have argued that the positive effects of worked examples are due to the fact that studies compare worked examples to unsupported problem-solving (Mwangi & Sweller, 1998). Research is needed to elucidate the impacts of these different forms of worked examples and their differential effects on student learning.

There are a few key gaps in the worked example literature. Nearly all worked example research has examined how worked examples improve performance *after* instruction (see Renkl, 2014). Little is known about the benefits of worked examples as 1) an introduction to novel solution strategies or the 2) potential benefits of pairing worked examples with direct teacher instruction to support mapping. Also, research should examine whether worked examples simply enhance problem-solving skills for students who already have knowledge of the solution strategies or if worked examples can encourage students to attempt novel problem-solving strategies. Moreover, it is necessary to examine if worked examples impact different cognitive skills such as procedural thinking (problem-solving) compared to critical, higher order thinking skills. Finally, worked examples are theorized to reduce students' cognitive load (Paas, 1992). However, what remains unknown is whether worked examples can ameliorate the negative effects of low cognitive resources during high pressure or anxiety provoking learning contexts.

1.2 Overview of Studies

This dissertation examines the influence of worked examples as a means to support learning of sophisticated solution strategies when solving proportional reasoning word problems 1) prior to any formal lesson on the solution strategies and 2) during instruction as a *double-dose* of visual support. Three in-classroom experiments were conducted and the studies build upon one another. The research focuses on youth in low socio-economic, ethnically diverse schools, between upper elementary and middle school, at a crucial transition point in which academic motivation tends to reduce dramatically and where US student performance continues to decrease relative to that of students from other nations (e.g., Eccles & Wigfield, 2002).

In Study 1 (Chapter 2), I report on a test of students' learning of novel solution strategies, and ability to map these to near-transfer problems when students were provided either with packets that include fully worked examples or partially worked examples. These gains are compared to students who did not receive any worked example but were asked to solve the same proportional reasoning problems. Students' accuracy in using two instructed solution strategies that varied in sophistication (i.e., equivalent fraction and unit ratio strategy) across three conditions were compared. To capture a more granular assessment of students' problem-solving gains rather than accuracy alone, students' problem-solving decisions and procedures were qualitatively coded, from their approach and initial setup and procedures to their final solution accuracy. To do so, I report a latent class analysis to quantify the nature of different observed patterns in students' problem-solving behavior, finding that fully worked examples can be used as a tool to introduce novel solution strategies but that fully worked examples support mapping of solution strategies, procedures, and arithmetic most successfully.

Building on the findings from Study 1, in the second study (Chapter 3) I examined the potential benefits of integrating fully worked examples in the classroom with direct teacher instruction. I tested whether a *double-dose* of instruction, that is a high-quality teacher-led lesson paired with worked examples *during* the lesson, supported students' learning of novel solution strategies more than a high-quality lesson alone. In their mathematics classrooms, students watched a previously-recorded, conceptually challenging mathematics lesson on proportional reasoning. Within classrooms, students were assigned to the treatment condition of worked examples or no worked examples. Problem-solving accuracy was examined to determine if providing worked examples supports performance more than high-quality instruction alone. I also examined if worked examples encouraged use of more sophisticated solution strategies and if this *double-dose* of visual support enhances mapping success of procedures from the worked example to the near-transfer problem more than the high-quality lesson alone.

The third study (Chapter 4) builds on Studies 1 and 2 and expands the lens to evaluate learning using a 2 x 2 design that manipulates two levels of visual support (high visual support vs low visual support) as well as two levels of pressure (pressure vs no pressure). The high visual support condition paired worked examples during instruction with other visual supports such as simultaneous presentation and linking gestures. The low visual support condition did not include worked examples, used sequential rather than simultaneous presentation of solution strategies and did not include linking gestures. Pressure and its associated anxiety has been shown to strain students' cognitive resources (Eysenck et al., 2007) and effect test performance (Beilock et al., 2010). We imposed pressure artificially to strain cognitive resources and examine if pressure affected *learning*. Finally, I tested the interactions to examine whether reducing cognitive load via visual support could improve learning outcomes for students with temporarily low processing resources due to the imposed pressure and anxiety. These intervention conditions were explored in two cognitive contexts: 1) higher order thinking about the multiple solution strategies and 2) problem-solving accuracy, approach, and procedural gains on post-tests. I examined these two cognitive contexts as a way to explore how pressure impacts conceptual (i.e., higher order thinking between solution strategies) versus procedural (i.e., problem-solving) learning and engagement. Research findings from all studies and implications are discussed in Chapter 5.

CHAPTER TWO: FULLY WORKED EXAMPLES SUPPORT PROPORTIONAL PROBLEM-SOLVING GAINS MORE THAN PARTIALLY WORKED EXAMPLES

1. Introduction

Mathematics is a subject in which studying examples is particularly advantageous, especially in complex domains like proportional reasoning. A worked example in mathematics is a problem that has been fully completed for students to study the procedures and steps of a solution strategy (Clark et al., 2011). A large body of work has documented the efficacy of presenting fully worked solutions to improve students' math accuracy in educational settings (e.g., Cooper & Sweller, 1987; Sweller & Cooper, 1985; Rittle-Johnson & Star, 2007). However, nearly all worked example research has examined how worked examples provide extra support for problem-solving *after* instruction (see Renkl, 2014). The effects of worked examples as a tool to *introduce* mathematics principles or strategies remains unknown.

Research on worked examples has also often focused on students' accuracy as the primary measure of learning. Surprisingly, little is known about how students approach problem-solving differently both in terms of the strategy employed and the problem-solving procedures. Exploring only the final solution overlooks improvements in students' approach and procedures (van Loon-Hillen et al., 2012). As an example, for many math subjects, different strategy approaches can indicate different levels of knowledge sophistication (Lamon, 1993; Geary et al., 1996). In the domain of proportional reasoning, repeated addition (also called the "build-up" strategy) is a less sophisticated strategy than multiplicative thinking (e.g., the "equivalent fraction" or "unit ratio" strategies); both can be correctly used to solve the same problem, yet they require different arithmetic and conceptual skills. Therefore, research that examines how worked example can promote the use of novel solution strategies and mapping of procedures is necessary. Finally, criticism in the worked example literature has argued that the regularly observed positive effects of worked examples are due to the fact that studies tend to compare worked examples to unsupported problem-solving (Koedinger & Aleven, 2007; McLaren et al., 2006; Mwangi & Sweller, 1998).

To address these gaps in knowledge and insufficient counterfactual conditions, the experiment reported here compared two variations of worked examples: fully worked and partially worked examples and tested their efficacy in introducing students to novel proportional reasoning solution strategies. Partially worked examples remove procedures of the solution strategy for the student to solve and have been found as an alternative effective learning tool especially as students build skills (e.g., Atkinson et al., 2003; Renkl et al., 2004). Even so, a direct comparison between fully worked and partially worked examples to introduce new solutions has not yet been examined. Taking a more traditional approach first, we examined students' accuracy for those who received mathematics packets with problems only (the control group) and compared that to those who received either fully worked examples or partially worked examples. The main goal of this work, however, was to capture a more granular and nuanced understanding of students' problem-solving gains from the worked examples and to examine if providing worked examples fundamentally change students' thinking, approach, and procedures. Thus, I next explored students' ability to map novel solution strategies and their procedures in near-transfer contexts using a detailed analysis of their problem-solving behavior, and tested how that varied by worked example condition using latent class analysis techniques.

1.1 Worked Examples and Multiple Solution Strategies

Learning from worked examples has been a major theme in educational research since the mid-1950s (e.g., Bourne et al., 1964; Bruner et al., 1956). An extensive body of literature in laboratories and classroom settings have shown that basic problem-solving, which often use a means-ends strategy, is not an effective way of learning when compared to instruction that pairs practice problems with worked examples (e.g., Cooper & Sweller, 1987; Owen & Sweller, 1985; Paas & Van Merrienboeer, 1994). When students are presented with a packet of practice problems, students typically employ novice and potentially incorrect strategy solutions.

Fully worked examples are a pedagogical tool to support learning gains; fully worked examples are problems that have been fully completed (i.e., the approach, setup, procedures, and solutions) to demonstrate a correct solution strategy (Atkinson et al., 2000; Renkl, 2014; Clark et al., 2011; McGinn et al., 2015). Students who are provided with fully worked examples are able to better apply instructed procedures as the worked examples free up cognitive resources and highlight the structural aspects of the solution strategy (Cooper & Sweller, 1987; Zhu & Simon, 1987; Carroll, 1994; Ward & Sweller, 1990; Rittle-Johnson & Star, 2007; Salden et al., 2010). In addition, providing students with worked examples decreases the amount of time it takes for learning to occur (Schwonke et al., 2009; Atkinson et al., 2003; Sweller & Cooper, 1985; Carroll, 1994). It also improves students' ability to solve problems that are very similar to those studied, harder than the studied examples, and problems that require transfer to different contexts (e.g., Bentley & Yates, 2017; Sweller & Cooper, 1985; Trafton & Reiser, 1993; Catrambone & Yuasa, 2006).

A key tenant of using worked examples has been to couple the pedagogical approach with teaching and comparing multiple solutions strategies to solve the same problem (Rittle-Johnson & Star, 2007; Gick & Holyoak, 1983; Namy & Gentner, 2002). Teaching multiple solution strategies is used by expert mathematics teachers and is a practice recommended in mathematics education standards (Richland et al., 2007; National Governors Association, Common Core State Standards in Mathematics, 2010). Specifically, comparing different solution methods to the same problems has been found to yield higher gains in conceptual knowledge and procedural flexibility than comparing different types of problems with the same solution types (Rittle-Johnson & Star, 2009). Teaching multiple solutions to the same problem type allows students to grapple with the multiple solutions features and key structural concepts; additionally, it highlights that there are many strategies and procedures that one can employ to solve problems. What is not known from this literature is how worked examples promote the use of novel solution strategies. Can worked examples fundamentally change the way students think about problem-solving? Research that examines whether worked examples can promote use of novel solution strategies is a necessary next step to answer questions such as these.

1.2 Variations in Worked Examples

Variations of worked examples have been tested in laboratory settings; incomplete or partial worked examples are one such variation that has been used to successfully transition students from studying examples to solving problems independently (see Renkl & Atkinson, 2003; Sweller et al., 2003). Students are presented with worked examples that set up the solution strategy but allow the student to engage and solve the procedures. This approach follows the principle that the student can gradually build skills during problem-solving while maintaining sufficient cognitive capacity to focus on understanding the key principles (Booth et al., 2015). Pass (1992) showed that because some students may not deeply study fully worked examples; some students may benefit more from partially worked out problems in which they have to complete steps of the problem-solving procedures and engage in more critical and active thinking and less passive thinking (see also, Van Merrienboer & Krammer, 1990).

There are two competing theories on the effectiveness of partially or fully worked examples as a tool to introduce novel solution strategies. The theory on exploration based instructional models (Hirsh-Pasek et al., 2009) suggest that partial worked examples will lead to the most learning gains. Research shows that engaging in exploratory problem-solving prior to receiving explicit instruction better prepares the child to learn from instruction than "instruct-then-practice" sequences (DeCaro & Rittle-Johnson, 2012; Schwartz et al., 1998; Schwartz & Martin, 2004). When students solve a challenging problem on their own prior to receiving explicit instruction, it leads to greater learning as students gauge their competence, attempt more strategies, and attend more to problem features (DeCaro & Rittle-Johnson, 2012). These findings, coupled with a desirable difficulty theory (Bjork & Bjork, 2011), suggest that creating difficulty for the learner may actually lead to more durable and flexible learning. According to this theory, partially worked examples may promote problem-solving skills more than fully worked examples.

Alternatively, evidence suggests that fully worked examples may be most effective. Atkinson and Renkl (2007) proposed that fully worked examples may be particularly beneficial for students with low prior knowledge or students at the beginning of cognitive skill acquisition. They also proposed that fully worked examples may be especially advantageous in conceptually demanding contexts, such as proportional reasoning. This leaves a critical gap in knowledge; research is needed to examine if fully or partially worked examples promote use of novel solution strategies.

Knowledge level and skills of the learner may also play a role into efficacy of the worked examples (e.g., Mayer, 1999; Mayer & Gallini, 1990). For example, novice learners typically focus on unimportant, surface or perceptual features of the word-problem context that are not relevant to higher level concepts or procedures (Gentner & Jeziorski, 1989; Catrambone & Holyoak, 1989) whereas more experienced learners can attend to the more structural features of the word problem, context, which allows them to use and implement more sophisticated solution strategies (e.g. Chi et al., 1981; Chi and Ohlsson, 2005). Renkl and Atkinson (2007) suggest that once learners attain a certain level of experience or skill level, fully worked examples are redundant may impose an even heavier working memory load without an additional benefit (Tuovinen & Sweller, 1999; Kalyuga et al., 1998). Therefore, worked examples may differentially impact students with different skill level.

The attention a student allocates to the instructional strategies is also a contributing factor problem-solving gains. Worked examples help modulate attention to solution strategies by reinforcing accurate procedural steps during problem-solving and drawing attention to the structural features of the solution strategy (e.g., Ward & Sweller, 1990). Therefore, assessing how worked examples help students notice and attend to procedural and structural features of solution strategies while taking into account individual difference factors of prior knowledge and attentional control is a crucial research goal.

1.3 Importance of Proportional Reasoning

Worked examples may be particularly beneficial in complex learning environments such as in proportional reasoning. Proportional reasoning is the ability to understand the multiplicative relationship between quantities and phenomena as systems of structured relationships that can be aligned, compared, and mapped together (Cramer & Post, 1993; Boyer et al., 2008; Richland & Simms, 2015). Proportional reasoning skills are considered a required skill in mathematics education as it is foundational for higher level mathematics and algebra, as well as for navigating everyday situations (e.g., National Mathematics Advisory Panel, 2008; Siegler et al., 2013). Despite this domain being deemed so crucial that the National Council of Teachers of Mathematics (1989) stated that it deserves "whatever time and effort must be expended to assure its careful development" (cited in Cramer & Post, 1993), many students experience pervasive difficulties with proportional reasoning (National Research Council, 2004).

Enhancing students' proportional skills is necessary to advance their overall ability to learn and retain mathematics skills. In order to evaluate this domain further, research should further explore the relevance of worked examples for student learning, including: a) the various strategies and levels of sophistication students can employ to solve proportional reasoning problems, b) students struggle with the key aspect (i.e., multiplicative thinking) of proportionality and c) as a tool to reduce students cognitive load in complex and multistep solution strategies.

First, research has identified a developmental pattern of proportional reasoning, which is associated with increasingly sophisticated problem-solving strategies (Lamon, 1993; Nabors, 2003; Parish, 2010). Poor understanding of proportional reasoning is associated with students' use of guessing or a visual judgment (Parish, 2010). An additive stage occurs when students use a subtraction or addition strategy, which reflects constant additives rather than an understanding of multiplicative relationships (Parish, 2010; Nabors, 2003). With increasing sophistication, students can show a basic understanding of proportional reasoning by using a repeated pattern strategy (e.g., buildup) in which students repeatedly add or subtract a number to each part of the ratio. The next stage of understanding is when students begin to use multiplication in place of repeated addition (Nabors, 2003). In this stage, students can also setup fractions and multiply or divide the numbers by the same value (i.e., equivalent fraction strategy). Even more sophisticated students can determine the unit rate of the proportions and subsequently apply that same rate to any desired quantity (i.e., unit ratio strategy), made up of division and multiplication (Nabors, 2003).

Second, research on proportional reasoning highlights a major procedural learning difficulty for students learning proportional reasoning. In the early learning stages, students typically use additive rather than multiplicative solutions when using an equivalent fraction strategy (e.g., to solve 6:14 = ?:35 they find the difference between the front relation and subtract it from 35; Fuson and Abrahamson, 2005; Sowder, 2007). As students gain more domain expertise, they learn that proportional reasoning involves understanding the multiplicative relationships between quantities (a/b =c/d) and is a form of relational reasoning that characterizes important structural relationships (Cramer & Post, 1993; Lesh et al., 1988). Basic arithmetic and procedural fluency of complex mathematical tools to execute arithmetic and computational operations can be obstacles to students' understanding of more complex problem-solving (Hecht et al., 2001).

Third, students must complete numerous procedures to arrive at the correct solution. In the middle school classroom, ratio and proportion problems are typically in the form of word problems (Jitendra et al., 2009). In order to solve real-word problems, students need to identify relevant structural relations among the problem characters, create a schema for problem-solving, generate a solution strategy, determine the appropriate operation, and develop an understanding of the broader goal of the word problem (Desoete et al., 2003; Verschaffel et al., 2009; Briars & Larkin, 1984). Therefore, problems that require students to take multiple steps to arrive at a solution and require students to remember, compare and manipulate multiple relationships at one time prove to be highly challenging.

For these reasons, proportional reasoning provides a clear, strong and measurable context to study pedagogical practices and offers a unique opportunity to examine variants of worked examples.

1.4 Current Study

The current research was designed to explore the processes by which students can learn novel, sophisticated solution strategies from worked examples. In this study, 5th and 6th grade students were introduced to two sophisticated solution strategies for solving word problems involving proportions - equivalent fraction and unit ratio strategies. Students in these grades were chosen as they have been introduced to use these strategies and solve proportional reasoning problems; however, they typically have not yet been formally introduced to these solution strategies (Common Core State Standards in Mathematics, 2010). We examined students' learning and how they employed novel procedures when provided with instructional packets that include i) fully worked (FW) examples or ii) partially worked (PW) examples, and compared problem solving to students in the iii) problems only (PO), control condition. For students in the PO condition, they were instructed to solve problems using any preferred strategy. This condition gives a clear baseline for the approaches and procedures students employ, without support, when solving proportional reasoning problems. When FW Examples were provided, students were given all of the problem-solving steps (i.e., approach, setup, procedures, and solutions). In the PW Example condition, students were presented with partially worked examples such that the approach and setup of the strategy were provided but students were required to complete the key procedure on their own. By examining students' success on near-transfer problems we were able to capture whether students produced the key component of proportional reasoning (multiplicative thinking).

The two solution strategies were chosen as they are the most sophisticated approaches to solving proportional reasoning problems that are regularly taught to children in grade six but are accessible to students in grade 5, and are both procedurally distinct and conceptually distinct. The difficulty level does not differ across the solutions, however, with the equivalent fraction strategy being typically taught first and involving less challenging division procedures than the unit ratio strategy. Testing the effects of worked examples on these two solution strategies allowed this study to elucidate whether efficacy of worked examples varied based on the sophistication of the solution strategy. Specifically, we assessed if worked examples differentially impacted the use of instructed solution strategies more for equivalent fraction versus unit ratio strategies.

In addition to assessing students' accuracy, a main aim of this paper was to capture learning gains beyond accuracy. In order to capture the nature and sequence of students' problem-solving, students qualitative written work was coded to determine if and to what extent students' attempted the novel strategy solution and mapped strategies and procedures in near-transfer contexts. Problem-solving steps including the i) approach of the solution strategy, ii) setup of the solution strategy, iii) procedures of the solution strategy, and iv) final solution accuracy were coded to capture students' full problem-solving ability. In order to analyze these problem-solving steps latent class analyses (LCA) was employed to identify and characterize the full sequence of steps students take to solve proportional reasoning problems. The LCA allows us to estimate distinct patterns of problem-solving steps, and group students by their problem-solving abilities. We examined also how worked example condition and individual differences in cognitive functioning predict problem-solving patterns. Finally, we explored whether individual differences in students' cognitive and mathematical abilities predicted use of solution strategies and success.

We hypothesized that i) worked examples encourage use of novel solution strategies, ii) FW Examples lead to the highest frequency of attempt for instructed strategies, the most success in problem-solving procedures, and higher overall accuracy compared to those in the PW Example or PO control condition, iii) PW Example condition will have more procedural errors at the point in which the worked example was left blank (i.e., the multiplicative step) compared to students in the FW Example condition, iv) students who have more prior knowledge of proportions (i.e., familiarity with the instructed solution strategies, either conceptually and procedurally) will perform more accurately across the problem-solving procedures and iv) the trends will look similar for more novel and sophisticated solution strategies (unit ratio strategy), but for unit ratio problem-solving there will be lower instances of attempt for instructed strategy, success in problem-solving procedures, and overall accuracy.

2. Methods

2.1 Participants

Fifth- and sixth-grade students from two K-8 schools in the Chicagoland area with primarily underrepresented Latinx or Black youth in low to middle income neighborhood schools were recruited to participate in this study. They were chosen because students at this age have been introduced to the math concepts necessary to solve proportional reasoning problems; however, they have not yet been formally introduced to solution strategies for solving proportional reasoning word problem (Common Core State Standards in Mathematics, 2010). Three students opted out of the study and two students were absent for more than half of the study; these students were excluded from analyses. One-hundred and eighty-eight students participated across the two schools. The final sample included 136 students in the 5^{th} grade and 52 students in the 6^{th} grade. Participants were 106 girls, 73 boys, and 9 unidentified. Parents and guardians were informed of the study a few weeks prior to data collection and were provided the opportunity to opt their child out. We also obtained children's written assent prior to data collection.

2.2 Teacher and Experimenter Demographics

All of the children's regular mathematics teachers were present in the classroom throughout the study and at least two experimenters were in the classroom during the study.

2.3 Design and Procedures

The research study was conducted during one in-classroom visit. Procedures were administered in their typical mathematics classroom, during their mathematic class period, and students were alongside their peers. First, students completed a group-administered *Prior Knowledge and Arithmetic Assessment*. To minimize variability across schools and teachers, participants were randomly assigned to the intervention within each classroom. Students randomly received a *Proportional Reasoning Problem-solving Packet* with FW Examples (n= 62), PW Examples (n= 64), or problems only (PO) control (n= 62). Finally, students completed a measure of executive functioning, the d2 test of attention (Brickenkamp & Zillmer, 1998), followed by a short demographics survey.

2.4 Materials

2.4.1 Proportional Reasoning Problem-solving Packet

Students completed a *Proportional Reasoning Problem-solving Packet* which included instruction and assessment. Students were assigned to one of three conditions: 1) FW Examples condition, 2) PW Examples condition, or 3) the PO control condition. In all three conditions, the worksheets contained three unique problems in which students were to find the values of an unknown number. Students solved the first three problems using one strategy. Then students were asked to solve the same three problems using a different strategy.

In the experimental conditions, the first and fourth problem were replaced with worked examples of two different solution strategies from two exemplar students. First, students either received a FW Example or a PW Example of Estella's equivalent fraction strategy (more detail regarding the instructed examples are presented below). Then, students were asked to solve two near-transfer problems using the equivalent fraction strategy. These near-transfer problems are referred to as the equivalent fraction problems. Next, the same three problems were presented. For the first problem, students were provided with either a FW Example or PW example of Eric's unit ratio strategy. The same two near-transfer problems were then shown again and students were asked to solve the problems this time with the unit ratio strategy. These near-transfer problems are referred to as the unit ratio problems.

2.4.1.2 Fully Worked Example Condition

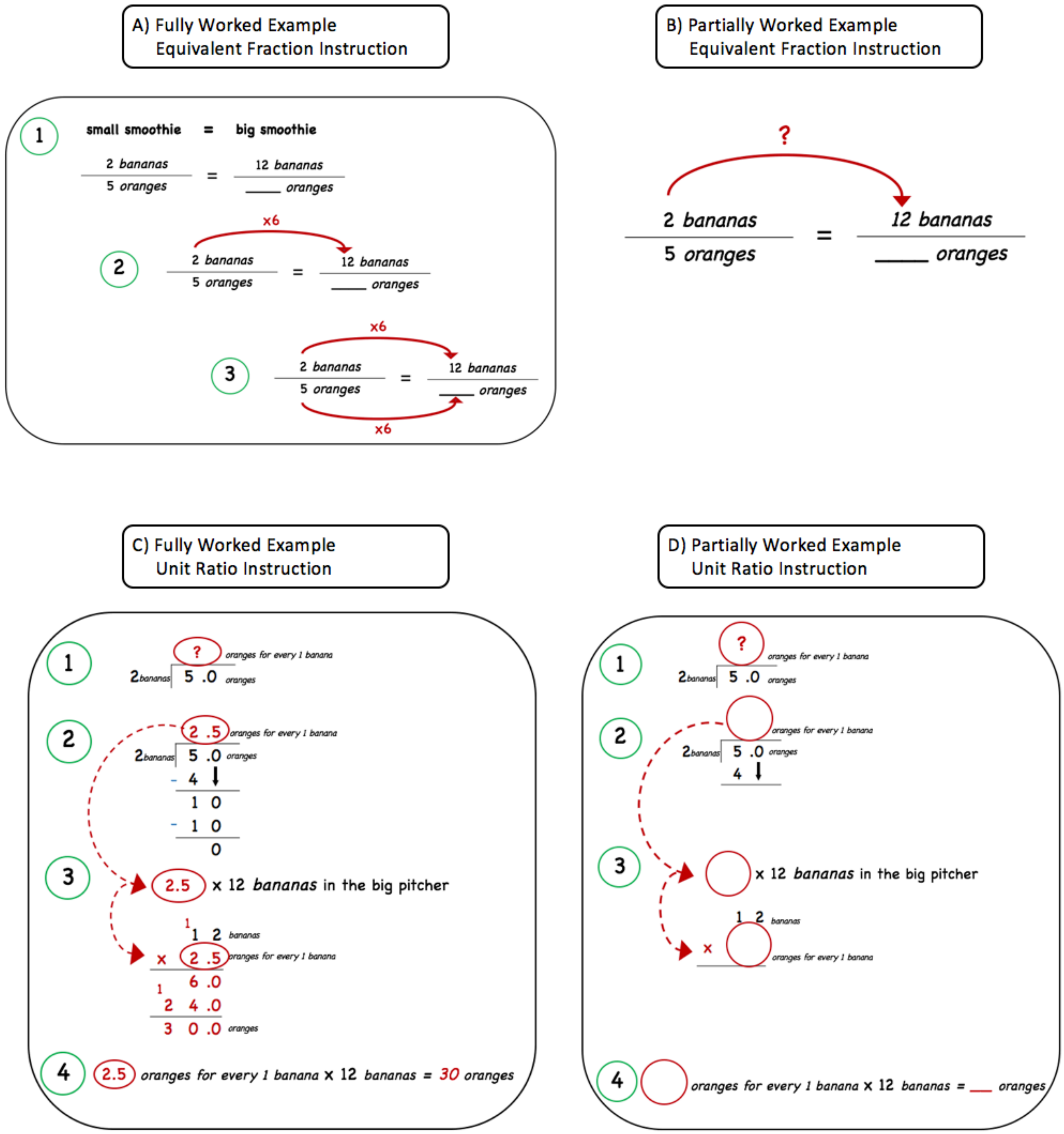
Students in the FW Example condition studied Estella's fully worked example of the equivalent fraction strategy and studied Eric's fully worked example of the unit ratio strategy (Figure 2.1, A and C). For the equivalent fraction strategy, the fully worked example included: 1) attempt of the equivalent fraction strategy, 2) structure setup, 3) number setup, 4) link, and 5) final solution accuracy. For the unit ratio strategy, the fully worked example included: 1) attempt of the unit ratio strategy, 2) division setup, 3) division numbers setup, 4) division accuracy, 5) link, 6) multiplication accuracy, and 7) final solution accuracy.

2.4.1.3 Partially Worked Example Condition

Students in the PW Example condition were provided with Estella's partially worked equivalent fraction strategy and Eric's partially worked unit ratio strategy (Figure 2.1, B and D). For the equivalent fraction strategy, the partially worked example included: 1) attempt of the equivalent fraction strategy, 2) structure setup, and 3) number setup. An arrow to link the known unit in the small recipe to the large recipe was provided, but the arithmetic (i.e., multiplication) to find the link was not provided; students were to determine this key procedure on their own. For the unit ratio strategy, the worked example included: 1) attempt of the unit ratio strategy, 2) division setup, 3) division numbers setup, and 4) link. Students were to solve the long division piece and decimal multiplication piece of the problem independently.

Figure 2.1

Fully Worked and Partially Worked Example Instruction for Equivalent Fraction and Unit Ratio Strategies



Note. Fully worked examples for the equivalent fraction strategy (A) and unit ratio strategy (C) were provided for students in the FW Example condition to study. Partially worked examples for the equivalent fraction strategy (B) and unit ratio strategy (D) were provided for students in the PW Example condition to solve.

2.4.1.4 Problem Only, Control Condition

The PO packet contained the same six problems as the FW and PW Example packet, but students were not shown any worked examples. Instead students were told to use any strategy they preferred for the first three problems. Then, they were asked to solve the same three problems using a different strategy of their choice.

2.4.1.5 Proportional Reasoning Problem-solving Coding Scheme

We examined learning gains from the worked examples using two different approaches. First, students' overall accuracy on proportional reasoning word problems was examined. Then, taking a more fine-grained approach, we examined the nature and sequence of students' problem-solving decisions, from their initial attempt of the solution strategy to their procedures and ultimately to their final solution accuracy. Study measures and coding protocol allowed us to examine a more nuanced approach to learning gains and explicitly examine how students are mapping procedures from the worked examples to on near-transfer problems which are similar and share features of the instructed problems.

The complete coding protocol is available on OSF and outlines the development of the codes, coding protocols and reliability checking. A brief outline of the coding protocol for the equivalent fraction strategy and unit ratio strategy is outlined next. For the all proportional reasoning word problems, students qualitative written work was coded to determine if and to what extent they followed the instructed worked example procedures.

Qualitative codes were used to gather quantitative data about the emergent patterns in problem-solving. To examine students' use of the equivalent fraction strategy on near-transfer problems, qualitative codes assessed if students: 1) attempted the instructed strategy, 2) structurally setup the problem as instructed, 3) numerically set up the problem as instructed, 4) linked known recipe units through multiplication or division, and 5) arrived at the accurate answer. A summary of the coding scheme for equivalent fraction word problems can be found in Table 2.1.

To examine students' use of the unit ratio strategy on near-transfer problems, codes assessed if students 1) attempted the instructed strategy, 2) structurally setup the division step as instructed, 3) numerically set up division step as instructed, 4) arrived at the accurate answer for the mid-point division step, 5) linked the unit ratio solution to the known large unit through multiplication or division, 6) arrived at the accurate mid-point answer for the multiplication step, and 7) arrived at the accurate final answer. A summary of the coding scheme for unit ratio word problems can be found in Table 2.1.

To establish inter-coder reliability, 20% of the data was coded by two research assistances. Krippendorff's alpha was used to assess inter-coder reliability for each variable coded (Hayes & Krippendorf, 2007). For the coding of the variable to be considered reliable it was required that the Krippendorff's alpha (an index that accounts for level of measurement and agreement expected by chance and is known to be conservative to be 0.70 or higher. Reliability for variables in the PO control condition were $\alpha = 0.96$, variables in the PW Example condition were $\alpha = 0.98$, and variables in the FW Example condition were $\alpha = 0.95$.

Table 2.1
Equivalent Fraction and Unit Ratio Strategy Coding Scheme for Problem-Solving Steps

Problem-solving Step	**Attempted Instructed Strategy**
Equivalent Fraction Strategy	
Attempt	Attempted equivalent fraction strategy (i.e., fraction or ratio) with numbers from problem *and* attempted to find relationship between numbers in problem
Structure Setup	Attempted fraction setup whether numbers are placed correctly or incorrectly
Numbers Setup	Compared unit of small recipe to unit of large recipe
Link	Linked known units from small recipe to large recipe through multiplication/division and applied link to unknown unit
Accuracy	Accurate solution
Unit Ratio Strategy	
Attempt	Attempted to find the unit ratio *and* attempted to apply that number (i.e., division and multiplication)
Division Setup	Attempted division first, indicated by any division symbol (e.g., decimal division or mixed fraction)
Division Numbers Setup	Divided numbers from the small unit to create the unit ratio
Division Accuracy	Accurate division solution
Link	Linked the unit ratio solution to large unit through multiplication/division (even if final number from the unit ratio division step was inaccurate)
Multiplication Accuracy	Accurate multiplication solution
Accuracy	Accurate solution

Note. The equivalent fraction and unit ratio coding schemes were applied to all proportional reasoning word problems in all packet conditions. Problem-solving steps are used as the key indicators for the LCA analysis.

2.4.2 Prior Knowledge Assessment

All students completed a pretest assessment to measure prior knowledge for solving proportional reasoning word problems. Students were asked to solve two proportional reasoning word problems using any strategy of their choice. The missing value could be solved by comparing like terms or across units (e.g., A smoothie recipe calls for 2 bananas and 5 oranges. To make a bigger pitcher of the smoothie with 12 bananas, how many oranges should you add?). The same coding scheme was applied for the prior knowledge of proportional reasoning word problems.

2.4.3 Arithmetic Assessment

Arithmetic skills that are needed to solve proportional reasoning word problems using an equivalent fraction strategy and unit ratio strategy were also assessed. Students solved four missing-value ratio equivalence problems to assess their understanding for multiplicative relationships (Cramer & Post, 1993). The missing value in the ratios were the back term numerator and back term denominator. Potential solution strategies for comparison also varied: numerator to numerator (e.g., $\frac{5}{6} = \frac{15}{\ }$), denominator to denominator (e.g., $\frac{2}{3} = \frac{\ }{12}$) or front term numerator to front term denominator (e.g., $\frac{2}{12} = \frac{3}{\ }$). Lastly, students solved one long division problem.

2.4.4 Attentional Control

The d2 Test of Attention is a group-administered, pen and paper measure which assesses students sustained levels of selective attention and attentional control (Brickenkamp & Zillmer, 1998; Rhonda & Ross, 2005). The assessment is normed with US and German children, adolescents and adults. Under a time pressure, participants are asked to search and cross off target characters (i.e., "d"s with two dashes surrounding it either above or below) from perceptually similar distractors (e.g., "d"s with one dash, "p"s with two dashes). The outcome of this task is the total number of items processed minus errors (TN-E). Across the literature, internal consistency is high ($\alpha \geq 0.8$) and test test-retest reliability is high ($\alpha > 0.8$; Clark, 2005). d2 TN-E scores correlate with other measures of attention and executive functioning: Stroop and Tower of London which support the validity of using this measure (Clark, 2005).

2.5 Analytical Plan to Examine Problem-Solving Gains

In one classroom intervention, we tested how presenting students with fully and partially worked examples for two novel and sophisticated solution strategies (equivalent fraction and unit ratio strategy) promote gains in problem-solving compared to students who did not receive any worked examples. First, we examined the impacts of intervention condition on problem-solving accuracy. For each strategy solution, we conducted two ordinal logistic regressions. The first set of regressions examined accuracy predicted by condition with fixed effects of classroom; the second set of regressions examined accuracy predicted by condition and controlled for individual differences of attentional control, prior knowledge of proportions, and arithmetic skills, with fixed effects of classroom.

To capture a more fine-grained assessment of students' problem-solving gains, we examined the nature and sequence of students' problem-solving from strategy approach to procedures to final solution. To do so, a three-step latent class analyses (LCA) was used. Due to the different problem-solving steps for equivalent fraction and unit ratio strategy, we conducted one LCA for equivalent fraction problems and one LCA for unit ratio problems in Latent GOLD 5.1 (Statistical Innovations, 2016). The goal of this approach was to identify and characterize the full sequence of steps students take to solve proportional reasoning problems, rather than accuracy only. Additionally, we did not anticipate that the intervention of worked examples would impact students in the same way. This latent class analysis allows us to estimate distinct groups of students and their distinct problem-solving abilities from approach to final solution accuracy.

2.5.2 Latent Class Analysis for Equivalent Fraction and Unit Ratio Problems

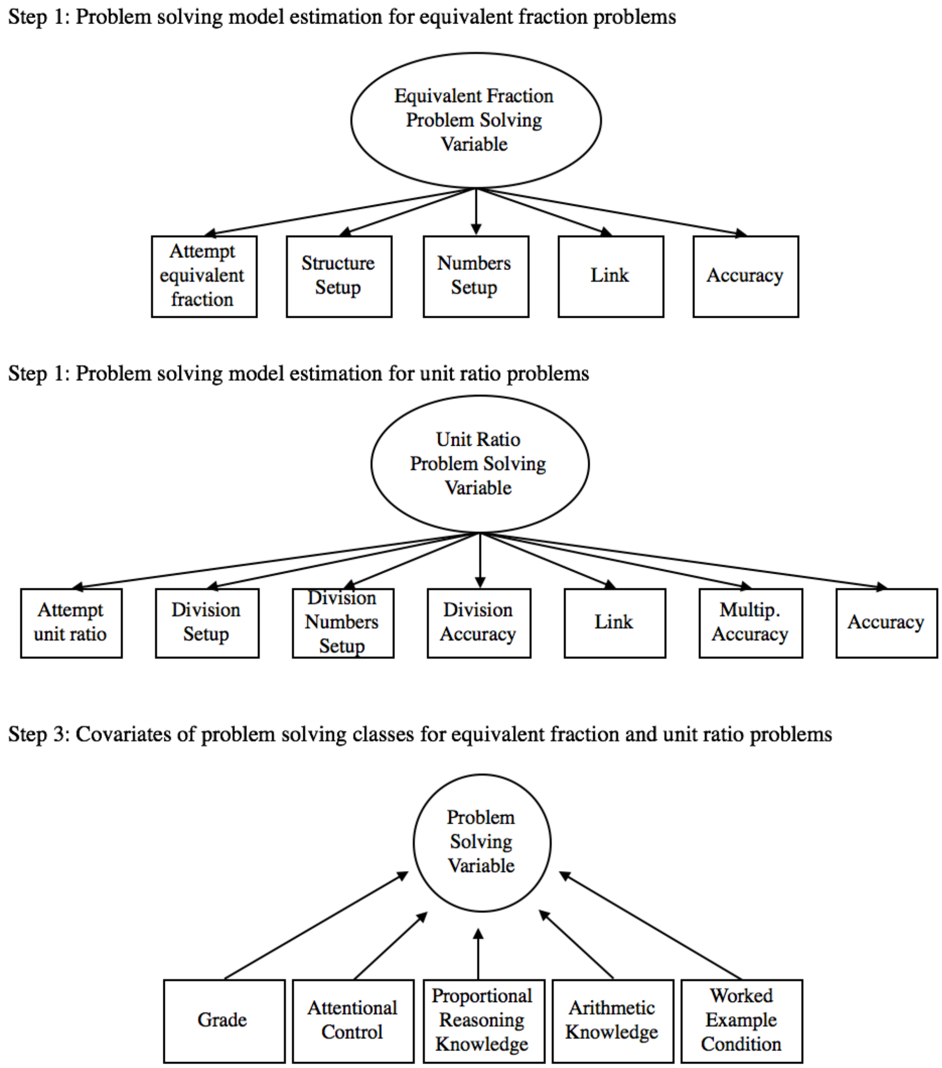
LCA classification methods aim is to identify the smallest number of distinct groups that represent patterns in the data; groups are represented by a categorical latent variable (Bray, Lanza,

& Collins, 2010). This contrasts with arbitrary cut-point methods such as median split, which both introduces bias into the analysis and is a-theoretical (Irwin & McClelland, 2003; Maxwell & Delaney, 1993). LCA also is useful for educational research, as learning often follows non-linear patterns (e.g., students do not progress through the groups linearly) and LCA can account for this (see Hickendorff, Edelsbrunner, McMullen, Schneider, & Trezise. 2018).

A three-step LCA involves analyzing data in three steps. First, multiple models are run with increasing numbers of latent classes (estimated with 100 sets and 500 iterations, using random seed) to determine how many distinct problem-solving groups there are. Statistical and theoretical criteria were used to compare seven models to determine the latent class model that best fit the data (Collins & Lanza, 2010; Vermunt & Magidson, 2016). Second, the optimal model is identified and individuals are classified into the groups. Variance was unconstrained across groups. This was a theoretical decision as we expected some groups would show more variance across steps and some groups would have little variance across steps. Third, the relationship between the latent classes and external variables are examined. This three-step approach allows for the methods to correct for error bias that is introduced when individuals are assigned into latent classes (Bakk, Tekle, & Vermunt, 2013; Bolck, Croon, and Hagenaars 2004; Vermunt, 2010).

The indicator variables used to characterize latent classes are the problem-solving steps. For the equivalent fraction strategy, we used the five problem-solving steps from the qualitative coding scheme as the indicators for the model (i.e., attempt, structure setup, numbers setup, link and accuracy; see Figure 2.2). Students were asked to solve two near-transfer assessment problems using this strategy; therefore, the count variables for each indicator ranged from 0-2. For the unit ratio LCA, the seven problem-solving steps from the qualitative coding scheme were used as indicators: attempt, division setup, numbers setup, division accuracy, link, multiplication accuracy, and problem accuracy (see Figure 2.2). All procedures were identical to the equivalent fraction LCA. For both the equivalent fraction LCA and the unit ratio LCA, the indicators used in the analyses were strongly related to the latent class variable, indicating that the indicators (i.e., qualitative coding scheme of problem-solving steps) were of high quality.

Figure 2.2
Conceptual Diagram and the Analytic Model



Note. Step 1 Model estimation: Latent Class Analysis (LCA) identifying the optimal model for equivalent fraction problems and unit ratio problems. Order of problem-solving indicators reflect the problem-solving steps (attempt instructed strategy, setup, link, accuracy). Step 3: Examination of the relationship between problem-solving groups and covariates. Step 2 of the analyses is not represented.

3. Results

3.1 Baseline data

A one-way ANOVA was conducted first to establish that the randomization was successful and that there were no differences between conditions. At pretest, there were no differences of prior knowledge of proportions ($F(2,185) = 0.04$, $p = 0.96$), arithmetic skills ($F(2,185) = 0.75$, $p = 0.48$), or attentional control ($F(2,177) = 0.16$, $p = 0.85$) between conditions.

3.2 Intervention Effects on Accuracy

Before considering students' overall problem-solving skills using the equivalent fraction and unit ratio strategy, we examined performance accuracy by condition for each solution strategy. In the PO condition, students were asked to solve the two proportional reasoning problems using one preferred strategy and then asked to solve the same two problems using a different strategy. At baseline, without any support, students in the sample were equally likely to solve 0, 1 or 2 problems correctly (Table 2). When asked to use a different strategy for the same problems, most students did not have a second solution strategy to employ. Without pedagogical support from worked examples, this condition highlights that students had some familiarity with proportional reasoning strategies but a limited understanding.

Students in the PW Examples and FW Examples conditions were asked to solve the first two problems using the equivalent fraction strategy and then the same problems using the unit ratio strategy. As shown in Table 2, students in the FW Example condition were most likely to solve both problems correctly for the equivalent fraction strategy and unit ratio strategy although students in the PW Example condition performed high as well.

Table 2.2
Proportion of Students' Accuracy for Equivalent Fraction and Unit Ratio Problems by Condition

	Problem Solution Accuracy							
	equivalent fraction problems				*unit ratio problems*			
	0 Correct	1 Correct	2 Correct	*Total*	0 Correct	1 Correct	2 Correct	*Total*
Problems Only	39%	26%	35%	*100%*	61%	0%	39%	*100%*
PW Examples	36%	19%	45%	*100%*	48%	22%	30%	*100%*
FW Examples	21%	19%	60%	*100%*	32%	23%	45%	*100%*

Note. Students were to solve two proportional reasoning word problems. Therefore, students could have 0, 1 or 2 correct solution answers.

For the equivalent fraction strategy problem-solving, we examined three research questions regarding accuracy: (i) whether the FW Example condition had higher accuracy compared to the PW Example condition and PO condition, (ii) whether the PW Example condition had higher accuracy compared to the PO condition, and (iii) whether children's individual differences in prior knowledge and attention control influenced students' accuracy scores using the equivalent fraction strategies.

3.2.1 Intervention Effects on Accuracy for Equivalent Fraction Problem-solving

To address the first question, we conducted an ordinal logistic regression model with classroom fixed effects to predict students' accuracy (0,1,2) from the experimental condition (see Table 2.3, Models 1 and 2). Children were more likely to have higher accuracy if they were in the FW Example condition than the PO control condition. Accuracy scores did not differ between those in the FW Example condition and PW Example condition but were trending toward significance. The identical models were run to compare the PO condition and PW Example condition; there were no significant differences in performance between the two groups, $p > 0.05$ (see Figure 2.3).

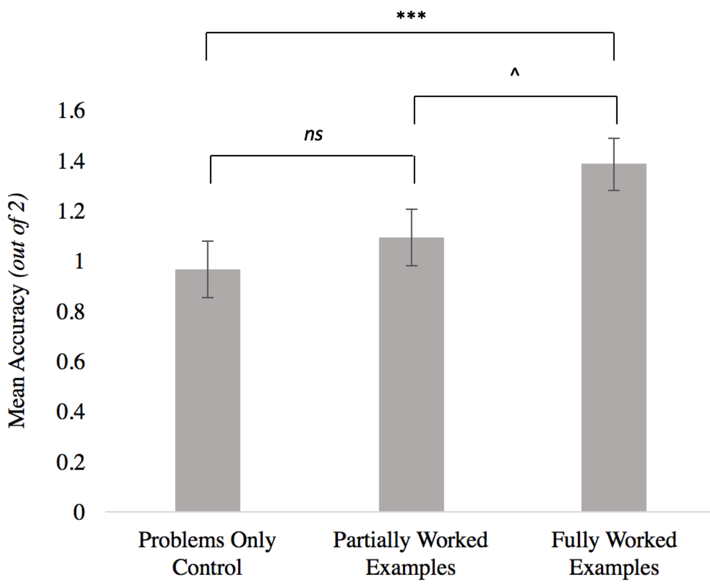
To address the third research question regarding individual differences, we ran an ordinal regression model with classroom fixed effects to predict students' accuracy (0,1,2) from experimental condition and examined the impact of individual differences of attentional control, prior knowledge of proportions, and prior arithmetic skills (see Table 2.3, Model 2). There was a significant positive effect of attentional control and prior knowledge of proportions on students' performance accuracy. This suggests that students who have more attentional control may be able to attend to the conceptual and procedural features of the worked examples and gain more than those with less attentional control.

Table 2.3
Equivalent Fraction Model Specifications for Accuracy by Condition using Classroom Fixed Effects

	Model 1					Model 2				
	exp (β)	β (logits)	*SE*	Wald z	*p*	exp (β)	β (logits)	*SE*	Wald z	*p*
PO Condition	**.36**	**-1.03**	**.36**	**-2.86**	**.001**	**.31**	**-1.17**	**.40**	**-2.92**	**.001**
PW Example Condition	*.51*	*-.67*	*.36*	*-1.87*	*.06*	*.48*	*-.74*	*.41*	*-1.82*	*.07*
Attentional Control	-	-	-	-	-	**1.84**	**.61**	**.34**	**3.55**	**.001**
Knowledge of Proportions	-	-	-	-	-	**2.72**	**1.00**	**.34**	**4.85**	**.001**
Arithmetic Skills	-	-	-	-	-	*1.36*	*.31*	*.08*	*1.81*	*.07*

Note. Model outcomes of performance accuracy for equivalent fraction problems. Students solved two problems and were instructed to use the equivalent fraction strategy (PW and FW Example condition) or a preferred solution strategy. Ordinal outcome is 0 correct, 1 correct, 2 correct. Model 1 regresses condition and controls for fixed effects of classroom. Model 2 regresses condition, covariates and controls for fixed effects of classroom.

Figure 2.3
Fully Worked Example Students Outperformed those in Problem Only Condition using Equivalent Fraction Strategy



Note. Model controls for attentional control, knowledge of proportions, arithmetic skills, and classroom fixed effects, $p < 0.001$ is represented by ***, $p < 0.10$ is represented by ^.

3.2.2 Intervention Effects on Accuracy for Unit Ratio Problem-solving

Next these same preliminary questions regarding accuracy for the more sophisticated strategy, the unit ratio strategy were examined to assess: (i) whether the FW Example condition had higher performance compared to the PW Example condition and PO condition, (ii) whether the PW Example condition had higher performance compared to the PO condition, and (iii) whether children's individual differences in prior knowledge and attention control influenced students' accuracy scores using the unit ratio strategy.

To address the first question, we conducted an ordinal logistic regression model with classroom fixed effects to predict students' accuracy (0,1,2) from the experimental condition (see Table 2.4, Models 1 and 2). Children in the FW Example condition had higher accuracy scores than those in the PO condition. Accuracy scores did not differ between those in the FW Example condition and PW Example condition. The identical models were run to compare the PO condition and PW Example condition; there were no significant differences in performance between the two groups, $p > 0.05$ (see Figure 2.4).

To address the third question, we ran an ordinal regression model with classroom fixed effects to predict students' accuracy (0,1,2) from experimental condition and examined the impact of individual differences of attentional control, prior knowledge of proportions, and prior arithmetic skills (Table 2.4, Model 2). There was a significant positive effect of attentional control and prior knowledge of proportions. This suggest that students who have more attentional control may be able to attend to the conceptual and procedural features of the worked examples and gain more than those with less attentional control.

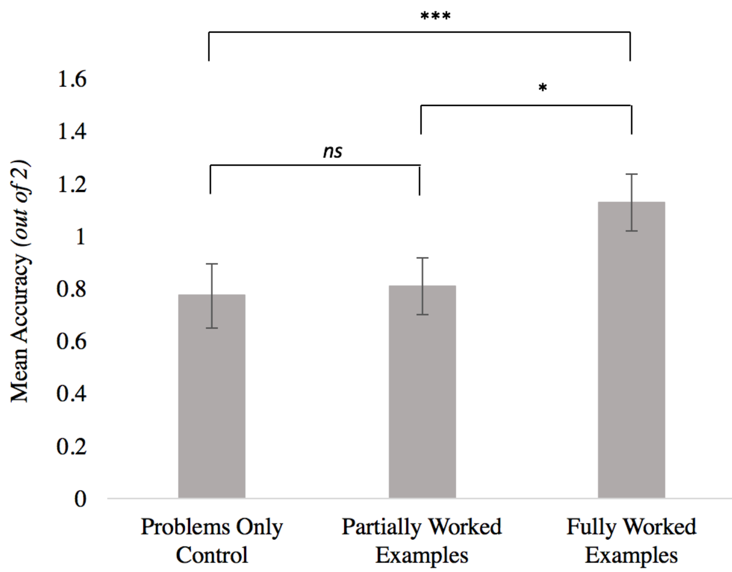
Table 2.4
Unit Ratio Model Specifications for Accuracy by Condition using Classroom Fixed Effects

	Model 1					Model 2				
	exp (β)	β (logits)	*SE*	Wald z	*p*	exp (β)	β (logits)	*SE*	Wald z	*p*
PO Condition	**.43**	**-.85**	**.36**	**-2.33**	**.02**	**.39**	**-.95**	**.40**	**-2.41**	**.02**
PW Example Condition	*.52*	*-.65*	*.34*	*-1.91*	*.06*	**.47**	**-.76**	**.38**	**-2.03**	**.04**
Attentional Control	-	-	-	-	-	**2.83**	**1.04**	**.34**	**3.08**	**.001**
Knowledge of Proportions	-	-	-	-	-	**3.68**	**1.30**	**.30**	**4.37**	**.001**
Arithmetic Skills	-	-	-	-	-	1.14	.13	.08	1.59	.11

Note. Model outcomes of problem accuracy for unit ratio problems. Students solved two problems and were instructed to use the unit ratio strategy (PW and FW Example condition) or a different solution strategy. Ordinal outcome is 0 correct, 1 correct, 2 correct. Model 1 regresses condition and controls for fixed effects of classroom. Model 2 regresses condition, covariates and controls for fixed effects of classroom.

Figure 2.4

Fully Worked Example Students Outperformed those in Partial Worked Example and Problem Only Condition using Unit Ratio Strategy



Note. Model controls for attentional control, knowledge of proportions, and arithmetic skills, $p < 0.001$ is represented by ***, $p < .05$ is represented by *.

3.3 Latent Class Analysis for Problem-solving Steps

Next, we conducted a latent class analysis (LCA) three-step approach to identify and characterize patterns of students' problem-solving approaches and procedures. These groups are characterized as a function of grade, attentional control, prior knowledge, arithmetic knowledge and worked example condition in Latent GOLD 5.1. We conducted two LCA analyses: one for equivalent fraction problems and one for unit ratio problems. First we will describe the latent class models for equivalent fraction problem-solving steps and subsequently describe the latent class models for the unit ratio problem-solving steps.

3.4.1 Step 1: Model Fit for Equivalent Fraction Problems

Model fit information for equivalent fraction problems can be found in Table 2.5. The four group model was best based on the BIC, CAIC, and AWE. The three-group model and four-group models are best for the ICL-BIC. Entropy values for the four-group model was high, indicating clear classification of individuals into latent groups. A conditional bootstrap procedure was conducted to compare the three-, four-, and five-group models (Vermunt & Magidson, 2016). The bootstrap likelihood difference analysis shows significantly more information was gained from the four group model than the three group model (Bootstrap 2LL Diff = 40.58, $p < .001$) and the four group model than the five group model (Bootstrap 2LL Diff = 18.75, $p = .004$). Therefore, the four-group model was selected for equivalent fraction problem-solving because of its statistical fit, parsimony, and theoretical interpretability in characterizing patterns of proportional reasoning problem-solving.

Table 2.5
Model Fit Indicators for LCA of Equivalent Fraction Problem-Solving

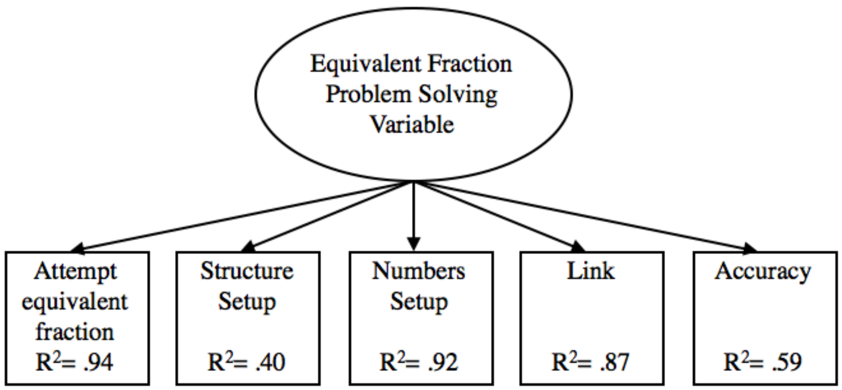
	LL	BIC	CAIC	AWE	ICL-BIC	Entropy R2
1-Group	-940.95	1934.26	1944.26	2016.63	1934.26	1.00
2-Group	-696.03	1475.84	1491.84	1617.49	1485.71	0.96
3-Group	-597.19	1309.59	1331.59	1506.30	1325.10	0.96
4-Group	**-576.90**	**1300.43**	**1328.43**	**1558.53**	**1327.91**	**0.94**
5-Group	-567.53	1313.09	1347.09	1617.10	1337.06	0.96
6-Group	-563.02	1335.50	1375.50	1688.17	1358.72	0.96
7-Group	-558.19	1357.26	1403.26	1766.38	1387.51	0.95

Note. Bold text indicates the selected model. LL, Log Likelihood; BIC, Bayesian Information Criterion; CAIC, Consistent Akaike Information Criteria; AWE, Average Weight of Evidence; ICL-BIC, a version of Integrated Classification Likelihood.

3.4.2 Step 2: Characterization of Equivalent Fraction Problem-solving Groups

The cluster loadings of each problem-solving indicator can be found in Figure 2.5. The loadings are obtained using a linear approximation of the class specific response probabilities (Vermunt & Magidson, 2005). For equivalent fraction problem-solving, accuracy is not the most influential indicator to class membership. Rather, attempt equivalent fraction (i.e., approach), number setup, and link are most predictive of class memberships.

Figure 2.5
Cluster Loadings of the Problem-Solving Indicators for Equivalent Fraction Problems Estimated in the Latent Class Analysis



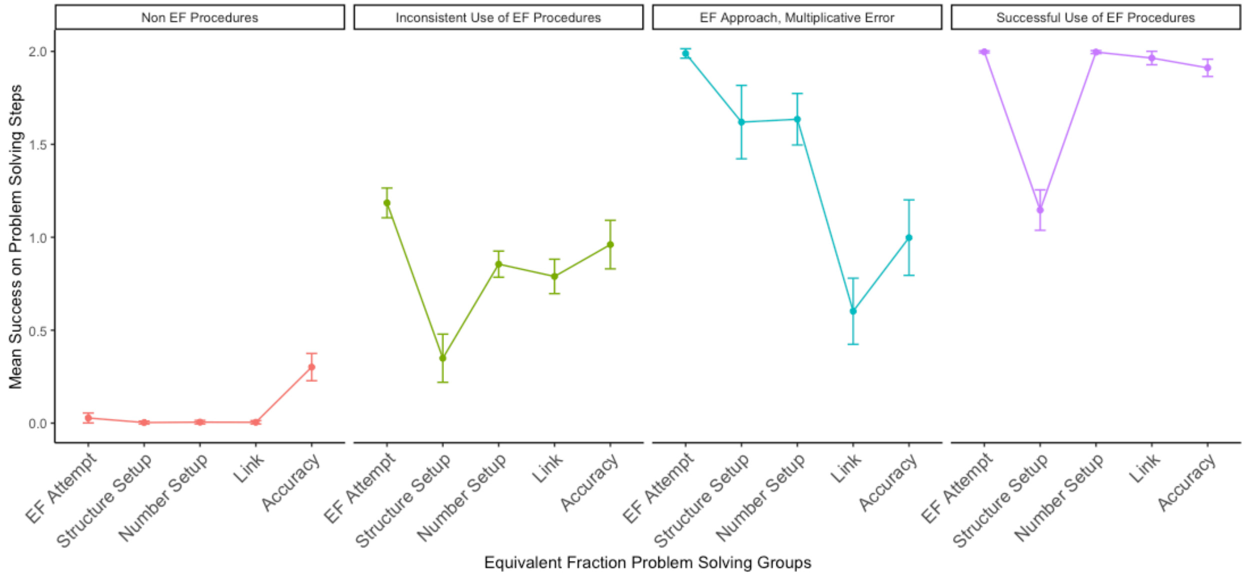
3.4.3 Latent Classes for Equivalent Fraction Problem-solving

The item-response and latent class probabilities estimated in Step 1 were used to calculate the posterior class membership probabilities, which indicate the probability of the student belonging to each group. Higher probabilities indicate that a students' response pattern is like the item response probability pattern for a specific group. Next, the posterior class membership probabilities were used to classify each student to the group for which they have the highest probability of belonging to (i.e., group assignment). Classification error probabilities were also calculated for all individuals.

Four latent classes (groups) emerged from the item-response probabilities. These groups represent students and their distinct problem-solving abilities (see Figure 2.6). One group was labelled 'Non Equivalent Fraction (EF) Procedures'. This group did not use the equivalent fraction strategy to solve the problems and had low overall accuracy; 30% of the total sample was assigned to this group. A second group was labeled 'Inconsistent Use of EF Procedures'. This group attempted the equivalent fraction strategy but inconsistently, had a sharp decline in the "setup" procedure meaning they did not use the same setup as the instructed example, and had moderate overall accuracy; 20% of the total sample was in this group. A third group was labelled 'EF Approach, Multiplicative Error'. This group was characterized by high rates for attempt of the equivalent fraction strategy, structural setup, and number setup with low rates of success in linking through multiplication/division resulting in low overall accuracy. This group had a sharp decline in the "link" procedures meaning they did not understand the key multiplicative feature for using the equivalent fraction strategy; 10% of the total sample was assigned to this group The most sophisticated group was characterized by successful problem-solving steps with a slight decline in setup however this did not affect the ability to complete the subsequent steps (i.e., they used alternative setup such as ratio or use of 'per'). This group was labeled 'Successful Use of EF Procedures'; 40% of the sample was assigned to this group.

Figure 2.6

Equivalent Fraction LCA Problem-Solving Groups



Note. Four problem-solving groups emerged from the LCA analysis for equivalent fraction problems: 'Non EF Procedures', 'Inconsistent Use of EF Procedures', 'EF Approach, Multiplicative Error', and 'Successful Use of EF Procedures'. Error bars indicate standard error. X-axis represent the sequence of problem-solving steps for the equivalent fraction problems: attempt equivalent fraction, structure setup, numbers setup, link, and accuracy. The mean on the y-axis represents mean success rate for each step of the problem-solving (0-2; representing the number of problems).

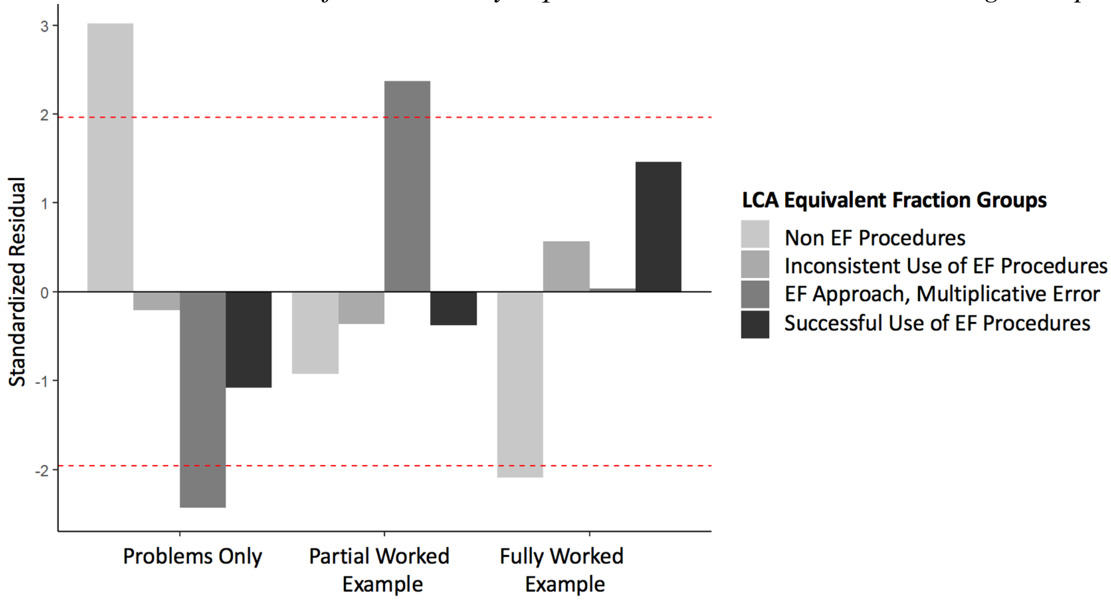
3.4.4 Step 3: Equivalent Fraction Problem-solving Group Membership by Experimental Condition

Next, we explored the relationship between experimental conditions and problem-solving group membership for the equivalent fraction problems. We explored whether students' progress through these groups in an ordinal fashion. To do so, we examined the ordinal-by-ordinal gamma statistic to determine whether equivalent fraction LCA groups and intervention conditions showed an ordered patterns of association. In this analysis, we considered an order of intervention condition of increasing support from PO control to PW Examples condition to FW Examples condition, and an order of Equivalent Fraction LCA groups of increasing sophistication from 'Non EF procedures', to 'Inconsistent Use of EF Procedures', to 'EF Approach, Multiplicative Error', to 'Successful Use of EF Procedures'. We also assessed standardized adjusted residuals to determine whether the cross-classification of subgroup-by-condition patterns occurred more/less often than expected by chance.

Cross-classification analysis showed equivalent fraction LCA groups and condition were significantly associated (χ^2 (6, n = 144) = 15.26, p = 0.018; γ = 0.23, p = 0.01); the significant, positive gamma statistic suggests an overall trend for children with the least amount of support (i.e., PO control) belonging to less sophisticated groups, and highest level of support (i.e., FW Example condition) likely to belong to more sophisticated equivalent fraction LCA groups. Examination of the standardized residuals (Figure 2.7) shows students in the PO condition, are most likely to belong to the 'Non EF Procedures' group. Students in the PW Example condition were likely to belong to the 'EF Approach, Multiplicative Error' group. Students assigned to the FW Example condition were unlikely to belong to the 'Non EF Procedures' and were trending, but not significantly more likely to belong to the 'Successful Use of EF Procedures' group. These data suggest that condition influenced the nature and sequence of students' problem-solving approach.

Figure 2.7

Standardized Residuals of Condition by Equivalent Fraction Problem-Solving Groups



Note. Standardized Residuals of students from each condition (e.g., Problems Only, Partially Worked Examples, Fully Worked Examples) by the LCA equivalent fraction problem-solving groups. If the residual is greater than 1.96, it indicates a greater than chance likelihood that the students assigned to that condition will belong to the equivalent fraction LCA group. If the residual is less than 1.96, then it indicates a lower than chance likelihood that a student assigned to the condition will belong to that particular equivalent fraction LCA groups. The pattern is not significant when -1.96 < residual < 1.96.

3.4.5 Covariates of Equivalent Fraction Problem-solving Group Membership

Finally, we investigated the relationship between proportional reasoning problem-solving groups and covariates. Grade, attentional control, prior knowledge of proportions, and arithmetic skills were entered as covariates. The Wald's tests in Table 2.6 indicates condition, knowledge of proportions and arithmetic skills predicted equivalent fraction problem-solving group membership. There was no effect of grade or attentional control on group membership. Coefficients indicate probability of belonging to four problem-solving groups. Higher positive values indicate a higher probability of belonging to a certain group, and lower negative values indicate a lower probability of belonging to a certain group. Coefficients suggest that students in the PO condition were likely to be in the 'Non EF Procedures' group. Students who were in the PW Example condition were most likely to be in the attempted 'EF Approach, Multiplicative Error'. This multiplicative error represents the first procedures students in this condition were to solve independently. Students in the FW Example condition were not likely to belong to the 'Non EF Procedures' group. Finally, students in the most sophisticated groups were likely to have higher arithmetic skills and knowledge of proportions. Students in the least sophisticated groups were likely to have low arithmetic skills.

Table 2.6
Experimental Condition and Cognitive Covariates of Equivalent Fraction Problem-Solving Group Membership

			Non EF Procedures		Inconsistent Use of EF Procedures		EF Approach, Multiplicative Error		Successful Use of EF Procedures	
	Wald	p-value	B	SE	B	SE	B	SE	B	SE
Condition	**16.76**	**.01**								
PO			**.87**	**.23**	.14	.26	**-.98**	**.47**	-.04	.25
PW Examples			-.40	.22	-.27	.24	**.89**	**.41**	-.22	.27
FW Examples			**-.48**	**.23**	.13	.26	.09	.45	.26	.26
5th Grade	5.09	.17	.29	.21	.12	.20	-.12	.29	-.29	.20
Attentional Control	5.68	.13	.00	.00	.00	.00	-.01	.00	.01	.00
Knowledge of Proportions	**12.81**	**.01**	-.61	.38	.34	.34	-.83	.69	**1.10**	**.34**
Arithmetic Skills	**11.67**	**.01**	**-.23**	**.08**	-.07	.09	**.22**	**.10**	.09	.07

Note. Bold indicates experimental significance. Conditions: Problems Only, Partially Worked Examples, Fully Worked Examples.

3.5.1 Step 1: Model Fit for Unit Ratio Problems

The same LCA was conducted for the unit ratio problems. Model fit information for the unit ratio problems can be found in Table 2.7. The fit criteria indicated the four-group model was the best fitting model based on the BIC, CAIC and ICL-BIC. The three-group model was identified as the optimal model by the AWE. Entropy values for the four-group model was acceptable, indicating good classification of individuals into latent groups. A conditional bootstrap procedure was conducted to compare the three-and four-group models (Vermunt & Magidson, 2016). The bootstrap likelihood difference analysis shows significantly more information was gained from the four-group model than the three-group model (Bootstrap 2LL Diff = 49.79, $p < .0001$) and an non-significant difference between four- and five-group model (Bootstrap 2LL Diff = 16.22, $p = .1$). Therefore, the four-group model was selected for unit ratio problem-solving because of its statistical fit, parsimony, and theoretical interpretability in characterizing patterns of proportional reasoning problem-solving.

Table 2.7
Model Fit Indicators for LCA of Unit Ratio Problem-Solving

	LL	BIC	CAIC	AWE	ICL-BIC	Entropy R2
1-Group	-1139.29	2351.89	2365.89	2467.20	2351.89	1.00
2-Group	-661.57	1438.35	1460.35	1622.95	1441.75	0.99
3-Group	-560.20	1277.49	1307.49	1528.90	1281.80	0.99
4-Group	**-535.30**	**1269.59**	**1307.59**	**1585.35**	**1272.37**	**0.99**
5-Group	-525.65	1292.18	1338.18	1674.19	1295.32	0.99
6-Group	-515.21	1313.19	1367.19	1777.08	1332.31	0.96
7-Group	-500.60	1325.86	1387.86	1857.25	1346.60	0.97
8-Group	-491.30	1349.15	1419.15	1945.44	1368.89	0.97

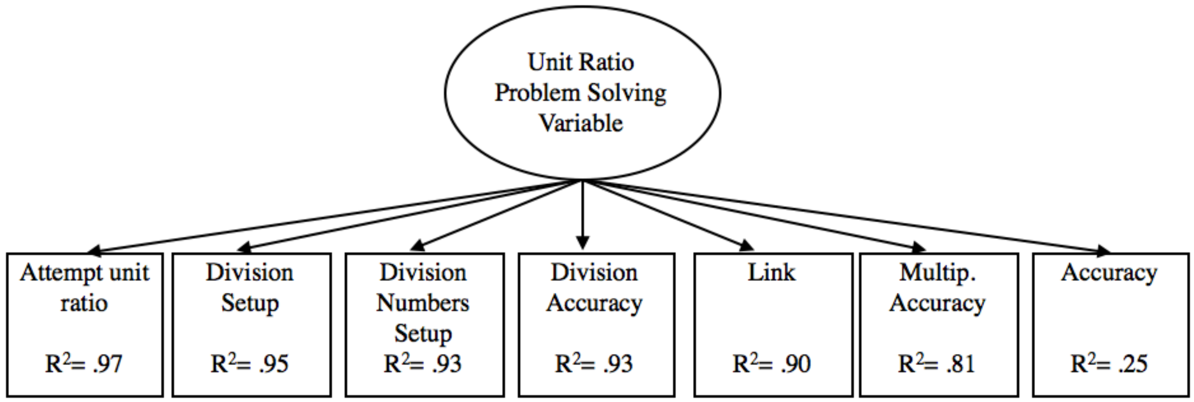
Note. Bold text indicates the selected model. LL, Log Likelihood; BIC, Bayesian Information Criterion; CAIC, Consistent Akaike Information Criteria; AWE, Average Weight of Evidence; ICL-BIC, a version of Integrated Classification Likelihood.

3.5.2 Step 2: Characterization of Unit Ratio Problem-solving Groups

The cluster loadings of each problem-solving indicator can be found in Figure 2.8. For unit ratio problem-solving, accuracy is not the most influential indicator to class membership problem-solving approach (attempt unit ratio), division setup, division numbers setup, division accuracy, and link can strongly predict class membership all with R^2 over .9.

Figure 2.8

Cluster Loadings of the Problem-Solving Indicators for Unit Ratio Problems Estimated in the Latent Class Analysis

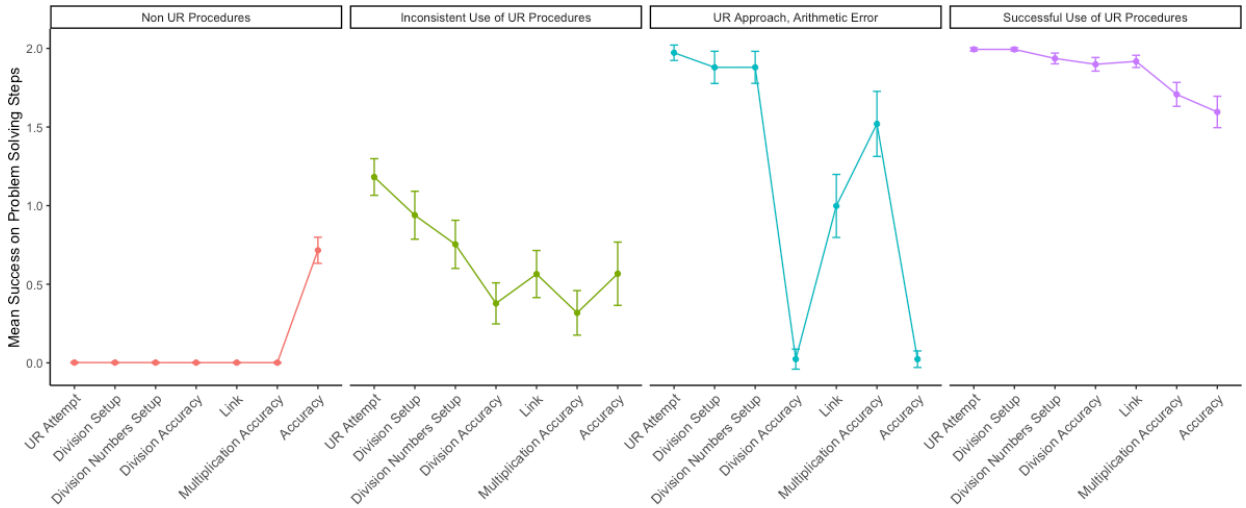


3.5.3 Latent Classes for Unit Ratio Problem-solving

For the unit ratio LCA, four latent classes (groups) emerged from the item-response probabilities and are summarized in Figure 2.9. One group was labelled 'Non UR Procedures'. This group did not use the unit ratio strategy to solve the problems and was characterized by low overall accuracy; 58% of the sample was assigned to this group. A second group was labeled 'Inconsistent Use of UR Procedures'; this group attempted the unit ratio strategy but was characterized by inconsistency across the full sequence of problem-solving steps; 9% of the sample was in this group. A third group was labelled 'UR Approach, Arithmetic Issue'; this group was characterized by high rates for unit ratio attempt and setup with low arithmetic competency for long division and low rates in the linking step (i.e., using the quotient from the first step and applying it to the larger recipe). The errors from these two key steps led to low overall problem accuracy; 6% of the sample was assigned to this group. The most sophisticated group was characterized by unit ratio attempt, successful problem-solving steps and high rates of solution accuracy. This group was labeled 'Successful Use of UR Procedures'; 27% of the sample assigned to this group.

Figure 2.9

Unit Ratio LCA Problem-Solving Groups



Note. Four problem-solving groups emerged from the LCA analysis for unit ratio problems solving steps: 'Non UR Procedures', 'Inconsistent Use of UR Procedures', 'UR Approach, Arithmetic Error' 'Successful Use of UR Procedures'. Error bars indicate standard error. X-axis represent the sequence of problem-solving steps for the unit ratio problems: attempt unit ratio, division setup, division numbers setup, accuracy, link, multiplication accuracy, and overall accuracy. The mean on the y-axis represents mean accuracy for each step of the problem-solving (0-2; representing the number of problems).

3.5.4 Step 3: Unit Ratio Problem-solving Group Membership by Experimental Condition

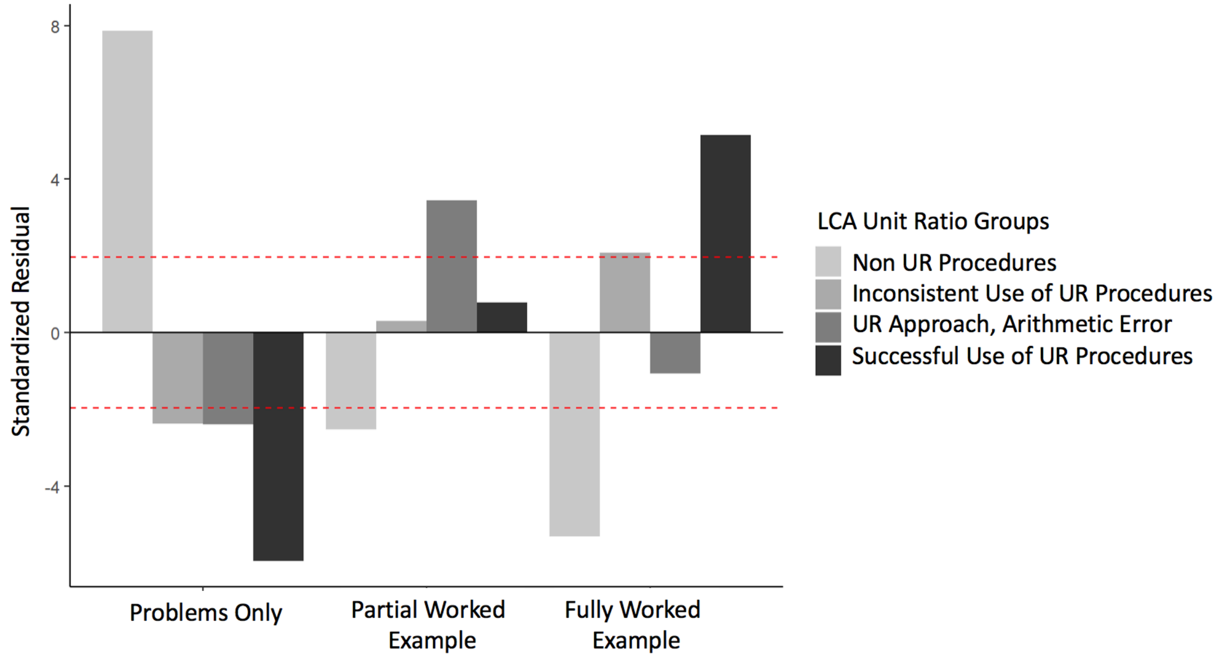
Next, we explored the relationship between experimental conditions and problem-solving group membership for the unit ratio problems. We examined the ordinal-by-ordinal gamma statistic to determine whether unit ratio LCA groups and intervention conditions showed ordered patterns of association. In this analysis, we considered an order of intervention condition of increasing support from PO to PW Examples to FW Examples, and an order of unit ratio LCA groups of increasing sophistication from 'Non UR Procedures', to 'Inconsistent Use of UR Procedures', to 'UR Approach with Arithmetic Error,' to 'Successful Use of UR Procedures.' we also assessed standardized adjusted residuals to determine whether the cross-classification of subgroup-by-condition patterns occurred more/less often than expected by chance.

Cross-classification analysis showed unit ratio LCA groups and condition were significantly associated (χ^2 (6, n = 144) = 75.41, p <0.001; γ = 0.73, p <0.001); the significant, positive gamma statistic suggests an overall trend for children with the least amount of support (i.e., PO control) belonging to less sophisticated groups, and highest level of support (i.e., FW Example condition) belonging to more sophisticated unit ratio LCA groups.

Examination of the standardized residuals (Figure 2.10) shows students in the PO condition, are most likely to belong to the 'Non UR Procedures' group, and unlikely to belong to the other three groups. Students in the PW Example condition were likely to belong to the 'UR Approach, Arithmetic Error' group and unlikely to belong to the 'Non UR Procedures' group. Students assigned to the FW Example condition were likely to belong to the 'Successful Use of UR Procedures' group and unlikely to belong to the 'Non UR Procedures.'

Figure 2.10

Standardized Residuals of Students in Condition by Unit Ratio Problem-Solving Groups



Note. Standardized Residuals of students from each condition (e.g., Problems Only, Partially Worked Examples, Fully Worked Examples) by the LCA unit ratio problem-solving groups. If the residual is greater than 1.96, it indicates a greater than chance likelihood that the students assigned to that condition will belong to the unit ratio LCA group. If the residual is less than 1.96, then it indicates a lower than chance likelihood that a student assigned to the condition will belong to that particular unit ratio LCA groups. The pattern is not significant when -1.96 < residual < 1.96.

3.5.5 Covariates of Unit Ratio Problem-solving Group Membership

Finally, we investigated the relationship between proportional reasoning groups and covariates. Grade, attentional control, prior knowledge of proportions, and arithmetic skills were entered as covariates. The Wald's tests in Table 2.8 indicates condition, attentional control, and knowledge of proportions predicted unit ratio problem-solving group membership. There was no effect of grade or arithmetic skills on group membership. Coefficients indicate probability of belonging to problem-solving groups where higher positive values indicate a higher probability of belonging to a certain group, and lower negative values indicate a lower probability of having a certain group. Coefficients suggest that students in the PO condition were likely to be in the 'Non

UR Procedures' group and unlikely to be in sophisticated problem-solving groups. Students who were in the PW Example condition were most likely to be in the 'UR Approach, Arithmetic Error' group and the 'Successful Use of UR Procedures' group, and not likely to be in the 'Non UR Procedures' group. Students in the FW Example condition were most likely to be in the 'Successful Use of UR Procedures' group and least likely to be 'Non UR Procedures' group. Finally, students in the 'Successful Use of UR Procedures' group were likely to have higher prior knowledge of proportions and higher attentional control. Students in the 'Non UR Procedures' group were likely to have lower arithmetic skills.

Table 2.8

Experimental Condition and Cognitive Covariates of Unit Ratio Problem-Solving Group Membership

			Non UR Procedures		Inconsistent Use of UR Procedures		UR Approach, Arithmetic Error		Successful Use of UR Procedures	
	Wald	p-value	B	SE	B	SE	B	SE	B	SE
Condition	**273.58**	**.001**								
PO			**2.87**	**.27**	.83	.55	**-1.23**	**.33**	**-2.47**	**.28**
PW Examples			**-1.43**	**.22**	-.81	.43	**1.45**	**.38**	**.79**	**.22**
FW Examples			**-1.45**	**.23**	-.01	.35	-.22	.42	**1.68**	**.23**
5th Grade	5.35	.15	.43	.20	.05	.34	-.54	.37	.06	.19
Attentional Control	**9.85**	**.02**	.00	.00	-.01	.00	.00	.01	**.01**	**.00**
Knowledge of Proportions	**8.87**	**.03**	.17	.29	.22	.47	**-.97**	**.47**	**.58**	**.25**
Arithmetic Skills	6.80	.08	**-.21**	**.09**	-.09	.12	.28	.16	.03	.08

Note. Bold indicates experimental significance. Conditions: Fully Worked Examples, Partially Examples, and Problems Only.

4. Discussion

This current study addressed two important theoretical and practical decisions when designing worked examples for classroom instruction: i) can worked examples be used to *introduce* students to novel solution strategies and change the way they think about and solve proportional reasoning problems and ii) are fully or partially worked examples the ideal tool to do so? The methodology used in this study, uniquely adds to the literature since we measured differences in intervention effects between the fully worked and partially worked examples, but also, importantly, captured students' proportional reasoning problem-solving skills and their ability to engage with sophisticated solution strategies without any forms of support beyond worked examples. Overall, findings from the current study suggest that fully worked examples enhance solution accuracy, as expected. More uniquely both fully and partially worked examples fundamentally changed the way students' deployed problem-solving approaches and encouraged students to employ novel solution strategies. These worked examples also support students' ability to map procedures thereby facilitating problem-solving success for novel, sophisticated solution strategies.

At baseline, without any support, students' accuracy levels for solving proportional reasoning problems was in line with our expectations (Common Core State Standards in Mathematics, 2010). More than half of the students in sample had some understanding of proportionality and solved the word problems correctly. However, students in this control condition used unsophisticated problem-solving strategies such as additive strategies (e.g., build up strategy). While many students had an elementary understanding of proportionality, many other students did not even seem to recognize the need for proportional thinking to solve the problems. More than 1/3 of the sample from the control condition used incorrect, non-proportional solution strategies (e.g., addition or subtraction) and did not answer any problems correctly. These findings in accuracy and problem-solving approaches from the control condition highlight that if students have an understanding of proportionality at this age it is very elementary. Moreover, it validates that this is the ideal timeframe to test the research question of using worked examples as a tool to introduce sophisticated solution strategies for solving proportional reasoning word problems.

There is a growing literature examining worked examples in classroom settings (see Booth et al., 2010; Rittle-Johnson & Star, 2009) which have shown strong effects of posttest accuracy when students study fully worked examples and when worked examples are presented in a partial format

(Baars, Visser, Van Gog, De Bruin, & Paas, 2013). Much of the previous work has compared learning outcomes between problem-solving only and worked example—problem conditions (Moreno, 2006). This work uniquely added to the literature by examining the effects of fully vs partially worked examples on problem-solving accuracy and approach. When students are provided with fully worked examples, accuracy rates almost double compared to the problems only control and partially worked example condition. These findings are in line with previous research on worked examples which suggests that studying fully worked exemplars are most impactful during the initial stages of skill acquisition (Renkl, 2014). However, these accuracy findings did not allow us to examine if there were differential effects between the fully and partially worked examples on students' problem-solving approach.

Moving beyond accuracy, we examined *how* students were solving the proportional reasoning problems. The same microanalytic approach was used to examine if the worked examples fundamentally changed the way students approached problem solving. A further question is what the mechanism might be by which worked examples supported changes in student thinking.

Theorists have argued that worked examples reduce extraneous cognitive load (Pass et al., 2003; Sweller, 2010) freeing up resources for students to attend to the procedures and structural components of the problem-solving approach (Sweller, Van Merrienboer, & Pass, 1998; Gog et al., 2004). Evidence form this study suggest students were not only able to implement and map procedures in near-transfer problem-solving contexts but that the worked examples fundamentally changed the way they thought about and approached proportional reasoning problem-solving. This was the case for both the fully worked and partially worked example conditions; students who received worked examples attempted the instructed, novel, and sophisticated solution strategies.

There were however key differences in problem-solving procedures between the fully worked and partially worked example conditions. For students in the partially worked example condition, they were likely to be in the "Error" group in that they made key errors in proportional thinking (i.e., they did not use multiplicative reasoning). This finding suggests evidence against the desirable difficulty and theory (Bjork & Bjork, 2011) and exploratory learning approach (DeCaro & Rittle-Johnson, 2012). During initial skill acquisition of novel, sophisticated solution strategies it seems important to provide students with the key conceptual features of the proportional reasoning strategies as students are not able to intuit the multiplicative nature of proportional reasoning without the appropriate level of support (Renkl, Atkinson, & Grobe, 2004). These results suggest that if using worked examples as a pedagogical tool to *introduce* solution strategies, fully worked examples are the most effective.

The multiple strategy solution methodology of this study also allowed for us to compare if worked examples encouraged students to employ solution strategies that are more familiar (equivalent fraction) or more novel (unit ratio strategy). For students who received fully worked examples, there were key differences between the two solution strategies. For the more familiar, equivalent fraction strategy, there was a tending effect for students to be in the most sophisticated problem-solving group. For the more novel approach, unit ratio strategy, there was a clear benefit of the fully worked examples as these students were highly likely to be in the most sophisticated problem-solving group.

Findings from this study showcase that both the coding scheme and analytic tool (LCA) allow for us to characterize the multiple steps of students' problem-solving procedures. Moreover, these techniques are in fact sensitive measures to capture key differences in the intervention on the nature and sequence of students' problem-solving. That is, the coding protocol and LCA captured the exact moment at which the partial worked examples did not provide the multiplicative instruction. This provides evidence that future research should not only explore problem-solving accuracy, but should examine students' problem-solving approach and procedures.

Overall, evidence from this work adds to the literature by examining how worked examples can be used as a pedagogical tool to *introduce* solution strategies prior to direct teacher instruction. Findings from this study add to the current literature on worked examples and further suggest that worked examples are a worthwhile instructional practice as promoted by the Institute of Education Sciences (Pashler, Bain, Bottge, Graesser, Koedinger, McDaniel, & Metcalfe, 2007). The study shows that in a single classroom session with only a mathematics booklet and one worked example for each solution strategy, that both partial and fully worked examples can fundamentally change the way students think about and solve proportional reasoning word problems. Providing exemplars not only encourage students to use novel solution strategies but provides the scaffolding for students to implement and map the problem-solving procedures with success. Moreover, there are benefits for all students, regardless of their prior knowledge and attentional control.

Findings from this study suggest, however, that if worked examples are to be used as a tool to *introduce* problem-solving strategies, that is best to provide students with fully worked examples rather than partially worked examples where key procedures are left out. While partial worked examples may have a benefit for supporting the transition to independent problem-solving (Renkl, 2017), these findings show fully worked examples are the ideal pedagogical tool to introduce novel strategies and can even work for highly sophisticated strategies. The limitation of partially worked examples may be due to the fact that students do not have enough understanding of multiplicity and proportionality and therefore cannot intuit the missing procedure. It may still be the case that partial worked examples scaffold learning and better support learners who already have a foundation of the problem-solving principles (Renkl, 2014).

What this study did not examine was how worked examples can be used as a tool during introduction of solution strategies in a broader time frame or in a longer mathematics packet. It may be that partial worked examples can be faded within one mathematics packet such that students are presented with fully worked examples and then worked examples with procedures eliminated over the course of the introductory packet. This procedure may provide opportunity for students to begin to grasp the solution strategy while also providing an opportunity for exploratory learning. In this study we examined the differences between fully and partially worked examples; however, future work should compare fully worked examples to fading in an introductory context with a longer mathematics packet.

In addition to using worked examples as a tool to introduce problem-solving strategies, future studies should examine how fully worked examples can be paired with teacher instruction as a *double-dose* to encourage use of novel strategies and support mapping of problem-solving steps.

Study 2, presented next, I tested fully worked examples paired with direct teacher instruction to examine if worked examples scaffold learning for new solution strategies *during* instruction.

CHAPTER THREE: WORKED EXAMPLES SUPPORT LEARNING OF NOVEL SOLUTION STRATEGIES DURING A PROPORTIONAL REASONING LESSON

1. Introduction

Scaffolding (Wood et al., 1976) first emerged to describe processes teachers use to support student learning. This support may be particularly beneficial in conceptually challenging mathematics lessons where new material is introduced and in lower achieving contexts (Broza & Kolikant, 2015; Anghileri, 2006). Example-based support is one such tool in which teachers can scaffold learning. Experts agree, example-based support like worked examples are powerful teaching tools, capable of supporting students during problem-solving; yet, most of this worked has examined the benefits of worked examples during problem-solving *after* instruction (Sweller & Cooper, 1985; Rittle-Johnson & Star, 2009; Atkinson et al., 2000; Kalyuga et al., 2001, 2003; Renkl, 2010, 2014, 2017). Surprisingly little is known about the potential benefits of integrating worked examples paired with direct teacher instruction during complex in-classroom mathematics lessons. It is possible that providing a *double-dose* of visual support in the way of instruction paired with worked examples might provide the scaffolding students need for learning of complex, new material. Moreover, research which examines the mechanisms for how worked examples support learning remains unknown; research which examines how students employ novel solution strategies during the lesson and how they map step-by-step procedures from worked examples to transfer problems would provide this insight.

Building on the findings from Study 1 (Chapter 2), this study examines the potential benefits of pairing fully worked examples in the classroom with direct teacher instruction. Specifically, we tested whether a *double-dose* of instruction, which in this context represents a high-quality teacher-led lesson paired with worked examples *during* problem-solving, supports students learning of novel solution strategies more than a high-quality lesson alone. The lesson used in this study increased the ecological validity from Study 1; students watched and solved problems within a previously-recorded, conceptually challenging mathematics teacher led lesson on proportional reasoning. Within classrooms, students were assigned to the treatment condition of high visual support (worked examples) during instruction and problem-solving or low visual support (no worked examples) during instruction and problem-solving. Worked examples were screenshots of the teacher's solution strategy and worked out procedures. Examples also included a brief conceptual overview of the solution strategy. In addition to examining students' problem-solving accuracy as a function of visual support, as in Study 1, a more detailed assessment of problem-solving was conducted to determine whether worked examples encouraged use of novel solution strategies and assessed their influence on students' ability to map procedures from the worked example to near transfer problems.

1.1 Impact of Worked Examples

Learning mathematical procedures for novel solution strategies is difficult, especially in complex and cognitively demanding mathematics contexts. While it is common in mathematics instruction and math textbooks to have students independently attempt newly trained strategies and procedures on new problems, this approach may not be very effective in supporting accurate procedural steps of the solution strategy, particularly without detailed, ongoing immediate feedback. A large body of work shows that sometimes more efficient and effective learning can be, often in less time, is achieved when problems to solve are replaced with worked examples for students to study (see Renkl, 2017, Sweller & Cooper, 1985; Cooper & Sweller, 1987; Carroll, 1994; Zhu & Simon, 1987; Atkinson et al., 2000). This approach is especially suited to foster the initial acquisition of cognitive skills (Renkl, 2014, 2017).

There is an impressive body of research on worked examples both in well-controlled laboratory settings (e.g., Cooper & Sweller, 1987; Sweller & Cooper, 1985), in classroom settings (e.g., Booth et al., 2015; McLaren et al., 2015; Reiss et al., 2008) and across ages (Christie & Gentner, 2010; Rittle-Johnson & Star, 2007), which collectively show that example-based learning can improve student achievement in many domains. Research has also explored the effects of worked examples paired with other methods to further enhance learning. For example, research which examines worked examples in combination with self-explanation prompts (e.g., Chi, 2000; Aleven & Koedinger, 2007; Hilbert et al., 2008), incorrect vs. correct examples (Booth et al., 2013; Ohlsson, 1996; Siegler, 2002), comparing different forms of worked examples (Rittle-Johnson & Star, 2007, 2009), alternating worked examples and practice problems (Trafton & Reiser, 1993), and even as an additional ingredient in Cognitive Tutors (e.g., Salden et al., 2009; 2010; Schwonke et al., 2009) can be effective supports while students practice their newly acquired strategy solutions.

There are a few important shortcomings of this work. Instructional design theory suggests effective examples scaffold learning most when they are highly integrated, employ multiple modalities in presentation (e.g., at the lesson level and practice problems), and emphasize labeling (Atkinson et al., 2000). Despite this extensive body of research, nearly all studies of worked examples have been conducted after teacher instruction, during homework, while the student practices the newly learned problem-solving procedures (Retnowati et al., 2010). It is unclear if and how worked examples can be used in combination with direct teacher instruction to promote use of novel solutions strategy and success with problem-solving procedures. Additionally, many of these studies have almost exclusively examined student accuracy to measure gains (Sweller & Cooper, 1985; Rittle-Johnson & Star, 2009), or conceptual understanding measures (Chi, 2001;

Booth et al., 2013), but none to date have examined the details of how problem-solving proceeds when learners aim to solve a problem using a worked example.

1.2 Proportional Reasoning as a Cognitively Demanding Mathematics Context to Examine the Effects of Worked Examples

Insight into the mechanisms underpinning how worked examples can support learning of novel solution strategies could be gained from examining students' step-by-step procedures and mapping from the worked example to transfer problems. Drawing on cognitive science research, worked examples can act as a source representation in which students map the novel strategy and procedures from the worked examples to the transfer problems (see Morrison et al., 2004; Gentner, 1983). Worked examples as a scaffold may be particularly effective in supporting use of novel solution strategies and procedures when teaching in the complex domain of proportional reasoning word problems as they i) are relationally complex, ii) require multiple steps/procedures, and iii) require multilevel mappings (across and within representations). Understanding the mechanisms of mapping may give us insight into how worked examples are supporting procedural gains in learning. Take for example a proportional reasoning word problem:

> Alex is making a strawberry cake. To make a small cake, the recipe calls for 2 eggs and 6 strawberries. Alex wants to make a bigger cake so he uses 8 eggs. How many strawberries does Alex need for the big cake?

In order to solve this problem, one must attend to the surface level features (e.g., the specific numbers and problem context) while also attending to the structural level principles (e.g., the mathematic approach, comparison within and between cakes, etc.). To solve the problem, students must attend to the multiple steps in alternative procedure approaches, determine what they expect to be the most effective solution strategy, set up the numerical symbolic equation structure, and conduct the procedures and arithmetic in order to arrive at a final solution. In addition to procedures, the learner must also consider the multiple levels of mapping. Proportional reasoning involves understanding the multiplicative relationships between quantities ($A/B = C/D$; Cramer & Post, 1993; Lesh et al., 1988). One must attend to the comparison between the ingredients within the small cake 2 eggs: 4 strawberries and the structural relationship between the ingredients across the cake recipes 2 eggs::8 eggs. Compare this to a long division problem (e.g. $6 \div 2$ formatted with mathematical symbols) where the problem is already set-up and clearly indicates the learner needs to complete a division procedure and the learner only needs to conduct the procedures to arrive at the final solution. Solving proportional reasoning word problems like the example above is evidently more complex and cognitively effortful, requiring not only setting up the symbolic structure before executing arithmetic procedures, but also mapping across two ratios inherent in the problem. Additionally, teachers find this curriculum area to be quite challenging to teach (Sowder, 2007). Therefore, worked examples during instruction may be particularly supportive as students may be provided with necessary scaffold to align and more easily map the procedures novel solution strategy to a near transfer problem.

In additional to the relational complexity, proportional reasoning offers a unique context to study the mechanisms of worked examples as students struggle with the key feature (i.e., multiplicative thinking) of proportionality. Secondarily, proportional reasoning is a unique yet complex domain as students can employ multiple strategy solutions that vary in their level of sophistication each with distinct steps, demands, and skills.

Students experience difficulty in proportional reasoning as they are shifting from an additive way to multiplicative way of thinking. Here we draw the comparison of proportional reasoning strategies to relational reasoning. While the process of solving analogy test items of the form A:B::C:D has been studied extensively (Sternber, 1977), there has been little experimental investigation of relational thinking in more complex mathematics domains. In proportional thinking domains, students' progress from guessing or a visual judgment to an, incorrect additive stage where students use subtraction or addition (Parish, 2010; Nabors, 2003). This can be viewed like a lower order version of relational reasoning in which students are comparing a single level of relational complexity (i.e., A:B or C:D). Increasing sophistication but not levels of relational complexity, students can show a basic understanding of proportionality and use a repeated pattern strategy (e.g., buildup) in which students repeatedly add or subtract a number to each part of the ratio. The next stage of understanding is when students begin to use multiplication in place of repeated addition (Nabors, 2003). In this stage, students can setup fractions and multiply or divide the numbers by the same value (i.e., equivalent fraction strategy). Even more sophisticated, students can determine the unit rate of the proportions thereby being able to apply that same rate to any desired quantity (i.e., unit ratio strategy), made up of division and multiplication (Nabors, 2003). These most sophisticated strategies reflect higher levels of relational thinking and complexity at multiple levels, within and across representations (i.e., A:B::C:D).

1.3 Individual Difference Factors

Developing and orchestrating classroom lessons in which learners fundamentally change the way they think about problem-solving in a conceptual domain, and learn procedures for new solution strategies is a difficult task especially during a cognitively demanding mathematical context with high relational complexity like in proportional reasoning. First, learning suffers when learners cognitive load is taxed (Cho, Holyoak, & Cannon, 2007; English & Halford, 1995; Morrison et al., 2001; Richland et al., 2006; Waltz et al, 2000; Paas et al., 2003). Second, learners' prior knowledge plays a key role in determining if he/she will be able to notice the key structural features (Gentner & Tattermann, 1991). Third, learners often fail to notice the relevance or importance of mapping procedures from a given example unless given very clear and explicit support cues to do so (see Gick & Holyoak, 1980; Alfieri et al., 2013; 1983; Gentner et al., 2003; Ross, 1989; Schwartz & Bransford, 1998). These challenges highlight that the learning context is essential in determining whether and how students notice and successfully execute mapping across multiple representations. This literature also suggests that worked examples may be a successful pedagogical tool to enhance mapping.

Proportional reasoning has a high level of relational complexity, that is systems of relations mapping to systems of relations. By definition a proportion involves higher order equivalent relation between two ratios, each with one level of complexity. Relational complexity can overload children's executive functioning resources. Halford (1993) suggested that the inability to process and map across multiple relations is in part due to children's executive functioning capacity which increases with age (Halford et al., 1998). Executive functioning is defined as the limited cognitive processing system that deploys resources to modulate the operation of various cognitive processes (see Miyake et al, 2000). Within executive functioning, two of the primary subsystems are particularly related to reasoning (see Krawczyk et al, 2010; Morrison et al., 2001): Working Memory (WM), the resources for holding information active within attention and manipulating that information (Engle & Kane, 2004), and attentional control, the processes of controlling attention away from irrelevant information (Diamond, 2002). WM plays an important role in coordinating mental representations. In this case, to mentally consider the relationship between the source representation (i.e., worked example) and the target representation (i.e., the near transfer problem) one must hold in mind and identify the similarities and differences across the representations, reorganize and represent these systems to align and map them together, and infer conceptual and schematic features (see Morrison et al., 2004; Morrison et al., 2011; Waltz et al, 2003). Attentional control is also integral to suppressing irrelevant yet potentially salient mappings (e.g., when comparing two math problems with the same story context, the surface level similarity is irrelevant; Richland, Morrison & Holyoak, 2006). Much research has examined that age related limitations or individual differences in working memory may explain why some students map relationships based on surface level features compared to structural level features or map on one level versus multiple levels (Richland et al., 2006).

Second, students' prior knowledge also plays an important role in their ability to structure map (Holyoak & Thagard, 1989; Holyoak, 2012). Without adequate domain knowledge of the key relationship or comparison between the source and the target, the student is unlikely to notice the structure mapping opportunity (Gentner & Rattermann, 1991; Goswami, 2001; Fyfe et al., 2012). Additionally, for mathematical reasoning tasks such as proportional thinking, learners must have adequate procedural fluency, that is, a high level of arithmetic skills so that their cognitive resources can be allocated to learning the more complex features of proportionality. Worked examples may be a resource for thereby reducing the procedural load (Paas, 1999) within the mathematical problem solving and allowing students to allocate resources to managing the relational complexity.

1.4 Current Study

The current study tests a classroom-relevant, highly implementable instructional practice shift that could have large benefits during learning: making worked examples visible during learning of novel solution strategies. This study tested systematic differences in learning based on variations in visual representations of novel solution strategies that were presented during instruction of proportional reasoning word problems. Students were randomly assigned to receive worked examples i) during videotaped instruction only or ii) a *double-dose* approach of worked examples made visible during videotaped instruction paired with a worked example during problem-solving. Using a novel microanalytic approach, we report a series of latent class analyses which allow us to examine the details how students employ novel solution strategies and to map procedures from exemplars between the two intervention conditions. This approach enables us to examine specific features of the relational reasoning and proportional procedures necessary for fully accurate solutions, which can provide a precise and detailed insight into how students are learning from these two worked example conditions.

Worked examples as a scaffold to enhance learning for novel solution strategies is theorized to support mapping from the source (worked example) to the target (near-transfer problem) representation at multiple levels. Students must: i) retrieve the right source representations (i.e., strategy approach), ii) align and setup the problem context, and iii) perform the mathematics, which means inferring how to execute instructed procedures, link between steps, and produce a final solution. We placed worked examples (a screenshot of the teacher's solution strategy) in the students' mathematics packets to test if students could align, map, and infer procedures more successfully than students who saw this information during the video lesson only. Beyond simply measuring improvements in accuracy, which were hypothesized, crucially we also explored what aspects of students' problem solving changed with the additional worked example resources. Therefore, findings from this experiment will yield both theoretical insight into the role of visual representations in complex multilevel mapping, and also provides practice relevant implications for everyday mathematics teachers.

The role of prior knowledge is also theorized to play a role in increasing the likelihood that learners can map correspondences across relational representations, so we used the same method as in Study 1, which enabled us to differentially examine the impacts of the single or double dose of worked examples on mathematical solution that relies on more familiar mathematics, the equivalent fraction strategy, or required a more novel set of procedures, the unit ratio strategy. These solution strategies varied in conceptual, procedural understanding, familiarity and sophistication. In this context, we hypothesize that worked examples will enhance students' accuracy and mapping between a previously instructed problem and a near-transfer practice problem by reducing cognitive load barriers. Additionally, we hypothesize that students' prior knowledge would predict accuracy and the nature and sequence of students' problem-solving procedures.

2. Methods

2.1 Participants

Fifth- and sixth-grade students from seven K-8 schools (17 classrooms) in the Chicagoland area with primarily underrepresented Latinx or Black youth in low to middle income neighborhood schools were recruited to participate in this study. They were chosen because students at this age have been introduced to the math concepts necessary to solve proportional reasoning problems; however, they have not yet been formally introduced to solution strategies for solving proportional reasoning word problem (Common Core State Standards in Mathematics, 2010). Students who were absent on the day in which the intervention was administered were excluded from analyses, yielding a sample of 224 students (166 5th grade, 58 6th grade; 100 females; 31 unidentified). Parents and guardians were informed of the study a few weeks prior to data collection and were provided the opportunity to opt their child out. We also obtained children's written assent prior to data collection.

2.2 Teacher and Experimenter Demographics

All of the children's regular mathematics teachers were present in the classroom throughout the study and at least two experimenters were in the classroom during the study.

2.3 Design and Procedures

This is a multisite randomized study; procedures were administered during two in-classroom visits over a one-week period. Students completed all procedures in their mathematics classroom, alongside their peers, and during their typical mathematics class period. The experimental manipulation was i) *double-dose:* worked examples while problem-solving or ii) no worked examples (control) while problem-solving during the mathematics lesson.

Session 1 (Day 1). Students completed a pretest assessing their initial understanding of proportional reasoning and arithmetic skills, material to be covered in the lesson. Students also completed a measure of selective and sustained attentional control, the d2 Test of Attention (Brickenkamp & Zillmer, 1998; Rhonda & Ross, 2005).

Session 2 (Day 2). Two to three days later, students watched a previously-recorded, conceptually challenging mathematics lesson on proportional reasoning which was administered individually on iPads. During the video lesson, students simultaneously completed a problem-solving packet. In this packet, the treatment manipulation was administered; students either received a mathematics packet with i) worked examples and problems to solve or a ii) packet with no worked examples and problems only.

2.4 Math Lesson

Researchers worked with a teacher and curriculum designer to create a lesson script introducing ratio and proportion. The 45-minute teacher-led lesson was recorded as a live, semi-scripted lesson on proportional reasoning strategies. The teacher taught a diverse class of fifth- and sixth-grade students who were recruited for the recording of the lesson. Recording a live lesson with real students allows for the natural variability of classroom instruction with students having real world conversations regarding the solution strategies. This experimental procedure enables the instructional stimuli to have high ecological validity while maintaining experimental control.

The lesson is designed based on a reform-based instructional model in which a teacher leads mathematical discussions where students compare and contrast solution strategies to a single problem (e.g., Carpenter et al, 1999; Kazemi & Hinz, 2014; Smith et al, 2009, Stigler & Hiebert, 1999). First, the teacher asks the students to solve a challenging proportional reasoning problem on their own prior to receiving explicit instruction (DeCaro & Rittle-Johnson, 2012; Schwartz et al., 2011; Schwartz & Martin, 2004). The teacher walked around the room and chose a student who used the equivalent fraction strategy; she asked the student to the board to describe the solution strategy to the class (e.g., Carpenter et al., 1999; Kazemi & Hinz, 2014; Smith et al., 2009; Stigler & Hiebert, 1999). Following, the teacher then led a discussion on the procedures and provided a higher-level conceptual overview of the equivalent fraction strategy. Students solved a near-transfer problem using the equivalent fraction strategy. The lesson employed a multiple solution strategy approach. Next, the teacher brought a second student up to the board to describe solving the same problem but this time using the unit ratio strategy. Following, the teacher led a discussion on the procedures and provided a higher-level conceptual overview of the unit-ratio strategy. Students solved a near-transfer problem using the unit ratio strategy. Then, the teacher summarized both solution strategies for both the instructed and near-transfer problems.

The math lesson was videotaped using multiple cameras capturing different angles of the teacher, the white boards, and the students. These angles allowed for manipulation of the raw video footage. To develop our study manipulations, videotapes were edited to produce the instructional stimuli.

2.5 Malleable factor: visual instruction

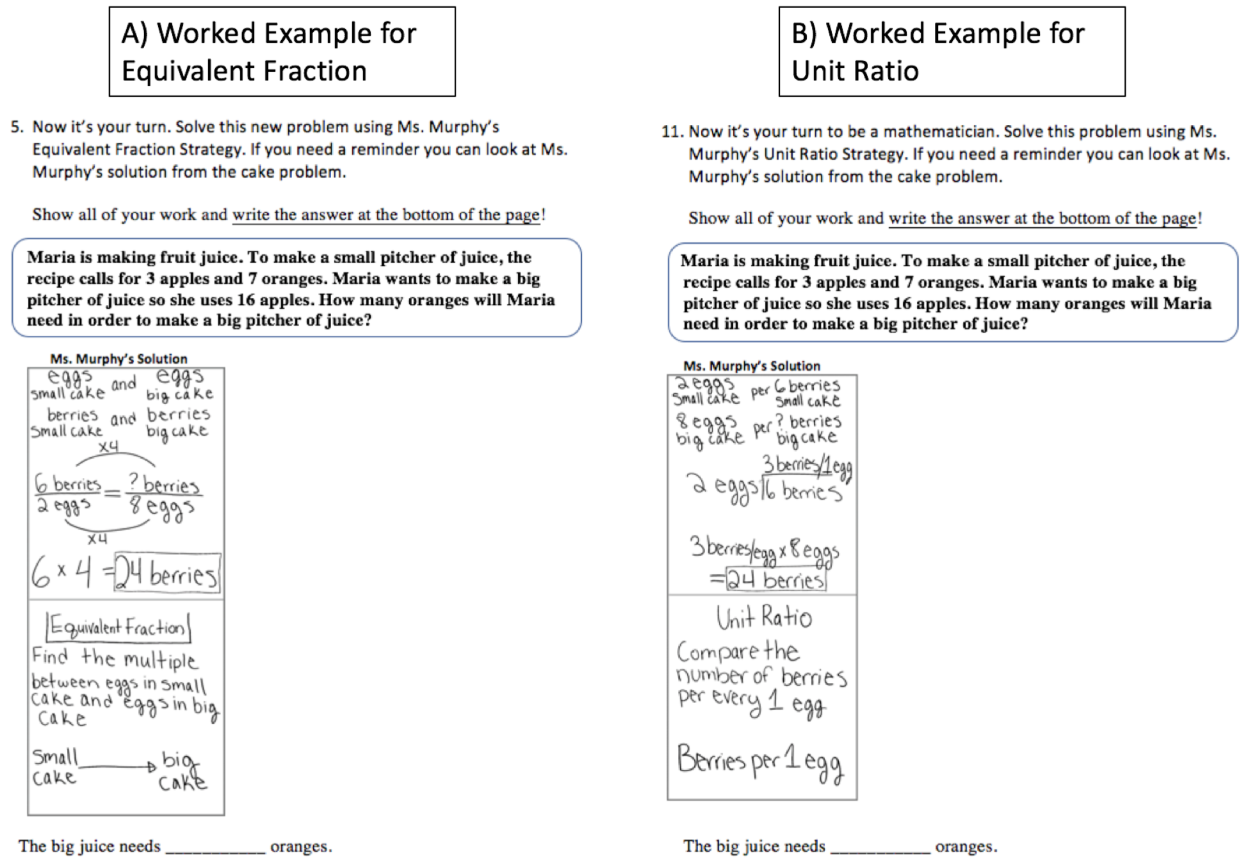
All students watched a high-quality video lesson on proportional reasoning. The audio stream of the students and teacher during the lesson was kept consistent across both conditions. A malleable factor of worked examples was then manipulated as a means to reduce burden on students' cognitive load and support mapping of solution strategies from the source to the target representation. Students were assigned to one of two experimental conditions: i) worked example during problem-solving or ii) no worked examples during problem-solving.

Students in the worked example condition received a screenshot of the solution strategy from the video during near-transfer problem-solving. This worked example was provided on the same page where students solved the problems; worked examples were spatially aligned to support mapping of procedures and reduce cognitive load (see Figure 3.1 for a worked example presented during of the equivalent fraction (A) and unit ratio (B) strategies). In addition to the fully worked out procedures of the solution strategies, the worked examples outlined the key comparisons to be made for the equivalent fraction strategy (i.e., "Find the multiple between eggs in the small cake and eggs in the big cake") and the unit ratio strategy (i.e., "Compare the number of berries per every 1 egg").

Students in the control condition did not have worked examples while problem-solving; students had to recall the solution strategy approach and procedures, which is a common approach in U.S. instruction (e.g., Richland et al., 2007).

Figure 3.1

Worked Examples Condition: Fully Worked Examples Spatially Aligned During Problem-Solving



Note. Worked examples are presented for students in the worked example condition on the near transfer problem. Students were asked to solve the problem using the equivalent fraction strategy (A) and unit ratio strategy (B). Students were provided with the strategy approach, setup, procedures, solution. Additionally, students were provided with the key conceptual framework for solving the problem.

2.6 Proportional Reasoning Problem-solving

During the video lesson, students were asked to solve one near-transfer proportional reasoning problem using the equivalent fraction strategy and then the same problem using the unit ratio strategy. For example, in the video lesson, after the student and teacher described the procedures of the equivalent fraction strategy, students in the study were asked to solve a near-transfer problem using the equivalent fraction strategy. After instruction of the unit ratio strategy on the same problem from a different student and the teacher, students in the study were asked to solve the same near-transfer problem using the unit ratio strategy.

Students in the worked example condition received the worked example while problem-solving whereas students in the no worked example condition were not provided with this *double-dose* of visual support. Students problem-solving approach, procedures, and final solution accuracy were used to assess whether the *double-dose* of visual support enhances learning of solution strategies and mapping of procedures during a conceptually challenging mathematics lesson.

2.6.2 Proportional Reasoning Problem-solving Coding Scheme

In order to assess learning gains from instruction to near transfer problems, we explored students' final solution accuracy using the equivalent fraction and unit ratio strategy on the same near-transfer proportional reasoning problem. We were also interested to examine students' ability to map solution strategies from instruction to the near-transfer problem. Therefore, we explored the nature and sequence of students' problem-solving decisions, from their initial attempt of the solution strategy to the setup and procedures and ultimately their final solution accuracy. The complete coding protocol is available on OSF and outlines the development of the codes, coding protocols and reliability checking. A brief outline of the coding protocol for the equivalent fraction strategy and unit ratio strategy is outlined next.

Qualitative codes were used to gather quantitative data about the emergent patterns in problem-solving. To examine students use of the equivalent fraction strategy, codes assessed if students: 1) attempted the instructed strategy, 2) numerically set up the problem as instructed, 3) related known units multiplicatively, 4) used the link found in the first relationship and applied it to the second relationship, and 5) arrived at the final accurate solution. A summary of the coding scheme for equivalent fraction word problems can be found in Table 3.1.

To examine students' use of the unit ratio strategy on the same near-transfer problem, codes assessed if students 1) attempted the instructed strategy, 2) numerically set up division step as instructed, 3) arrived at the accurate answer for the mid-point division step, 4) completed two required steps to the solution strategy (division then multiplication), 5) linked the unit ratio solution to the known large unit through multiplication or division, 6) numerically set up the multiplication step as instructed, and 7) arrived at the accurate final answer. A summary of the coding scheme for unit ratio word problems can be found in Table 3.1.

To establish inter-coder reliability, 20% of the data were coded by two research assistances. Krippendorff's alpha was used to assess inter-coder reliability for each variable coded (Hayes & Krippendorf, 2007). For the coding of the variable to be considered reliable it was required that the Krippendorff's alpha, an index that accounts for level of measurement and agreement expected by chance, was .70 or above, which is known to be conservative. Reliability for coding in the problem-solving steps (see Table 3.1) was $\alpha = 0.84$. Prior knowledge variables were coded $\alpha = 0.95$.

Table 3.1
Equivalent Fraction and Unit Ratio Strategy Coding Scheme for Problem-Solving Steps

Problem-solving Step	**Attempted Instructed Strategy**
Equivalent fraction strategy	
Attempt	Attempted equivalent fraction strategy (i.e., fraction or ratio) with numbers from problem and attempted to find relationship between numbers in problem
Number Setup	Attempted fraction setup creating proper or improper fractions within recipes
Procedure	Compared unit of small recipe to unit of large recipe using multiplication
Link	Solution calculated from the first operation used in the second operation
Accuracy	Accurate solution
Unit Ratio Strategy	
Attempt	Evidence of reducing a ratio and attempted to apply that number (i.e., division and multiplication)
Division Numbers Setup	Divided within the recipe to find the unit ratio
Division Accuracy	Accurate division solution
UR 2 Steps	Completed two steps of procedures: attempted division first, to find unit ratio and then multiplication to apply to large recipe
Link	Linked the unit ratio solution to large unit through multiplication/division (even if final number from the unit ratio division step was inaccurate)
Multiplication Numbers Setup	Set up multiplication using the quotient from the division step and apply it to the large recipe unit
Accuracy	Accurate solution

Note. This equivalent fraction and unit ratio coding schemes were applied to the near-transfer problem during the lesson when students were instructed to use the equivalent fraction strategy or unit ratio strategy. Problem-solving steps are used as the key indicators for the LCA analysis.

2.7 Individual Difference Measures

A set of measures assessing individual differences in executive functioning and achievement orientation were administered to all participants.

2.7.1 Prior Knowledge Assessment

All students completed a pretest assessment to measure prior knowledge for solving proportional reasoning word problems. Students were asked to solve two proportional reasoning word problems using any strategy of their choice. The missing value could be solved by comparing like terms or across units (e.g., A smoothie recipe calls for 2 bananas and 5 oranges. To make a bigger pitcher of the smoothie with 12 bananas, how many oranges should you add?). The same coding scheme from the assessment was applied for the prior knowledge of proportional reasoning problems.

2.7.2 Arithmetic Assessment

Arithmetic skills that are needed to solve proportional reasoning word problems using an equivalent fraction strategy and unit ratio strategy were also assessed. Students solved five missing-value ratio equivalence problems to assess their understanding for multiplicative relationships (Cramer & Post, 1993). The missing value in the ratios were the back term numerator and back term denominator. Potential solution strategies for comparison also varied: numerator to numerator (e.g., $\frac{5}{6} = \frac{15}{\ }$), denominator to denominator (e.g., $\frac{2}{3} = \frac{\ }{12}$) or front term numerator to front term denominator (e.g., $\frac{2}{12} = \frac{3}{\ }$). Lastly, students solved two long division problems and three decimal multiplication problems.

2.7.3 Attentional Control

The d2 Test of Attention is a group-administered, pen and paper measure which assesses students sustained levels of selective attention and attentional control (Brickenkamp & Zillmer, 1998; Rhonda & Ross, 2005). The assessment is normed with US and German children, adolescents and adults. Under a time pressure, participants are asked to search and cross off target characters (i.e., "d"s with two dashes surrounding it either above or below) from perceptually similar distractors (e.g., "d"s with one dash, "p"s with two dashes). The outcome of this task is the total number of items processed minus errors (TN-E). Across the literature, internal consistency is high ($\alpha \geq 0.8$) and test test-retest reliability is high ($\alpha > 0.8$; Clark, 2005). d2 TN-E scores correlate with other measures of attention and executive functioning: Stroop and Tower of London which support the validity of using this measure (Clark, 2005).

2.7.4 Working Memory (WM)

we assessed students' baseline working memory capacity with the Shortened Symmetry Span Task (Foster et al., 2015) formatted for play on a 5th generation Apple iPad with a 9.7-inch active screen. Instructions were presented on the screen and narrated through headphones. Participants were required to hold target information (highlighted cells in a matrix) in memory while attending to distracting information (whether or not an unrelated matrix is symmetrical across its vertical axis). Participants completed four test trials which randomly varied in length from two to five cells to be memorized, for a total possible score of 14 cells correctly memorized in order.

2.8 Analytical Approach

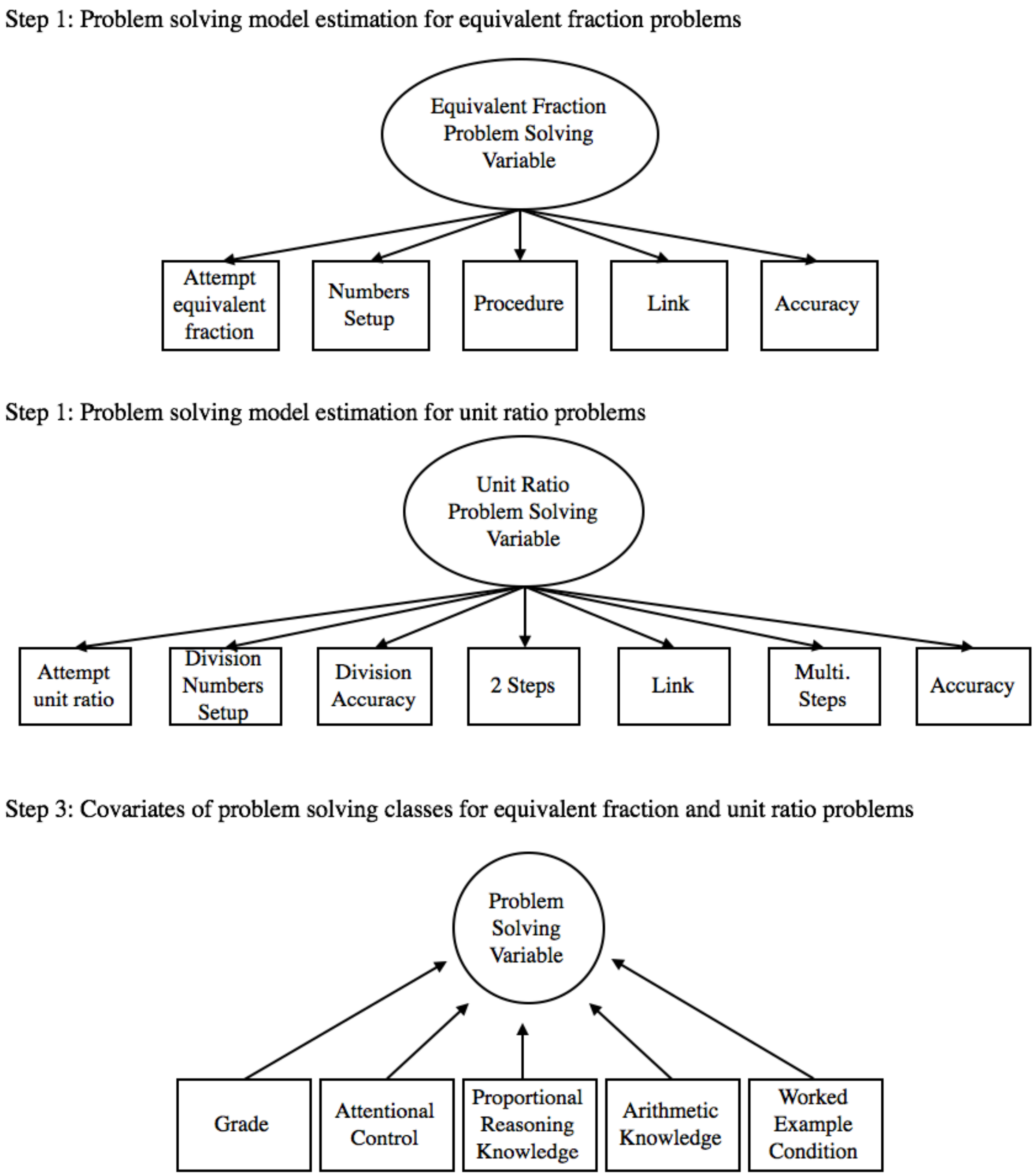
First, we examined the impacts of the worked examples during instruction on problem-solving accuracy. To do this, we conducted two binary logistic regressions for each solution strategy. For both equivalent fraction problems and unit ratio problems, the first regression examined accuracy predicted by condition; the second set of regressions examined accuracy predicted by condition and individual factors of attentional control, working memory, prior knowledge of proportions and arithmetic skills.

In order to capture a more fine-grained assessment of students' problem-solving gains, we examined the full nature and sequence of problem-solving from approach to procedures to final solution accuracy. To do so, we used a novel latent class approach. This analysis allowed us to capture students' complete problem-solving attempt; additionally, it allowed for us to examine different groups of students within the sample and their district problem-solving ability (Collins & Lanza, 2010). Due to the different nature of problem-solving steps for equivalent fraction and unit ratio strategy, this three-step LCA approach was conducted twice in Latent GOLD 6.0. For the equivalent fraction problem, we used the five problem-solving steps as the ordinal indicators for the model: attempt, number setup, procedure, link and accuracy. The LCA for the unit ratio problem included seven ordinal indicator variables: attempt unit ratio, division number setup, division accuracy, completed two steps, link multiplication number setup, and accuracy (see Figure 3.2 Step 1). The problem-solving steps which were the indicators used in the analyses were strongly related to the latent class variable, suggesting that the indicators were high quality.

This three-step approach involves analyzing data with a stepwise fashion: i) identification of latent classes from 1 to 7 (step 1), ii) classification of individuals into groups identified by the optimal latent class model (step 2), and iii) estimation of the relationship between the latent classes and external variables using maximum-likelihood approach estimates (step 3; Vermunt, 2010).

Figure 3.2

Conceptual Diagram and the Analytic Model



Note. Step 1 Model estimation: Latent Class Analysis (LCA) identifying the optimal proportional reasoning problem-solving model for equivalent fraction problems and unit ratio problems. Order indicates the order of proportional reasoning problem-solving steps (attempt instructed strategy, setup, link, accuracy). Step 3: Examination of the relationship between problem-solving groups and covariates.

Note that the data set contains missing values of covariates (i.e., attentional control, working memory, prior knowledge of proportions and arithmetic skills). For the logistic regressions, we addressed these omissions via multiple imputation the method of predictive mean matching, implemented using the *mice* library in R (van Buuren & Groothuis-Oudshoorn, 2011). See Section S 1.1 in Supplemental materials for details of the imputation procedure and alternate analyses with complete cases only.

3. Results

3.1 Pretest Accuracy

A one-way ANOVA was conducted first to establish that the randomization was successful and to confirm there were no differences between conditions at pretest. At pretest, there were no differences in proportional reasoning prior knowledge between conditions for 5th graders (F (1,164) = 0.10, p = 0.76) or 6th graders (F (1,56) = 0.13, p = 0.72). The proportion of students' problem-solving strategies and accuracy scores at each test point by condition are summarized in Table 3.2. At pretest, collapsing across grades, students used mostly incorrect, non-proportional reasoning strategy solutions (e.g., build up) and had inaccurate solutions. There was a higher proportion of 6th grade students who attempted proportional reasoning strategies compared to 5th grade students. Sixth grade students were more accurate than 5th grade students. Across both grades, a small proportion of students used the instructed equivalent fraction or unit ratio strategies.

3.2 Intervention Effects on Problem-solving Strategy and Accuracy

During the lesson, although students were directed to use the just instructed equivalent fraction and unit ratio strategy, students' problem-solving approaches varied by condition and grade. When instructed to use the equivalent fraction strategy, more students in the high visual support (worked example condition) attempted the equivalent fraction strategy compared to students in the low visual support (no worked example, problem only) condition. Students in the 6th grade used the instructed strategies more than students in the 5th grade. Accuracy scores were similar across conditions; 6th grade students had higher accuracy than 5th grade students. When instructed to use the unit ratio strategy, student approaches varied by grade but not condition. Students in the 6th grade attempted the instructed strategy more and had higher accuracy than the 5th grade students.

Table 3.2

Proportion of Students' Solution Strategies and Accuracy by Condition and Grade

	Problem Solving Attempt				Problem Solution Accuracy	
	at pretest					
	Attempt Equivalent Fraction or Unit Ratio Strategy	Attempt Other Proportional Reasoning Strategy	Attempt Non-Proportional Reasoning Strategy	Blank	Correct	Incorrect
5th Grade	.11	.25	.49	.14	.34	.66
6th Grade	.14	.36	.24	.26	.48	.52
Mean	**.13**	**.31**	**.37**	**.20**	**.41**	**.59**

	when instructed to use the equivalent fraction strategy																	
	Attempt Equivalent Fraction Strategy			Attempt Other Proportional Reasoning Strategy			Attempt Non-Proportional Reasoning Strategy			Blank			Correct			Incorrect		
	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**
5th Grade	.58	.39	**.48**	.12	.18	**.15**	.19	.38	**.28**	.11	.05	**.16**	.12	.10	**.11**	.88	.90	**.89**
6th Grade	.69	.70	**.69**	.17	.03	**.10**	.07	.24	**.15**	.07	.03	**.35**	.28	.21	**.25**	.72	.79	**.76**
Mean	**.62**	**.46**	**.54**	**.13**	**.14**	**.13**	**.16**	**.35**	**.25**	**.10**	**.04**	**.07**	**.20**	**.16**	**.18**	**.80**	**.85**	**.82**

	when instructed to use the unit ratio strategy																	
	Attempt Unit Ratio Strategy			Attempt Other Proportional Reasoning Strategy			Attempt Non-Proportional Reasoning Strategy			Blank			Correct			Incorrect		
	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**	WE	PO	**Mean**
5th Grade	.42	.37	**.40**	.11	.17	**.14**	.30	.35	**.33**	.17	.11	**.14**	.27	.23	**.25**	.73	.77	**.75**
6th Grade	.48	.52	**.50**	.24	.24	**.24**	.07	.14	**.11**	.21	.10	**.16**	.38	.35	**.37**	.62	.65	**.64**
Mean	**.45**	**.45**	**.45**	**.18**	**.21**	**.19**	**.19**	**.25**	**.22**	**.19**	**.11**	**.15**	**.33**	**.29**	**.31**	**.68**	**.71**	**.69**

Note. This table represents the proportion of students by condition and grade for different problem-solving approaches and solution accuracy at i) pretest, ii) when instructed to use the equivalent fraction strategy, and iii) when instructed to use the unit ratio strategy). WE signifies the high visual support (worked example) condition. PO signifies the low visual support (no worked example, problem only) control group. Bold represents means across conditions and grades.

3.3 Intervention Effects on Equivalent Fraction Problem-solving Accuracy

To examine whether visual support affects performance accuracy for the equivalent fraction strategy, a binomial logistic regression with imputed missing data was conducted (see Table 3.3, Model 1 and 2). Model 1 tests the independent effects of condition on accuracy, controlling for fixed effects of classroom; Model 2 adds in theoretical relevant covariates including attentional control, working memory, knowledge of proportions, arithmetic skills, and grade. High visual support via worked examples was not a significant predictor of students' accuracy scores $p > .05$. There is suggestive evidence that working memory and grade impact students' accuracy scores. Separate regression models were run to examine the effects of condition by grade (see Supplementary Materials S 1.2, Table S 1.1). Results are consistent with the overall sample model; there is no effect of condition on accuracy for equivalent fraction problems. See Supplementary Materials Table S 1.2 for an alternative analysis that used only complete cases.

3.4 Intervention Effects on Unit Ratio Problem-solving Accuracy

For the more sophisticated, unit ratio strategy, the same series of binomial logistic regressions were conducted (see Table 3.3, Model 3 and 4). High visual support via worked examples was not a significant predictor of students' accuracy scores $p > .05$. There is suggestive evidence that arithmetic skills impact students' accuracy scores. Separate regression models were run to examine the effects of condition by grade (see Supplementary Materials S 1.2, Table S 1.1). Results are consistent with overall model; there is no effect of condition on accuracy for unit ratio problems. See Table S 1.2 in the Supplemental Material for an alternative analysis that used only complete cases.

Table 3.3
Model Specifications of Accuracy when using the Equivalent Fraction and Unit Ratio Strategies

	Equivalent Fraction Strategy Accuracy									Unit Ratio Strategy Accuracy							
	Total Sample									Total Sample							
	Model 1				Model 2					Model 3				Model 4			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*		b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.30	.39	-.77	.44	.31	.40	-.78	.44	High Visual Support	.18	.30	-.60	.55	.25	.32	-.78	.44
Attentional Control	-	-	-	-	.00	.00	.42	.68	Attentional Control	-	-	-	-	.00	.00	1.62	.11
Working Memory	-	-	-	-	*-.15*	*.09*	*-1.71*	*.09*	Working Memory	-	-	-	-	-.09	.07	-1.30	.19
Knowledge of Proportions	-	-	-	-	.13	.17	.73	.46	Knowledge of Proportions	-	-	-	-	.20	.14	1.42	.16
Arithmetic Skills	-	-	-	-	.11	.09	1.19	.23	Arithmetic Skills	-	-	-	-	**.16**	**.07**	**2.25**	**.02**
Grade	-	-	-	-	**.87**	**.44**	**1.98**	**.05**	Grade	-	-	-	-	.22	.37	.61	.54

Note. High visual support (worked example) condition is compared to the low visual support condition (no worked example, problem only). Covariates are standardized.

3.5 Latent Class Analysis for Problem-solving Steps

There were no effects of condition on student performance potentially due to the difficulty and novelty of the solution strategies. Therefore, to further examine students' learning gains from the lesson and determine if worked examples provide extra support when problem-solving, we used a latent class analysis to examine the nature and sequence of students' problem-solving from attempt and setup to final solution accuracy. We conducted a LCA three-step approach to identify and characterize patterns of problem-solving steps as a function of grade, attentional control, working memory, prior knowledge, arithmetic knowledge and visual support condition in Latent GOLD 6.0. We conducted two LCA analyses: one for the equivalent fraction problem and one for the unit ratio problem.

The first step of the three step LCA analysis is building a latent class model for a set of indicator variables (i.e., problem-solving steps). Then cases are assigned to latent classes. Third, the latent classification scores are related to external variables of interest, correcting for the classification error to prevent bias (Bolck, Croon and Hagenaars, 2004; Vermunt, 2010; Bakk, Tekle and Bermunt, 2013).

First we will describe the latent class models for equivalent fraction problem-solving steps and subsequently describe the latent class models for the unit ratio problem-solving steps.

3.5.1 Step 1: Model Fit for Equivalent Fraction Problems

Model fit information for equivalent fraction problems can be found in Table 3.4. The AIC and SABIC indicate the two and three-group models fit best, the BIC indicates the 2-group model fits best. A conditional bootstrap procedure was conducted to compare the two-, three-, and four-group models (Vermunt & Magidson, 2016). The bootstrap likelihood difference analysis shows significantly more information was gained from the three-group model than the two-group model (Bootstrap 2LL Diff = 11.64, p= 0.04); there were no significant differences between the three-group model and the four-group model (Bootstrap 2LL Diff = 5.87, p = 0.12). Therefore, the equivalent fraction problem three-group model was deemed the best model for statistical fit, parsimony, and theoretical interpretability in characterizing patterns of proportional reasoning problem-solving.

Table 3.4
Model Fit Indicators for LCA of Equivalent Fraction Problem-Solving

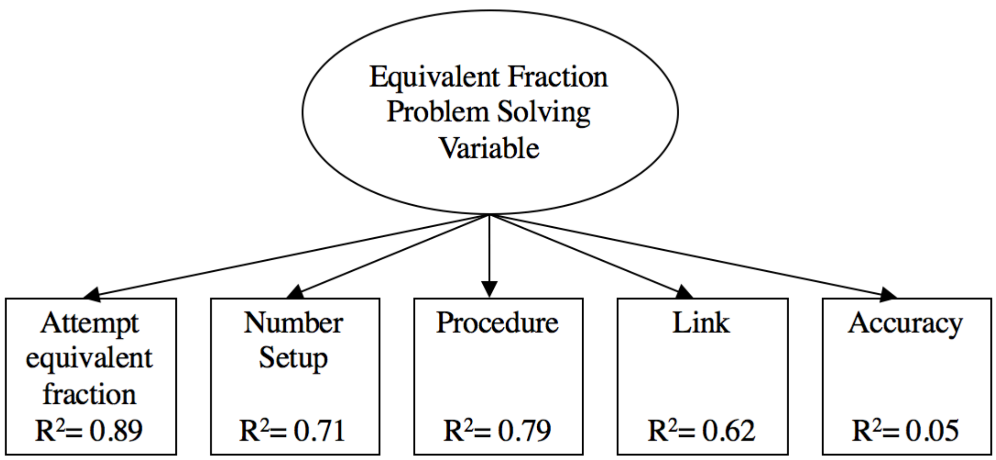
	LL	BIC	AIC	SABIC	CLC	ICL-BIC	Entropy R^2
1-Group	-685.54	1398.19	1381.09	1382.35	1371.09	1398.19	1.00
2-Group	-459.95	979.52	941.90	944.66	930.33	989.96	0.97
3-Group	**-454.13**	**1000.40**	**942.25**	**946.53**	**982.01**	**1074.15**	**0.84**
4-Group	-451.19	1027.05	948.38	954.16	1013.87	1138.54	0.80
5-Group	-450.68	1058.56	959.37	966.65	1016.33	1173.52	0.80
6-Group	-450.46	1090.63	970.91	979.71	1103.27	1292.99	0.69
7-Group	-450.32	1122.88	982.64	992.94	1102.77	1349.69	0.70

Note. Bold text indicates the selected model. LL, Log Likelihood; BIC, Bayesian Information Criterion; CAIC, Consistent Akaike Information Criteria; AWE, Average Weight of Evidence; ICL-BIC, a version of Integrated Classification Likelihood.

3.5.2 Step 2: Characterization of Equivalent Fraction Problem-solving Groups

The cluster loadings of each problem-solving indicator can be found in Figure 3.3. These can be interpreted similar to factor loadings in factor analyses. The loadings are obtained using a linear approximation of the class specific response probabilities (Vermunt & Magidson, 2005). For equivalent fraction problem-solving we find that problem accuracy is not an influential indicator to class membership. Rather, attempt equivalent fraction, number setup, and procedure are most predictive of class memberships.

Figure 3.3
Cluster Loadings of the Problem-Solving Indicators for the Equivalent Fraction Problem Estimated in the Latent Class Analysis



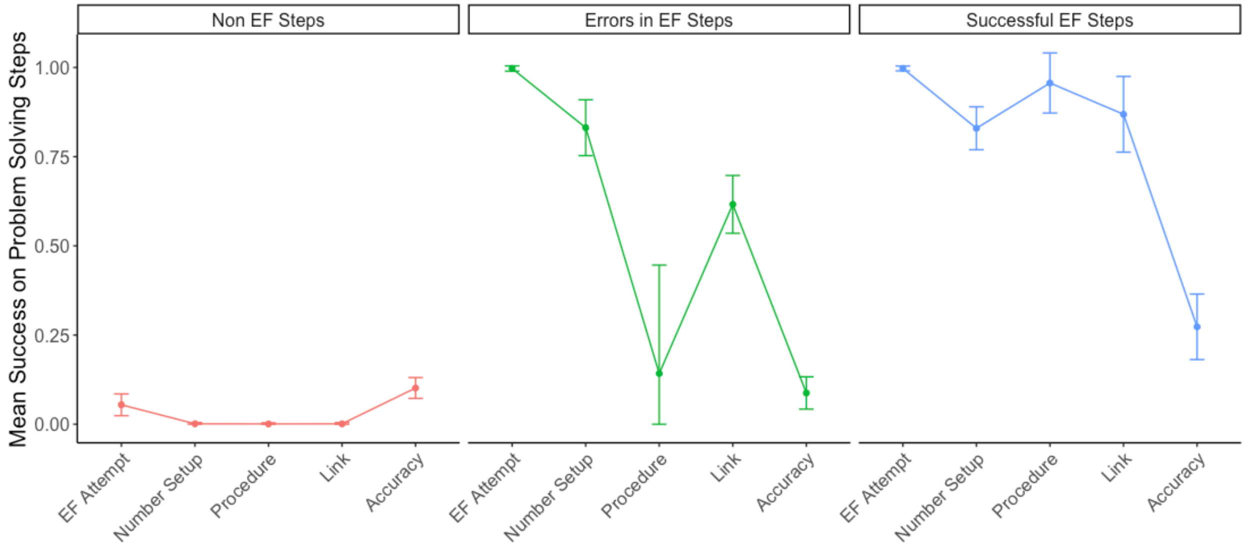
3.5.3 Latent Classes for Equivalent Fraction Problem-solving

Next we looked at the specific latent classes that emerged from the item-response probabilities for the equivalent fraction problem; the groups are summarized in Figure 3.4. These latent classes represent students' distinct problem-solving abilities from setup to solution accuracy. The groupings allow us to breakdown the specific steps required for problems solving and allows us better understand the processes by students are mapping the instructed procedures to the near transfer problem.

One group was labeled 'Non Equivalent Fraction (EF) Steps'; this group was characterized by no attempt of the solution strategy and low overall accuracy for each step in the equivalent fraction problem-solving; these students did not map the source to the target in any of the steps; 50% of the total sample were in this group. A second group was labelled 'Errors in EF Steps' and 25% of the total sample was assigned to this group. This group was characterized by high rates for attempt of equivalent fraction and numerical setup with low rates of procedural success (i.e., linking through multiplication/division) resulting in low overall accuracy. Students in this group mapped only the first few steps of problem-solving resulting in more surface level attempts. A third group was characterized by structural mapping on multiple levels; these students successfully attempted the equivalent fraction strategy, setup, procedure, and link with a decline in accuracy. This decline in accuracy can, in part, be explained by the difficulty students have when solving for multiplication problems with decimals. This group was labeled 'Successful EF Steps'; the latent class probability for this group was 25%. The latent class probability can be interpreted as the proportion of the sample assigned to this group.

Figure 3.4

Equivalent Fraction LCA Problem-Solving Groups



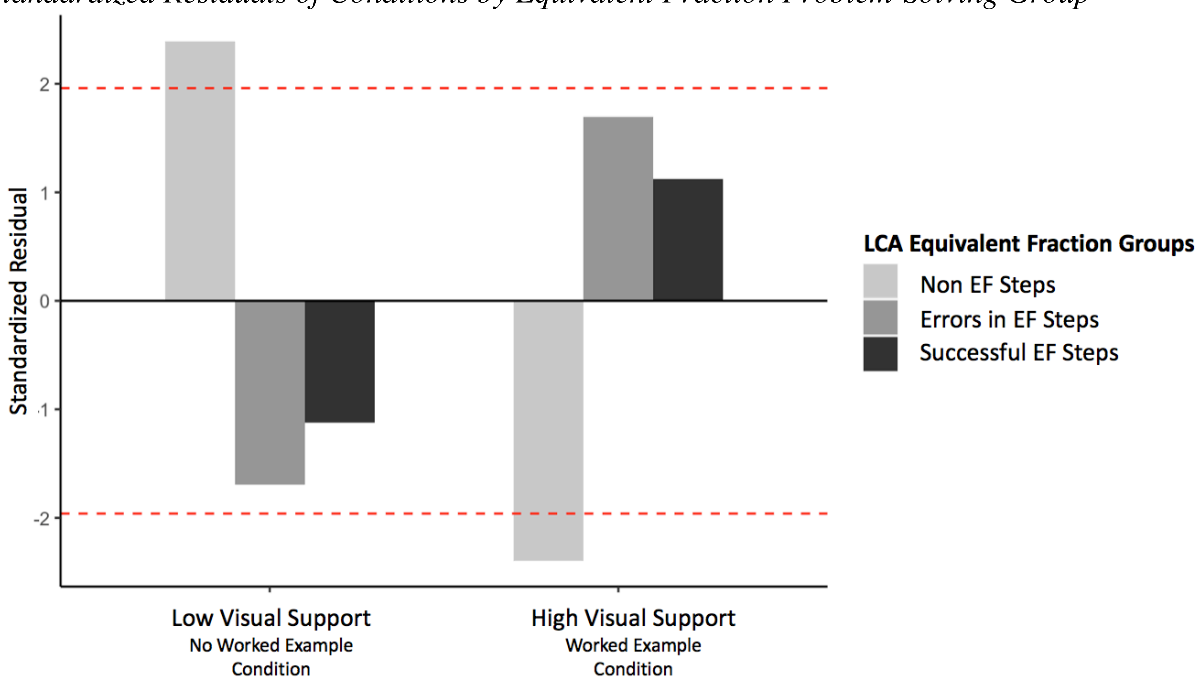
Note. Three problem-solving groups emerged from the LCA analysis for the equivalent fraction strategy: 'Non EF Procedures', 'EF Procedural Difficulty', 'Successful EF Steps. Error bars indicate standard error. X-axis represent the sequence of problem-solving steps for the equivalent fraction problems: attempt equivalent fraction, number setup, procedure, link, and accuracy. The mean on the y-axis represents mean accuracy for each problem-solving step.

3.5.4 Step 3: Equivalent Fraction Problem-solving Group Membership by Experimental Condition

Next, we explored the relationship between experimental conditions and problem-solving group membership for the equivalent fraction problem. We examined the ordinal-by-ordinal gamma statistic to determine whether equivalent fraction LCA groups and intervention conditions showed ordered patterns of association. In this analysis, we considered an order of intervention condition of increasing support from low visual support to high visual support, and an order of Equivalent Fraction LCA groups of increasing sophistication from 'Non EF Steps', to 'Errors in EF Steps, to 'Successful EF Steps.' we also assessed standardized adjusted residuals to determine whether the cross-classification of subgroup-by-condition patterns occurred more/less often than expected by chance.

Cross-classification analysis showed equivalent fraction LCA groups and condition were significantly associated (χ^2 (2, n = 226) = 5.98, p = 0.05; γ = -0.24, p = 0.04); the significant, positive gamma statistic suggests an overall trend for children with least support belonging to less accurate problem-solving procedures, and higher levels of support not likely to belong to the less accurate problem-solving groups. Examination of the standardized residuals (Figure 3.5) shows students in the high visual support condition, are not likely to belong to the 'Non EF Steps' group and were trending but not significant to belong to the 'Successful EF Steps' group. Students in the low visual support condition are most likely to belong to the 'Non EF Steps' group.

Figure 3.5
Standardized Residuals of Conditions by Equivalent Fraction Problem-Solving Group



Note. Standardized Residuals of students from each condition (e.g., high visual support and low visual support) by the LCA equivalent fraction problem-solving groups. If the residual is greater than 1.96, it indicates a greater than chance likelihood that the students assigned to that condition will belong to the equivalent fraction LCA group. If the residual is less than 1.96, then it indicates a lower than chance likelihood that a student assigned to the condition will belong to that particular equivalent fraction LCA groups. The pattern is not significant when -1.96 < residual < 1.96.

3.5.5 Covariates of Equivalent Fraction Problem-solving Group Membership

Finally, we investigated the relationship between proportional reasoning problem-solving groups and covariates using Latent GOLD 6.0 step 3 analysis. Grade, arithmetic skills, attentional control, and working memory were entered as covariates. Experimental condition was evaluated for its effects on equivalent fraction problem group membership. Model outcomes are shown in Table 3.5.

The Wald's tests in Table 5.5 indicates condition, grade, and arithmetic skills predicted equivalent fraction problem-solving group membership. There was no effect of attentional control or working memory on group membership. Coefficients indicate probability of belonging to three problem-solving groups: 'Successful EF Steps', 'Errors in EF Steps, and 'Non EF Procedures'. Higher positive values indicate a higher probability of belonging to a certain group, and lower negative values indicate a lower probability of belonging to a certain group.

Examination of the coefficients in Table 3.5 and proportion of students by condition suggest that students in the high visual support condition were not likely to be in the 'Non EF Steps' group. That is, providing students with more visual support suggests to prevent people from belonging to the lowest performing group. Grade is also a strong predictor of group membership. 6^{th} grade students were more likely to be in the 'Successful EF Steps' group than 5^{th} grade students. Lastly, students' arithmetic knowledge was predictive of group membership. Students with higher arithmetic skills were more likely to be in the 'Errors in EF Steps' group and not likely to be in the 'Non EF Steps' group.

Table 3.5
Experimental Condition and Cognitive Covariates of Equivalent Fraction Problem-Solving Group Membership

			Non EF Steps		Errors in EF Steps		Successful EF Steps	
	Wald	p-value	B	SE	B	SE	B	SE
High Visual Support	**8.33**	**.02**	**-.29**	**.10**	.25	.13	.04	.13
6th Grade	**8.26**	**.02**	.19	.12	.18	.16	**-.37**	**.14**
Arithmetic Skills	**9.71**	**.01**	**-.11**	**.04**	**.16**	**.06**	-.05	.05
Attentional Control	.09	.96	.00	.00	.00	.00	.00	.00
Working Memory	3.92	.14	-.09	.05	.09	.07	.01	.06

Note. Bold indicates experimental significance. Conditions: High Visual Support and Low Visual Support.

3.6.1 Step 1: Model Fit for Unit Ratio Problems

The fit criteria indicated the four-group model was the best fitting model based on the BIC, AIC, SABIC, and CLC (Table 3.6). The three-group model was identified as the optimal model by the ICL-BIC). Both class error and entropy values for the four-group model was acceptable, indicating good classification of individuals into latent groups. A conditional bootstrap procedure was conducted to compare the three-, four-, and five-group models (Vermunt & Magidson, 2016). The bootstrap likelihood difference analysis shows significantly more information was gained from the four group model than the three group model (Bootstrap 2LL Diff = 46.43, $p < .0001$); there was not a significant difference between the four-and five-group model (Bootstrap 2LL Diff = 6.08, $p = 0.26$). Therefore, for the following analyses, we chose the four-group model for the equivalent fraction problem because of its statistical fit, parsimony, and theoretical interpretability in characterizing patterns of unit ratio proportional reasoning problem-solving.

Table 3.6
Model Fit Indicators for LCA of Unit Ratio Problem-Solving

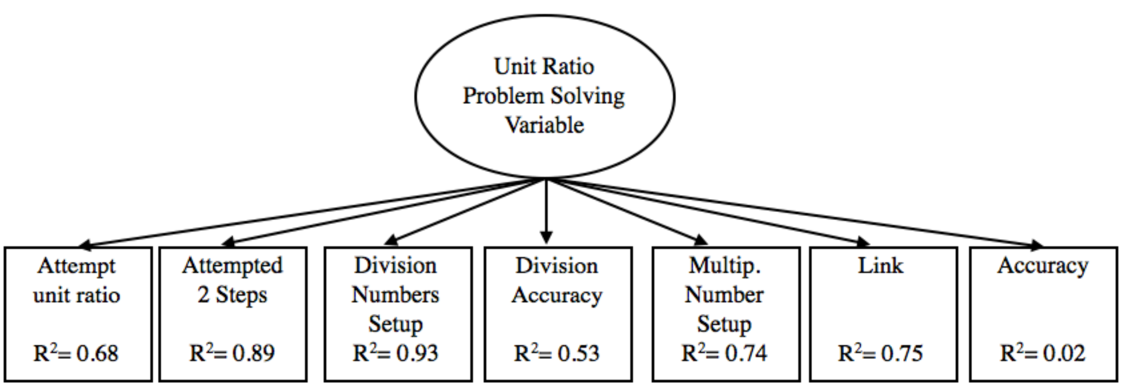
	LL	BIC	AIC	SABIC	CLC	ICL-BIC	Entropy R^2
1-Group	-766.13	1570.21	1546.26	1548.02	1532.36	1570.21	1.00
2-Group	-562.71	1206.74	1155.43	1159.20	1139.95	1221.26	0.95
3-Group	-519.82	1164.30	1085.63	1091.41	1046.24	1170.91	0.98
4-Group	**-496.60**	**1161.23**	**1055.20**	**1062.99**	**1009.22**	**1177.26**	**0.96**
5-Group	-493.56	1198.52	1065.11	1074.91	1013.27	1224.67	0.95
6-Group	-492.93	1240.63	1079.86	1091.67	1112.67	1303.69	0.80
7-Group	-492.82	1283.77	1095.64	1109.46	1158.20	1414.05	0.75

Note. Bold text indicates the selected model. LL, Log Likelihood; BIC, Bayesian Information Criterion; CAIC, Consistent Akaike Information Criteria; AWE, Average Weight of Evidence; ICL-BIC, a version of Integrated Classification Likelihood.

3.6.2 Step 2: Characterization of Unit Ratio Problem-solving Groups

The cluster loadings of each problem-solving indicator can be found in Figure 3.6. These can be interpreted similar to factor loadings in factor analyses. For unit ratio problem-solving we find that accuracy is not an influential indicator to class membership. Attempted 2 steps of the unit ratio procedures, division numbers setup, multiplication numbers set up, and link strongly predict class membership all with R^2 over 0.70.

Figure 3.6
Cluster Loadings of the Problem-Solving Indicators for the Unit Ratio Problem Estimated in the Latent Class Analysis



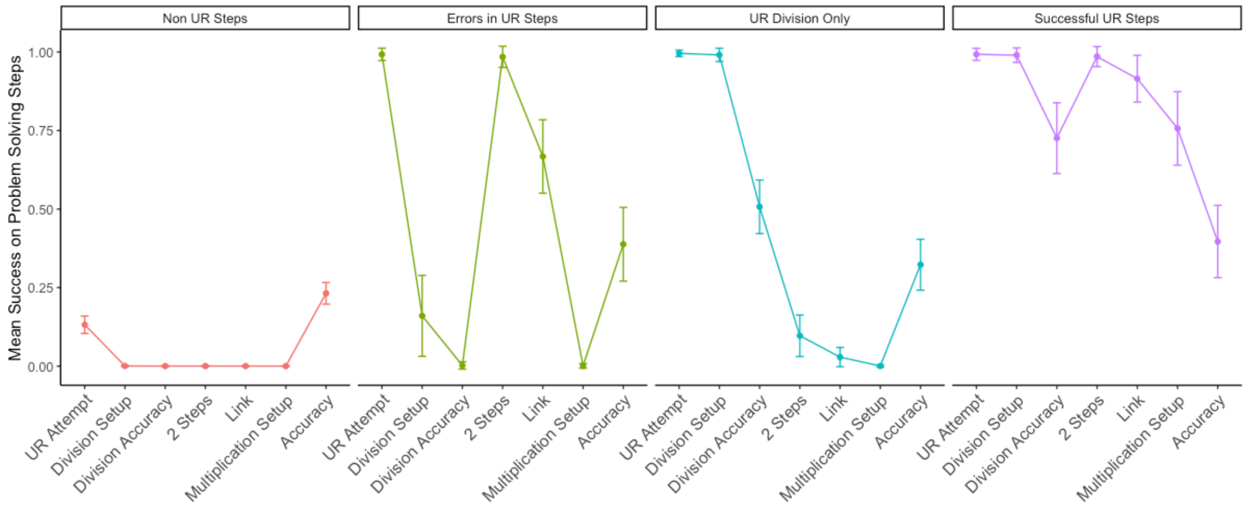
3.6.3 Latent Classes for Equivalent Fraction Problem-solving

Four latent classes emerged from the item-response probabilities for the unit ratio problem; the groups are summarized in Figure 3.7.

The first group was labeled 'Non UR Steps; this group was characterized by no attempt of the solution strategy and low overall accuracy for each step in the unit ratio problem-solving steps; 66% of the total sample were in this group. The second group was labeled 'Errors in UR Steps'; this group was characterized by high rates of unit ratio attempt but procedural errors across the full sequence of problem-solving steps mapping again based on surface level features; 9% of the total sample were in this group. A third group was labelled 'UR Division Only' and 16% of the total sample was assigned to this group. This group was characterized by high rates for attempt of unit ratio and division numerical setup with low rates of division accuracy. These students completed surface level but not structural level mappings. They mapped the division procedure from the source to the target but did not map that there are two crucial steps for problem-solving using the unit ratio strategy (i.e., division and multiplication) resulting in low overall accuracy. The most sophisticated group characterized by successful problem-solving attempt across all steps. There was a decline in overall accuracy. These students structurally mapped the source to the target and aligned and executed the multipole steps to problem-solving. This group was labeled 'Successful UR Steps'; 9% of the sample was assigned to this group.

Figure 3.7

Unit Ratio LCA Problem-Solving Groups



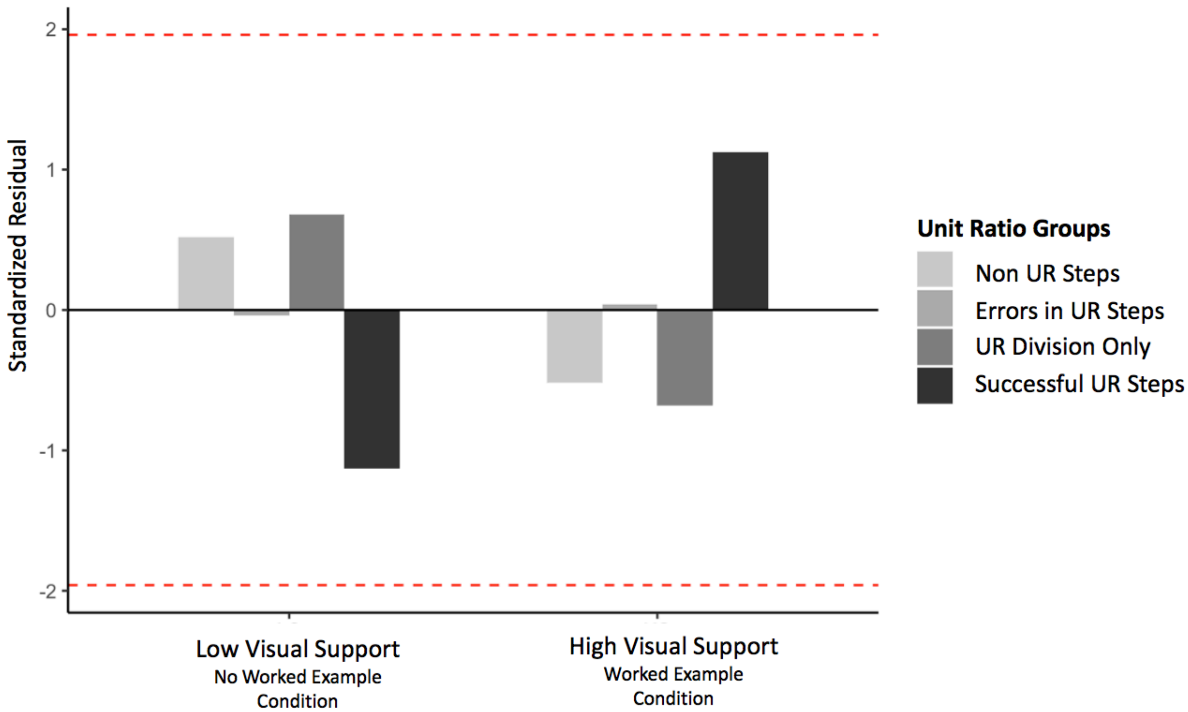
Note. Three problem-solving groups emerged from the LCA analysis for the equivalent fraction strategy: 'Non UR Steps', 'Errors in UR Steps', 'UR Division Only', and 'Successful UR Steps'. Error bars indicate standard error. X-axis represent the sequence of problem-solving steps for the equivalent fraction problems: attempt unit ratio, number setup, procedure, link, and accuracy. The mean on the y-axis represents mean success for each problem-solving step.

3.6.4 Step 3: Unit Ratio Problem-solving Group Membership by Experimental Condition

Next, we explored the relationship between experimental conditions and problem-solving group membership for the unit ratio problem. We examined the ordinal-by-ordinal gamma statistic to determine whether equivalent fraction LCA groups and intervention conditions showed ordered patterns of association. In this analysis, we considered an order of intervention condition of increasing support from low visual support to high visual support, and an order of Unit Ratio LCA groups of increasing sophistication from 'Non UR Steps', to 'Errors in UR Steps', to 'UR Division Only', to 'Successful UR Steps'. We also assessed standardized adjusted residuals to determine whether the cross-classification of subgroup-by-condition patterns occurred more/less often than expected by chance.

Cross-classification analysis showed equivalent fraction LCA groups and condition were significantly associated (χ^2 (3, n = 226) = 1.57, p = 0.67; γ = 0.08, p = 0.50); the non-significant, positive gamma statistic suggests and no overall trend between condition and LCA groups. Examination of the standardized residuals (Figure 3.8) shows no overall trends of experimental condition and LCA groups.

Figure 3.8
Standardized Residuals of Students in Condition by Unit Ratio Problem-Solving Groups



Note. Standardized Residuals of students from each condition (e.g., high visual support and low visual support) by the LCA unit ratio problem-solving groups. If the residual is greater than 1.96, it indicates a greater than chance likelihood that the students assigned to that condition will belong to the equivalent fraction LCA group. If the residual is less than 1.96, then it indicates a lower than chance likelihood that a student assigned to the condition will belong to that particular equivalent fraction LCA groups. The pattern is not significant when -1.96 < residual < 1.96.

3.6.5 Covariates of Unit Ratio Problem-solving Group Membership

Finally, we investigated the relationship between proportional reasoning groups and covariates for which we used the step-three maximum likelihood method. Grade, arithmetic skills, attentional control, and working memory were entered as covariates. Experimental condition was evaluated for its effects on unit ratio problem group membership. Model outcomes are shown in Table 3.7.

The Wald's tests indicate condition, attentional control, and knowledge of proportions predicted unit ratio problem-solving group membership. There was no effect of grade or arithmetic skills on group membership. Coefficients indicate probability of belonging to problem-solving groups: "Non UR Steps", "Errors in UR Steps", "UR Division Only", and "Successful UR Steps". Higher positive values indicate a higher probability of belonging to a certain group, and lower negative values indicate a lower probability of having a certain group. Examination of the coefficients suggest that visual support via a worked example does not improve problem-solving steps when attempting the more sophisticated, unit ratio strategy. However, students' arithmetic skills are a predictor to determine if group membership such that students with lower arithmetic skills are more likely to be in the "Non UR Steps" group and students with higher arithmetic skills are likely to be in the "Successful UR Steps" group.

Table 3.7
Experimental Condition and Cognitive Covariates of Unit Ratio Problem-Solving Group Membership

			Non UR Steps		Errors in UR Steps		UR Division Only		Successful UR Steps	
	Wald	p-value	B	SE	B	SE	B	SE	B	SE
High Visual Support	1.93	.59	-.11	.12	.14	.16	-.13	.20	.09	.20
6th Grade	1.39	.71	.04	.13	-.01	.17	.18	.24	-.21	.20
Arithmetic Skills	**11.80**	**.01**	**-.17**	**.05**	-.01	.07	-.01	.09	**.19**	**.08**
Attentional Control	7.23	.07	.00	.00	.00	.00	.00	.00	.01	.00
Working Memory	1.81	.61	-.04	.05	-.02	.08	.11	.09	-.04	.09

Note. Bold indicates experimental significance. Conditions: Low visual support (dummy variable) and High Visual Support.

4. Discussion

This research was intended to assess whether a *double-dose* of high-quality instruction paired with visual support through worked examples enhances learning and mapping of novel solution strategies and procedures more than high quality instruction alone. Problem solving tends to be difficult to change, with misconceptions often persisting even following traditional classroom instruction (Booth, Koedinger, & Stigler, 2007). This study, however, suggests that worked examples support engagement with novel solution strategies and enhance problem solving in new contexts. Though the manipulation was short term, with administration including only one classroom lesson, the changes in problem solving that were visible through a deep analysis of students' solution attempts as they engaged with the instruction were noteworthy. Therefore, our data encourage the interpretation that worked examples should be considered as a tool to implement in more learning contexts.

The study itself also serves as a model for a new approach to studying the effects of instructional interventions. We not only looked at students' accuracy in the problem-solving context but also used rigorous, qualitative codes with quantitative reliability and latent class analytics to gather close data on how students approached each problem (i.e., with proportionally or using non-proportional thinking) and the specific steps students took to solve problems. This microanalytic approach allowed for us to examine the specific steps at which students mapped or failed to map the source representation (i.e., the instructed problem) to the target representation (i.e., the near-transfer problem). Mathematics instruction with multiple solutions also allowed us to examine how these interventions support mapping in more familiar problem-solving strategies (i.e., equivalent fraction strategy) or more sophisticated and novel problem-solving approaches (i.e., unit ratio strategy).

Results from this study revealed no differences in students' accuracy scores by intervention condition for either the equivalent fraction problem or unit ratio problem. This was not entirely surprising as the near-transfer problems required arithmetic for the link and final solution with non-whole numbers and non-divisible relationships which represent an area students struggle with (Fuson & Abrahamson, 2005). Additionally, we did not specifically instruct on arithmetic procedures. Future lessons may place a greater emphasis on instruction for the arithmetic required when solving proportion reasoning problems.

Findings from the LCA analysis, however, suggest that there are distinct groups of students within sample characterized by their distinct problem-solving abilities. Key subgroups arose for both the equivalent fraction strategy and unit ratio strategy. A surprising finding was that over half of the students, mostly students from the no worked example condition, did not even attempt the instructed strategies on the near-transfer problems, despite being explicitly asked to do so. This suggests that without support via worked examples, students may be nervous or reluctant to try new solution strategies. During a complex lesson which involves problem-solving, students may not feel equipped to approach the problem using the just instructed, novel strategy. However, when there is some level of support via worked examples and teachers scaffold learning, students feel more comfortable, confident, and are willing to attempt the novel solution strategy. These findings indicate just how important instruction and scaffolding are to support students on their journey to fundamentally change the way they are thinking and the strategies they use.

Another group that emerged from the LCA was characterized by key misconceptions in proportional thinking within problem solving. For the equivalent fraction strategy, students had difficulty at the "procedure" step. This highlights that the multiplicative nature of proportional reasoning (instead of using addition) is highly complex and not easily mastered during a short video lesson (Parish, 2010 & Nabors, 2003). For the unit ratio strategy, the key misconception was that students compared the wrong relations- more specifically, the students' setup the equations for the strategy but did not compare numbers to create the "unit". Another key group that emerged from the unit ratio strategy problem-solving was students that mapped relations only on one level. Students in this group applied instructed proportional approach and attended to the surface level features, but they did not successfully and structurally map all steps of the source reference to the target reference. Multiple levels and multistep procedures of this complex strategy led students to attend only to surface level features (the first set of procedures) and not the structural level reasoning that was required.

Finally, across both strategies, the highest performing group that emerged from the LCA analyses was characterized by students mapping problem-solving approaches and procedures accurately but ultimately not arriving at the accurate solution. For the equivalent fraction strategy, students in the worked example condition were most likely to be in this group. For the unit ratio strategy, the trend was similar but the effects were not significant. Students in the 5th and 6th grade have been introduced to multiplicative problem-solving strategies and therefore, the equivalent fraction strategy builds on more familiar mathematics concepts both conceptually and procedurally. This may in part explain why there are strong effects for the equivalent fraction strategy compared to the unit ratio strategy.

Across all groups, students struggled with their arithmetic skills in long division and multiplication with decimals; while this is in line with previous literature (Rittle-Johnson, Siegler, Alibali, 2001), research and curricula should focus on developing students' skills in this area. This also suggests future research and curricula should critically examine when and how non-divisible numbers are incorporated into lesson learning environments and more carefully attend to the pedagogical features of doing so.

Previous research has shown that prior knowledge and familiarity with the source representation are key factors that predict whether the students will be able to map the source representation to the target representation (Richland, Holyoak, & Stigler, 2004). Findings from this study provide further evidence that procedural fluency and prior knowledge in the domain is predictive of learning gains (Fujimura, 2001; Gentner & Tattermann, 1991).

Findings from this study provide further evidence that the ability to carry out a series of actions to solve a problem can be supported via worked examples (Rittle-Johnson, Stiegler, & Alibali, 2001) but only for the more familiar, equivalent fraction strategy. However, for more complicated and more novel solution strategies, visual support and worked examples do not provide enough scaffolding to support students to try new solution strategies and map key procedural steps.

Future research should examine the effects of worked examples during instruction in a longer, more extensive learning format. Within one mathematics class session, we were only able to administer one instructed example for each solution strategy. In an ideal setting, multiple problems paired with multiple worked examples would be useful to examine how a *double-dose* of support not only enhances procedural gains but also encourages use of novel solutions strategies. While trending, the lack of significance for the unit ratio strategy suggested that for this sample at this age, the unit ratio strategy may have been to sophisticated. Future studies should examine these effects for students at varying ages.

One aspect of Study 3, presented next, builds off the findings from Study 2 by examining how visual support impacts learning in two cognitive domains. Specifically, Study 3 examines how visual support may be used as a tool to support 1) higher order thinking about the multiple solutions presented and 2) the development of problem solving skills on an immediate post-lesson test and a delayed, follow-up tests.

CHAPTER FOUR: PRESSURE IMPACTS LEARNING OF COGNITIVELY DEMANDING AND CONCEPTUALLY COMPLEX MATERIAL

1. Introduction

Mathematics is a domain in which anxiety and performance pressure are often heightened. Performance pressure, defined as an anxious desire to perform at a high level in a given situation (Hardy et al., 1996) leads students to explicitly monitor their anxious behaviors, consuming cognitive resources, leaving fewer resources for task-related processes (see Beilock & Carr, 2001; Maloney et al., 2013). The worries and anxious ideation from performance pressure can overload resources of the executive functioning system (e.g., Attali et al., 2011; Levitt et al., 2016) can result in underperformance in cognitively demanding testing contexts (see Maloney et al., Beilock, 2013).

Despite its relevance and importance, mathematical cognition under pressure has not been comprehensively researched or evaluated, specifically the effects of pressure on mathematics *learning*. If students experience pressure during a cognitively demanding mathematics lesson, their ability to engage in higher order reasoning and learn novel solution strategies and their procedures may be impaired. Higher order thinking, a type of relational reasoning where two representations are linked together in some way (Lewis & Smith, 1993; Richland & Simms, 2015) may require more cognitive processing than other tasks but may also have greater benefits for learning (Bloom, 1956; Common Core State Standards for Mathematics, 2010; National Research Council, 2012). Thus when combined, the effects of pressure and its implications on engaging with higher order thinking may be consequential but are not well understood.

If pressure does impair *learning* by reducing available executive functioning horsepower, do pedagogical approaches that offload working memory have the potential to reduce the negative effects of pressure? An extensive body of research in mathematics learning has shown that providing high visual support through different modalities such as worked examples, simultaneous presentation, and linking gestures, reduce students' cognitive resource load and improve learning outcomes (Richland et al., 2007; Paas, 1999; Pass et al., 2003). Yet, no research to date has examined if these visual supports can reduce the negative effects of pressure during learning and how it may differentially impact engagement in higher order thinking and problem-solving.

In this study, we report on an experiment that was designed as a two by two study manipulating visual support (worked examples as described in Study 2 with the addition of simultaneous presentation and linking gestures), and the presence or absence of an imposed pressure. We tested the hypotheses that there will be a main effect of high visual support such that students with high visual support will have 1) more instances of higher order thinking regarding the multiple instructed solution strategies, 2) higher problem solving solution accuracy, and 3) more learning and retention of novel solution strategies (i.e., higher rates of attempt and accurate setup of instructed solutions). Additionally, in this study we impose pressure as a way to tax cognitive resources and test the hypothesis that high visual support can improve outcomes for students with temporarily low processing resources due to imposed pressure during a cognitively taxing lesson. We used the mathematical context of proportional reasoning because this is a domain that is particularly affected by cognitive load and complex strategy learning.

Instances of higher order thinking were assessed by students' responses to questions that prompted comparison, abstraction and inference within and across solution strategies (Richland & Simms, 2015). We assessed problem solving gains by exploring students' performance on proportional reasoning problems immediately after the lesson and one week's delay. In addition to examining students' accuracy, we also took a more nuanced and detailed approach to problem solving gains by examining the nature and sequence of students' problem solving, that is to determine if students are attempting the instructed solution strategies and to what success, as in Studies 1 and 2. To do so, we used a latent transition analysis to capture how students engage with and apply the instructed procedures and how this changes over time.

1.1 Individual Differences in Pressure and Anxiety

For many students, there is a high desire to perform their best in the fear of the consequences that may follow such as poor grades, reduced educational or employment outcomes, disapproval from parents or teachers or such related negative outcomes. Most elementary and middle school students report regularly experiencing high levels of stress and pressure in the classroom (Bauwens & Hourcade, 1992). In the domain of mathematics, this pressure and anxiety can lead to individuals taking fewer math courses, earning lower grades in the courses they do take, and performing more poorly compared to their equal counterparts who do not have math anxiety (see Hembree, 1990). Studying the effects of these negative relations to mathematics, including formalized mathematics anxiety, which is defined as fear or apprehension about doing math-related tasks (Richardson & Suinn, 1972) is important for math learning and, more broadly, engagement with STEM.

A large body of research has examined students' performance in high pressure testing situations to examine anxiety on mathematical cognition. For instance, Ashcraft and Kirk (2001) found that students' mathematics anxiety undermines performance but those same students in non-anxiety-provoking contexts are highly competent. Stereotype threat, that is, the heightened awareness of a negative stereotype about a social group in a specific task has been consistently found to result in poorer performance by members of that social group when the stereotype is activated (Steele, 1997). Students who have math anxiety or are worried about confirming negative racial or gender stereotypes about their mathematics ability may be especially susceptible to the negative impacts of performance pressure (Maloney et al., 2013). Lastly, Beilock and Carr (2005) found that under high pressure and high-stakes contexts in mathematics, students can *choke,* or perform more poorly than expected than one's skill level and misrepresent students' knowledge or capabilities (see Beilock, 2010). Too much pressure can lead to intrusive thoughts and worries which can tax cognitive resources and limit engagement during math exercises (Maloney et al., 2014; Lyons et al., 2018). While in many cases, the pressure can be harmful, the valence of the stress and arousal plays a key role into whether the pressure can be supportive (see Yerkes & Dodson, 1908; Sapolsky, 2015). As an example, when the stakes are too low, students may lack engagement, effort, and motivation (Sullivan et al., 2006; Martin et al., 2012; Hancock, 2001).

The mechanisms by which stressors from pressure can harm performance include both *trait* and *state* anxiety. Trait math anxiety (e.g., math anxiety or testing anxiety; Beilock et al., 2004; Hancock, 2001) is characterized by stable differences between individuals in their baseline levels of anxiety, whereas state anxiety (e.g., performance pressure; Baumeister, 1984) is a temporary arousal due to threatening, context specific stimulus. The influence of trait and state anxiety depend, in part, to the mathematics context. For domains that are easy or highly-practiced there is no effect of pressure (Beilock et al., 2004; Beilock & DeCaro, 2007). In contexts that are cognitively demanding and require one's working memory capacity, performance suffers when students are more anxious (Trezise & Reeve, 2018).

Almost all research into the effects of anxiety or pressure on math performance have focused on math problem-solving or very basic core number tasks (e.g. Ashcraft & Kirk, 2001; Maloney et al, Ansari & Fugelsang, 2011, Trezise & Reeve, 2018). Research which examines how students learn novel solution strategies and fundamentally change the way they think about and approach problem-solving remains unknown. Additionally, recent evidence suggests that math anxious people have less reflective thinking (Maloney & Retanal, 2020). Reasoning about relationships (higher order thinking) while enables learners to transfer ideas and think flexibly, however, it is more cognitively effortful (see Gick & Holyoak, 1983; Richland & Simms, 2015; Rittle-Johnson & Star, 2007, 2011). This leaves a gap in our understanding if pressure differentially affects procedural gains in problem-solving compared to more conceptual and comparative thinking such as is required for higher order thinking. Moreover, there are no studies to date that assess pressure on a large scale in an ecological valid way—that is in children's learning environments during cognitively complex and effortful contexts such as in the domain of proportional reasoning.

1.3 Individual Differences in Cognitive Resources

The main mechanism by which pressure and anxiety affect performance is through the individual's cognitive capacity or executive functioning system. The executive functioning system is a set of cognitive resources including cognitive capacity, working memory, and inhibitory or attentional control (Diamond, 2013). When a learner is anxious or under pressure, one's attention is split between the pressure and the task. Under high pressure or high anxiety contexts, learners cognitive resource system is taxed (e.g., Ashcraft & Kirk, 2001; Maloney et al., 2013) as they tend to divert more attention to the threat-related ideas than to task-related ideas (Eysenck et al., 2007). The available resources are allocated in verbal worry and emotion regulation efforts, leaving less resources available for problem-solving in cognitively complex tasks, lowering performance and efficiency (e.g., Ashcroft & Kirk, 2001; Beilock & Carr, 2005; Beilock & DeCaro, 2007). Research has shown that high working memory students, under pressure, are most vulnerable to state (Beilock, 2008; Beilock & Carr, 2005) and trait (Ramirez et al., 2013; Ramirez et al., 2016) related decreases in performance. This is because high working memory students often use more complex strategies to solve problems so when pressure is applied, these students are not able to use these sophisticated strategies and their performance declines (Beilock & DeCaro, 2007). Lower working memory students, however, use the same less sophisticated strategies whether or not they are under pressure.

Individual differences are also susceptible to the *learning context* and mathematics domain. In reasoning about cognitively complex mathematics domains such as proportional reasoning, a students' cognitive capacity is key factor to consider. For example, while some students spontaneously draw connections and engage relationally when learning new material, which is required during proportional reasoning, many students do not.

There are a few key reasons that some students may benefit from reasoning opportunities more than others. One explanation may be due to individual differences in reasoning skill. Reasoning requires the ability to represent information as a system of relationships (e.g., between symbols in an equation, from the word problem to the equation), manipulate that information to draw higher-order relationships between phenomena (e.g., similarities or differences between multiple problem types or solution strategies), and draw connections and relations between these representations (see Alfieri et al., 2013; Polya, 1957; Richland et al., 2012; Ross, 1989; Ross & Kennedy, 1990; Reed, 1987, 1989). This is both complex and cognitively effortful. Another factor is that engaging in these conceptually rich learning environments is cognitively effortful. It requires that learners hold complex representations in mind, manipulate those representations to find higher order relationships, and execute many multi-step solution strategies. Therefore, a student with low cognitive resources may fail to effectively hold in mind the multiple relationships, ignoring the higher order connections, and rather simply memorize procedures (Begolli & Richland, 20215). Moreover, students who are feeling pressure to perform may be disproportionality challenged to engage successfully in these learning practices due to overloaded cognitive resources (e.g., Meyer, 2010; Yeung et al., 1997).

Finally, reasoners domain-general and domain-specific skills may also lead some students to easily notice higher order relations more consistently than others (Holyoak, 2012). Domain-general reasoning skills are known to vary among individuals but maintain reliability over time (Raven, Raven & Court, 2004). Domain-specific skills, such as prior mathematics knowledge, has also been a key factor to determine if students are prepared to benefit from opportunities to reason from new content (Rittle-Johnson et al., 2012). Therefore, proportional reasoning makes for an interesting case to investigate the relationship between individual differences and learning contexts.

1.4 Teaching Practices to Promote Learning in Conceptually Demanding Contexts

Proportional reasoning, the ability to understand the multiplicative relationship between quantities and phenomena as systems of structured relationships that can be aligned, compared, and mapped together (Cramer & Post, 1993; Boyer et al., 2008; Richland & Simms, 2015) is a required skill in mathematics education as it is foundational for higher level mathematics and algebra, as well as for navigating everyday situations (e.g., National Mathematics Advisory Panel, 2008; Siegler et al., 2013). Many students however experience pervasive difficulties with proportional reasoning (National Research Council, 2004) in part due to its conceptual difficulty and the demands it places on the executive functioning system.

Developing classroom lessons in which learners successfully reason relationally is complex due to the need to remember, compare and manipulate multiple relationships at one time (National Research Council, 2004). When comparing multiple solutions, students often fail to connect the conceptual similarities across the structure mapping unless provided with explicit support (see Alfieri et al., 2013; Gick & Holyoak, 1980, 1983; Gentner et al., 2003). Instruction that engages students in more comparative thinking and provides perceptual cues such as gesture to explicitly draw the relationship has been shown to support learning (Alfieri et al., 2013; Richland & Begolli, 2018).

To support this conceptually complex reasoning, expert teachers often use comparison as an effective learning process for mathematics (Common Core Standards in Mathematics, 2010; Richland et al., 2007). In mathematics, comparisons can be between multiple representations or solution strategies (e.g., Richland & McDonough, 2010; Rittle-Johnson & Star, 2007) or between multiple exemplars of the same solution strategy (Kellman et al., 2010; Richland & McDonough, 2010). These comparisons help convey new, complex information by drawing parallels to concrete examples (see Kilpatrick et al., 2001; Alfieri et al., 2013; Alibali et al., 2014; Richland & McDonough, 2010; Rittle-Johnson & Star, 2007, 2009; Star & Rittle-Johnson, 2008). When comparing problem types within strategies, students develop stronger procedural skills. When comparison across solution strategies students gain a broader and more sophisticated schematic for solving problem types in many ways (Watson & Mason, 2006).

Another reason for students' pervasive difficulties in proportional reasoning is that reasoning taxes the cognitive load system. A growing literature suggests that making worked examples visible during learning opportunities can reduce cognitive load and support problem-solving (see Sweller et al., 1998). When comparing solutions, making compared solutions visible, and leaving them visible throughout the comparison is one mode for promoting deep, conceptual reasoning (Nokes-Malach et al., 2013; Richland & McDonough, 2010). Benefits of making these representations visible have been shown in experiments across ages, including from young children (Christie & Gentner, 2010) to later elementary students (Rittle-Johnson & Star, 2007). Matlen and colleagues (2011) for example, found that elementary-aged students were more likely to learn and retain geoscience concepts when text passages describing the concepts were accompanied by visual representations of both the source and target as opposed to just showing the target. This simultaneous presentation prompted students to compare the two representation and reduced the cognitive burden of having to recall each representation. Similar findings in the domain of mathematics have shown learning gains when source and target representations are displayed simultaneously versus sequentially, leading to gains in procedural knowledge, flexibility, and conceptual understanding (e.g., Begolli & Richland, 2015; Christie & Gentner, 2010; Rittle-Johnson & Star, 2007).

Most research on pedagogical supports have only investigated worked examples *after* instruction, during practice or testing; there has been little attention to how these pedagogical supports can be implemented during *learning*. Moreover, no work to our knowledge has examined how these supports could potentially reduce the negative effects of imposed pressure.

1.5 Current Study

This project is a novel integration of literatures on mathematics education instruction, relational reasoning, mathematics anxiety and pressure in order to explore new ways that cognitive processing resources may be related to achievement gaps associated with U.S. mathematics instruction. The research focuses on youth in low SES, ethnically diverse schools at a crucial transition point, between upper elementary and middle school, at which academic motivation tends to reduce dramatically (e.g., Eccles & Wigfield, 2002) and where US student performance starts to decrease relative to other nations.

Pressure and its associated anxiety is well established to impair *test* performance (Beilock, 2010). In this study we examine the role of pressure and its corresponding anxiety during a *learning* context in a particularly challenging and cognitively taxing domain of proportional reasoning. We also examined if reducing cognitive load via instruction with high visual support can improve learning outcomes. We explored these questions in two learning contexts: 1) higher order thinking of multiple solution strategies during the lesson, 2) performance accuracy on the post-lesson test and follow-up test, 3) the nature and sequence of students' problem-solving approaches on the post-lesson test and follow-up test, and 4) if students' higher order thinking mediates the relationship between condition and performance gains in accuracy and problem-solving. In order to capture the nature and sequence of students' problem-solving, students qualitative written work was coded to determine if and to what extent students' attempted the novel strategy solution and mapped strategies and procedures in near-transfer contexts; latent transition analysis was used to identify and characterize the full sequence of steps students took to solve proportional reasoning problems.

To address these research questions, we used an experimental methodology that uses measurements designed to optimize experimental rigor and ecological validity. We used stimuli that mirror a true classroom experience – a mathematics video-lesson recorded in a real classroom with real students providing their solution strategies. The high quality instruction on proportional reasoning was video-edited to support or strain cognitive resource through variations in the visual support during instruction. The audio stream of the students and teacher during the lesson was kept consistent across both videos. In a 2 (pressure vs no pressure) x 2 (high visual support vs low visual support) we explore how pressure and visual support impact students' engagement with the lesson and overall learning.

2. Methods

2.1 Participants

Fifth- and sixth-grade students from seven K-8 schools (17 classrooms) in the Chicagoland area with primarily underrepresented Latinx or Black youth in low to middle income neighborhood schools were recruited to participate in this study. They were chosen because students at this age have been introduced to the math concepts necessary to solve proportional reasoning problems; however, they have not yet been formally introduced to solution strategies for solving proportional reasoning word problem (Common Core State Standards in Mathematics, 2010). Students who were absent on the day in which the intervention was administered were excluded from analyses (n=37), yielding a sample of 441 students (330 5th grade, 111 6th grade; 190 females; 45 unidentified). Parents and guardians were informed of the study a few weeks prior to data collection and were provided the opportunity to opt their child out. We also obtained children's written assent prior to data collection.

2.2 Teacher and Experimenter Demographics

All of the children's regular mathematics teachers were present in the classroom throughout the study and at least two experimenters were in the classroom during the study.

2.3 Design and Procedures

This is a multisite randomized study; procedures were administered during three visits to each classroom over a two-week period. Students completed all procedures in their classroom, alongside their peers in their everyday mathematics class. This is a 2 (pressure: no and imposed pressure between subject conditions) x 2 (visual support: low and high between subject conditions) x 3 (phase: pretest, post-lesson test, and follow-up test repeated measures) mixed design approach. Counterbalancing was performed for condition assignment as well as for the mathematics packets students completed on the pretest, post-lesson test, and follow-up test.

Session 1 (Day 1). Students completed a pretest assessing their initial understanding of proportional reasoning and arithmetic skills, material to be covered in the lesson. Students also completed a measure of selective and sustained attentional control, the d2 Test of Attention (Brickenkamp & Zillmer, 1998; Rhonda & Ross, 2005).

Session 2 (Day 2). Two to three days later, students watched a previously-recorded, conceptually challenging mathematics lesson on proportional reasoning which was administered individually on iPads. During the video lesson, students simultaneously completed a problem-solving packet. The two manipulations of the study were administered during the lesson via the instructional video and in the problem-solving packet. Pressure was administered before the lesson through the video instruction and in the problem-solving packet (more details below). Visual support was administered during the lesson through the video instruction and in the problem-solving packet (more details below). All students completed a post-lesson test immediately following the lesson. Finally, students completed a Situational Interest Scale (Chen et al., 1999). There was a pressure debrief at the end of this day to remove the imposed pressure.

Session 3 (Day 3). One week after the lesson, students completed a follow-up posttest on proportional reasoning, a working memory Symmetry span measure (Unsworth, Heitz, Schrock & Engle, 2005), and a mathematics anxiety scale, the MARS-E (Suinn, Taylor, & Edwards, 1988).

2.4 Math Lesson

Researchers worked with a teacher and curriculum designer to create a lesson script introducing ratio and proportion. The 45-minute teacher-led lesson was recorded as a live, semi-scripted lesson on proportional reasoning. The teacher taught a diverse class of fifth- and sixth-grade students who were recruited for the recording of the lesson. Recording a live lesson with real students allows for the natural variability of classroom instruction with students having real world conversations regarding the solution strategies. This experimental procedure enables the instructional stimuli to have high ecological validity while maintaining experimental control.

The lesson is designed based on a reform-based instructional model in which a teacher leads mathematical discussions where students compare and contrast solution strategies to a single problem (e.g., Carpenter et al, 1999; Kazemi & Hinz, 2014; Smith et al, 2009, Stigler & Hiebert, 1999). First, the teacher asks the students to solve a challenging proportional reasoning problem on their own prior to receiving explicit instruction (DeCaro & Rittle-Johnson, 2012; Schwartz et al., 2011; Schwartz & Martin, 2004). The teacher walked around the room and chose a student who used the equivalent fraction strategy; she asked the student to the board to describe the solution strategy to the class (e.g., Carpenter et al., 1999; Kazemi & Hinz, 2014; Smith et al., 2009; Stigler & Hiebert, 1999). Following, the teacher then led a discussion on the procedures and provided a higher-level conceptual overview of the equivalent fraction strategy. Students solved a near-transfer problem using the equivalent fraction strategy. We employed a multiple solution strategy approach in this lesson. Next, the teacher brought a second student up to the board to describe solving the same problem but this time using the unit ratio strategy. Following, the teacher led a discussion on the procedures and provided a higher-level conceptual overview of the unit-ratio strategy. Students solved a near-transfer problem using the unit ratio strategy. Then, the teacher summarized both solution strategies for both the instructed and near-transfer problems.

The math lesson was videotaped using multiple cameras capturing different angles of the teacher, the white boards, and the students. These angles allowed for manipulation of the raw video footage. To develop our study manipulations, videotapes were edited to produce the instructional stimuli.

2.5 Imposed Pressure Manipulation

Students were assigned to one of four experimental conditions administered during Day 2 of the study via the video lesson and the problem-solving packet: i) no pressure, low visual support (control), ii) no pressure, high visual support, iii) imposed pressure, low visual support, and iv) imposed pressure, high visual support.

Pressure was manipulated at the very beginning of the proportional reasoning lesson through a combination of false negative feedback, social-evaluative pressure (Beilock & Carr, 2001), and expressive free writing targeted to heighten the level and awareness of the pressure. Students in the high pressure condition first received the false negative feedback: the student is instructed (on the iPad) to turn the page in their packet to reveal their pretest score. Students saw their [false] pretest score of ~30% (range 25 – 35%), handwritten in their packet. Next, students received the social-evaluative pressure. The students were told they would be completing a video about ratios and instructed to do as well as possible as their performance from the lesson could determine whether the class would have a pizza party. Finally, for the expressive free writing, children were prompted to reflect and write their thoughts regarding the mathematics lesson and test. Students received the instruction: "It's very important to us that we know how kids are feeling about this math lesson". Students were then prompted to write about their feelings and thoughts, from a list of negative mathematics and performance related topics (e.g., "what the test will be like", "if you will learn enough to get a good grade", "what the class will think of you if you lose the pizza party for them", "a time in the past when you got a bad grade in math", "what your parents and teachers think if you don't do well in math"). At the end of Day 2, a pressure debrief was provided to the students. Researchers told the students that the scores on their math packet were made up, the score on this test would not affect their grades in anyway, and that everyone would be receiving a pizza party. This was designed so that students would not feel pressure on

Day 3 for the follow-up test. The goal here was to manipulate the learning context on Day 2 but not influence test pressure on Day 3.

Students in the low pressure condition were told to read a summary of the aims of the study in their packet. Next, students were told they would be completing a video about ratios and informed that they would complete some questions at the end of the lesson based on what they had learned. Finally, students received the instruction: "It's very important to us that we know how kids feel about school" and were asked to identify their favorite school subject from a list and then to write down all the things you like about that subject.

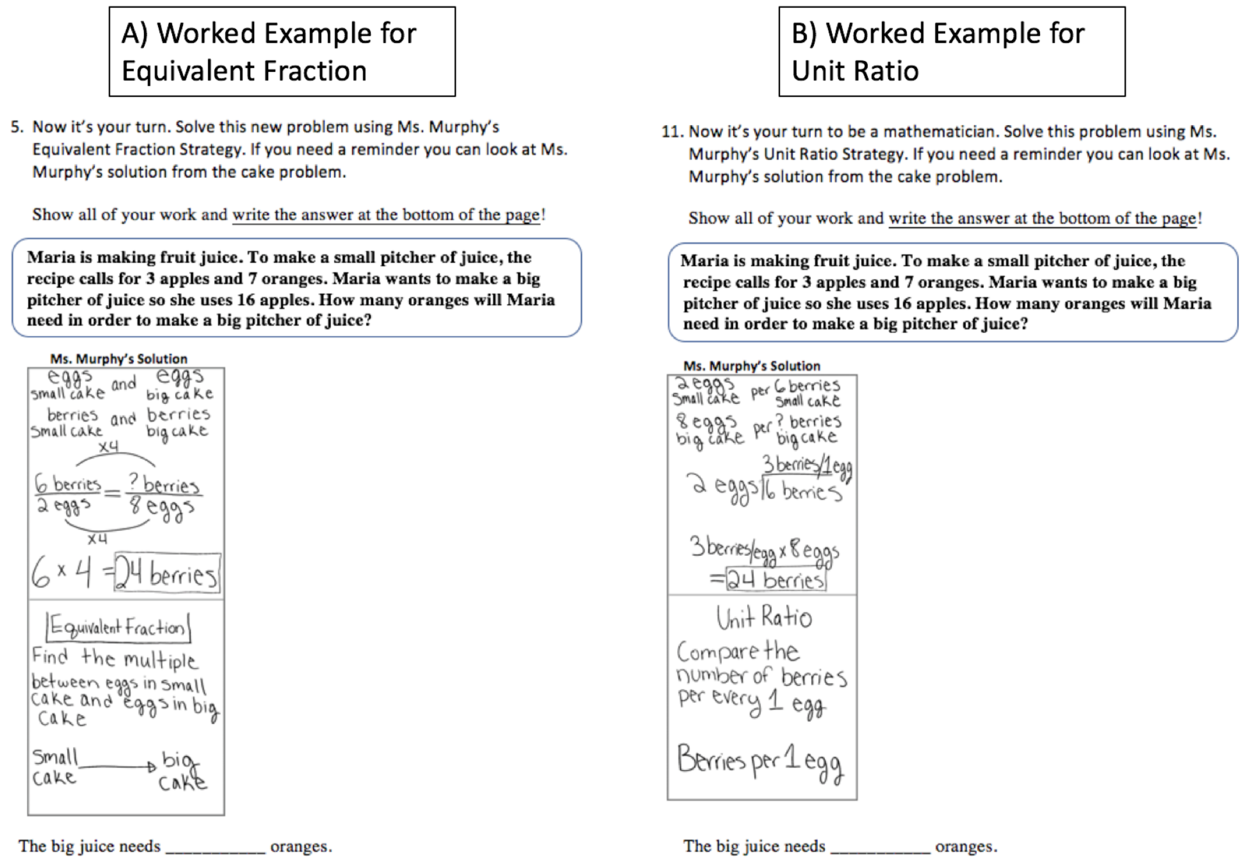
2.6 Visual Support Manipulation

Visual support was manipulated as a malleable factor as a means to reduce burden on students' cognitive load via two pedagogical practices: visual support during video instruction and worked examples in the mathematics packet during the lesson. The visual support in the recorded lesson included simultaneous and spatially aligned presentation of solution strategies; the teacher also used liking gestures when comparing across problem types or solution strategies (Begolli & Richland, 2016). Worked examples were provided in the problem-solving packet during the lesson when students were asked to solve problems or compare solution strategies. For example, when solving proportional reasoning word problems, there was a worked example of the instructed solution strategy from the video and it was spatially aligned to support mapping of procedures when students solved near-transfer problems (see Figure 4.1 for worked example during problem-solving using the equivalent fraction (A) and unit ratio (B) strategy)). In addition to the fully worked out procedures of the solution strategies, the worked examples outlined the key comparisons to be made for the equivalent fraction strategy (i.e., "Find the multiple between eggs in the small cake and eggs in the big cake") and the unit ratio strategy (i.e., "Compare the number of berries per every 1 egg"). When students were asked to compare problem types or solution strategies in written or multiple choice questions, worked examples of both solution strategies were provided.

The low visual support condition included teacher instruction with sequential presentation of solution strategies and no linking gestures in the video. Additionally, these students did not have worked examples while problem-solving or comparing solution strategies which required students to recall the solution strategy approach and procedures. The audio stream of the students and teacher during the video lesson was kept consistent across both conditions; all students heard the same instruction.

Figure 4.1

High Visual Support Condition: Worked Example During Near Transfer Problem-Solving



Note. Worked examples were presented for students in the worked example condition on the near transfer problem. Students were asked to solve the problem using the equivalent fraction strategy (A) and unit ratio strategy (B). Students were provided with the strategy approach, setup, procedures, solution. Additionally, students were provided with the key conceptual framework for solving the problem.

2.7 Problem-solving Packet During Lesson

Consistent with the video lesson and reform-based instructional models, students in the study solved a challenging proportional reasoning problem on their own before any instruction (e.g., Carpenter et al., 1999; Kazemi & Hinz, 2014; Smith et al., 2009). Students interacted with the lesson as they would normally in a class by solving problems interleaved with teacher instruction (Begolli & Richland, 2016).

Following, the teacher led a discussion on the procedures and higher-level conceptual overview of the equivalent fraction strategy, students were instructed to solve a near-transfer problem using this strategy. Then, students were asked conceptual and higher-order thinking questions regarding the equivalent fraction strategy. For example, they were asked to compare the similarities and differences of solving the two problems using the equivalent fraction strategy and to identify the key relation between the units in the problem. Engagement with higher-order thinking (HOT) questions during the lesson were coded as an outcome measure and a potential mediator to learning during the lesson, see below.

The identical procedure occurred for the unit ratio strategy; the teacher led a discussion on the procedures and higher-level conceptual overview of the unit ratio strategy, students completed the same near-transfer problem but this time using the unit ratio strategy. The same free-response and multiple choice questions targeted to engage students higher-order thinking were presented following instruction of the unit ratio strategy. Finally, students were asked higher order thinking questions that engaged reasoning across solution strategies (e.g., why particular strategies would be best for specific problems).

2.8 Higher-order thinking questions during the instructional intervention

Higher order thinking (HOT) during the lesson was measured by students' ability to draw higher order relationships between phenomena (e.g., similarities or differences between multiple problem types or solution strategies), and draw connections between these representations, invoking relational and analogical reasoning processes (see Alfieri, Nokes-Malacha, & Schunn, 2013; Polya, 1957; Richland, Stigler & Holyoak, 2012; Ross, 1989 Ross & Kennedy, 1990). The videotape is paused after each approximately five-minute instructional phase and participants were asked to respond to questions that assess whether they had been reasoning relationally while watching the previous section, including being asked to critique, compare, contrast, or draw inferences about discussed solution strategies.

Questions to probe higher order thinking occurred during three district parts in the lesson: higher order thinking for the equivalent fraction strategy, higher order thinking for the unit ratio strategy, and higher order thinking for comparing across strategies. We used free response questions and multiple choice questions to measure students' engagement with higher order thinking for these three categories. Free response questions were coded for language that articulates comparison, connections across contexts, inferences, hierarchical categorizations, or abstractions (see Lewis & Smith, 1993; Richland & Simms, 2015). Higher order thinking was examined as an outcome measure to examine lesson performance. Higher order thinking was also examined as a mediator to problem-solving learning.

2.9 Assessments

A pretest, immediate posttest, and delayed posttest was administered to assess learning. Each test included 6 missing value proportional reasoning word problems, for example:

> A smoothie recipe calls for 2 bananas and 5 oranges. To make a bigger pitcher of the smoothie with 12 bananas, how many oranges should you add?.

Students were instructed to solve to find the missing value. Students were given instructions for the strategy to use for each problem: problem 1, any strategy of choice; problem 2, equivalent fraction; problem 3, unit ratio; problem 4-6, any strategy used in problems 1-3.

As outlined above, there were 3 alternative versions of the packets. The problems for each of the packets were matched by using clear defined rules to ensure equivalence between packets. Students were instructed to use the two solution strategies when solving the six test problems. For problem 1: all numbers in the problem were between 3 and 12 and the solution was a double digit number under 30. If the key link was calculated correctly, it was a whole number for both solution strategies. For problem 2: all numbers in the problem were between 3 and 12 and the solution was a double digit number under 30; if students correctly used an equivalent fraction strategy, any calculated numbers were whole numbers. If the unit ratio strategy was used, solution strategy required multiplication/division of non-whole numbers. For problems 3 and 4: one number in the problem was a single digit number above 5, two numbers were double digit numbers up to 20, the problem could be solved using long division and multiplications but using fractions would be extremely difficult, solutions were non-whole numbers. For problem 5, there were two problems embedded within it. Two of the numbers in the problem were between 4 and 12, the small recipe number was always the value 10 and a large recipe number was always the value 22; if equivalent fraction or unit ratio were used the calculation at step 1 would result in a number *X*.5 (e.g 3.5) and all solutions were whole numbers.

2.10 Proportional Reasoning Problem-solving Coding Scheme

In order to assess learning gains, students' overall accuracy on proportional reasoning word problems was examined. We were also interested to examine students' ability to apply the strategies and employ the procedures taught in the lesson on new problems. Therefore, we explored the nature and sequence of students' problem-solving decisions, from their initial attempt of the solution strategy to the setup and procedures and ultimately their final solution accuracy for the transfer problems. For the proportional reasoning word problems, students written, qualitative work was examined to determine if and how students followed the instructed procedures. The complete coding protocol is available on OSF and outlines the development of the codes, coding protocols and reliability checking. A brief outline of the coding protocol for the strategy solutions is outlined next.

Students were instructed to choose and solve the best strategy (i.e., equivalent fraction or unit ratio strategy) for all six problems. To examine students' use of the equivalent fraction and unit ratio strategy, a combined coding scheme was created to assess key problem-solving procedures: 1) proportional approach, 2) instructed attempt (equivalent fraction or unit ratio strategy), 3) numerical setup, 4) procedure, 5) linked quotient from the first step to the second step, and 6) accuracy. A summary of the coding scheme for equivalent fraction word problems can be found in Table 4.1. These scores were aggregated across problems (0-6) for each step and time point. For example, there was a score between 0 and 6 for pretest attempt of solution strategy.

To establish inter-coder reliability, 20% of the data was coded by two research assistances. Krippendorff's alpha was used to assess inter-coder reliability for each variable coded (Hayes & Krippendorf, 2007). For the coding of the variable to be considered reliable it was required that the Krippendorff's alpha (an index that accounts for level of measurement and agreement expected by chance and is known to be conservative to be 0.70 or higher. Reliability for variables in the problem-solving variables on pretest were $\alpha = 0.82$, on the post-lesson test on day 2 problem-solving variables were $\alpha = 0.89$, and delayed, follow-up test on day 3 problem-solving variables were $\alpha = 0.90$. Prior knowledge arithmetic variables were $\alpha = 0.95$.

Table 4.1
Coding Scheme for Problem-Solving: Equivalent Fraction and Unit Ratio Strategies Combined

Problem-solving Step	**Equivalent Fraction Setup**	**Unit Ratio Setup**
Approach	Approached the problem using a proportional thinking strategy (e.g., equivalent fraction, unit ratio, build up, multiplicative, other)	
Attempt	Attempted equivalent fraction strategy (i.e., fraction or ratio) with numbers from problem and attempted to find relationship between numbers in problem	Attempted to find the unit ratio and attempted to apply that number (i.e., division and multiplication)
Numerical Setup	Set up creating proper or improper fractions within recipes	Set up comparing different ingredients within a recipe.
Procedure	Compared unit of small recipe to unit of large recipe using multiplicative reasoning	Completed two steps of procedure: attempted division first, to find a unit ratio and then multiplication to apply to large recipe
Link	Solution calculated from the first operation used in the second operation	
Accuracy	Accurate solution	

Note. This combined coding scheme across solution strategies was applied to all problem-solving questions across all three testing points (i.e., pretest, immediate posttest, delayed posttest). Problem-solving steps are used as the key indicators for the LTA analysis.

2.11 Individual Difference Measurements

A set of measures assessing individual differences in executive functioning, math anxiety, and achievement orientation were administered to all participants.

2.11.1 Arithmetic Assessment

Arithmetic skills that are needed to solve proportional reasoning word problems using an equivalent fraction strategy and unit ratio strategy were assessed. Students solved five missing-value ratio equivalence problems to assess their understanding for multiplicative relationships

(Cramer & Post, 1993). The missing value in the ratios were the back term numerator and back term denominator. Potential solution strategies for comparison also varied: numerator to numerator (e.g., $\frac{5}{6} = \frac{15}{\ }$), denominator to denominator (e.g., $\frac{2}{3} = \frac{\ }{12}$) or front term numerator to front term denominator (e.g., $\frac{2}{12} = \frac{3}{\ }$). Students also solved two long division problems and three decimal multiplication problems.

2.11.2 Attentional Control

The d2 Test of Attention is a group-administered, pen and paper measure which assesses students sustained levels of selective attention and attentional control (Brickenkamp & Zillmer, 1998; Rhonda & Ross, 2005). The assessment is normed with US and German children, adolescents and adults. Under a time pressure, participants are asked to search and cross off target characters (i.e., "d"s with two dashes surrounding it either above or below) from perceptually similar distractors (e.g., "d"s with one dash, "p"s with two dashes). The outcome of this task is the total number of items processed minus errors (TN-E). Across the literature, internal consistency is high ($\alpha \geq 0.8$) and test test-retest reliability is high ($\alpha > 0.8$; Clark, 2005). d2 TN-E scores correlate with other measures of attention and executive functioning: Stroop and Tower of London which support the validity of using this measure (Clark, 2005).

2.11.3 Working Memory (WM)

we assessed students' baseline working memory capacity with the Shortened Symmetry Span Task (Foster et al., 2015) formatted for play on a 5th generation Apple iPad with a 9.7-inch active screen. Instructions were presented on the screen and narrated through headphones. Participants were required to hold target information (highlighted cells in a matrix) in memory while attending to distracting information (whether or not an unrelated matrix is symmetrical across its vertical axis). Participants completed four test trials which randomly varied in length from two to five cells to be memorized, for a total possible score of 14 cells correctly memorized in order.

2.11.4 Trait (Habitual) Mathematics Anxiety

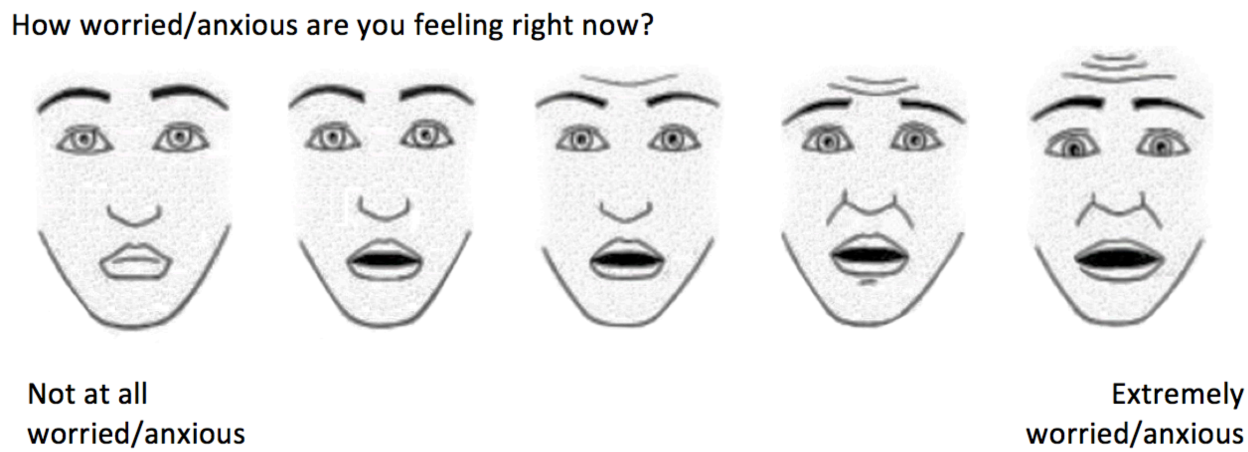
Modified Abbreviated Math Anxiety Scale (mAMS, Carey et al., 2017) is a 9-item self-report mathematics anxiety questionnaire for children between 8 and 13 years. The mAMAS is modified from the adult AMAS (Hopko et al., 2003). Participants use a 5-point Likert scale to indicate how anxious they would feel in different math situations. The scale has good reliability (Ordinal alpha = .89, Cronbach alpha = .85, Carey et al., 2017), that is equally high across age-groups.

2.11.5 State Anxiety

State anxiety affects cognitive processing because it increases worry (Eysenck et al., 2007). Students' state anxiety was captured on every page in their mathematics packet to assess changes in state anxiety over the lesson (see Figure 4.2). Capturing anxiety with this specificity allows us to examine rates of state anxiety immediately after the imposed pressure manipulation was administered, during word problem-solving and high order thinking questions, and finally we can take an average across the mathematics packet to gather overall state anxiety levels during the instruction.

Figure 4.2

State Anxiety Measures Administered During the Study



Note. Students were to choose a state level feeling (0 not worried/anxious at all to 4 extremely worried/anxious) after completing each page of mathematics packets.

2.11.6 Situational Interest Scale

The scale consists of 24 five-point Likert type items (5 = strongly agree, 1 = strongly disagree) to measure levels of students' feelings about the activity in terms of Novelty, Challenge, Attention Demand, Exploration Intention, and Instant Enjoyment dimensions (Chen et al., 1999). This scale was used as an outcome measure and as a potential mediator to problem-solving gains. Findings of students' situational interest are reported in the Supplementary Materials S 2.2.

2.12 Analytical Approach

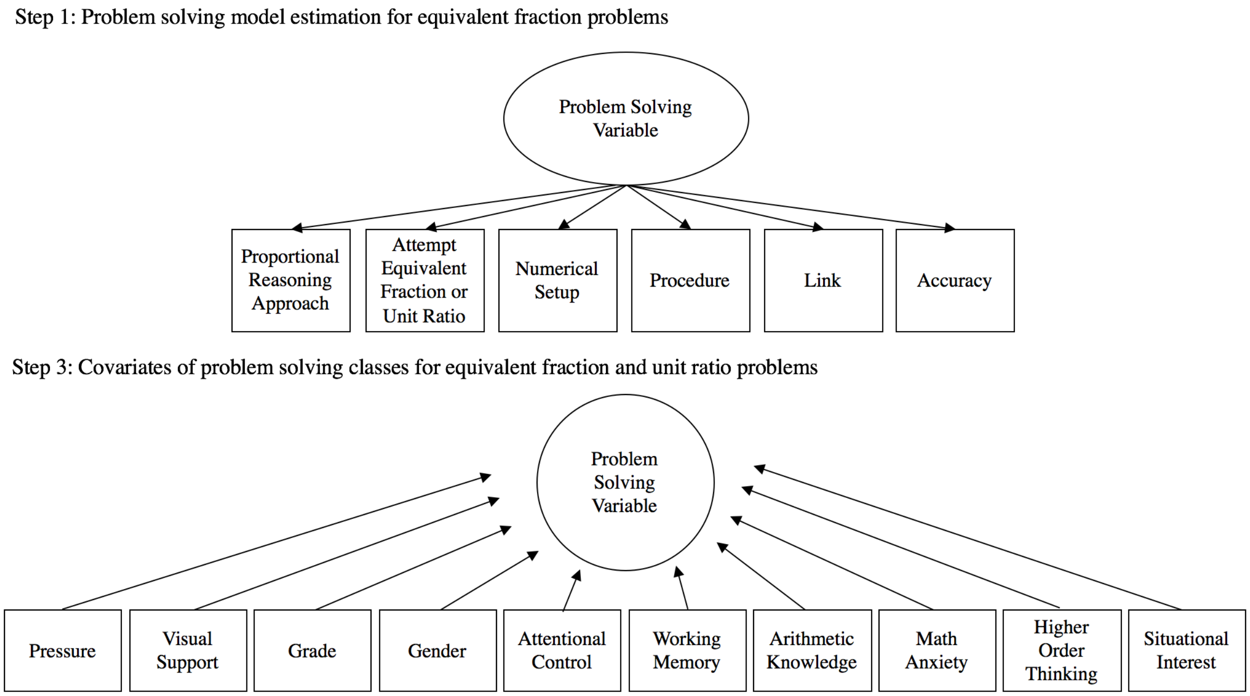
In this study, there are two experimental manipulations: pressure and visual support. To build on the findings from Study 2, first we examined the effects of visual support in the absence of pressure. Then we examined whether visual support is a tool to reduce the extraneous cognitive load from the imposed pressure. Pressure and visual support are dummy indicators.

To examine these interventions during the cognitively taxing and conceptually challenging lesson on proportional reasoning, we explored different types and levels of learning and engagement during the lesson and on posttests (i.e., immediately following the lesson and a week's delay). We ran a series of ANOVAs and regression analyses to examine the impact of condition on students' 1) higher order thinking during the lesson, 2) performance accuracy gains on the post-lesson test, and 3) performance accuracy gains on the follow-up test. We also examined nature and sequence of students' problem-solving gains (i.e., from approach to setup to procedures to final solution accuracy) on both the 4) post-lesson test and 5) follow-up test.

To capture the nature and sequence of students' proportional reasoning problem-solving, we coded students' qualitative problem-solving. To analyze this multistep data across three time points, we used a novel analytic approach: a three-step latent transition analysis (LTA) in Latent GOLD 6.0 (Statistical Innovations, 2021). LTA is a longitudinal version of latent class analysis; LTA allows us to examine the multiple steps students take in order to solve the proportional reasoning problems and assess these steps across time points. The multiple problem-solving steps are indicator variables used in the LTA (i.e., approach, attempt of instructed strategy, setup, procedure, link, accuracy; see Figure 4.4, Step 1). The assessment measures included 6 problems to solve therefore, the count variables for each indicator will range from 0-6.

There are three key steps in LTA. First, an optimal number of latent groups are estimated using these indicators. This first step determines how many latent groups best represent the data, these groups represent students' problem-solving approaches and successes. Our study is longitudinal with three time points; the same latent groups exist at all time points but the participants can move between groups at each time point (e.g., a student may belong to group A at time point 1 and move to group B at time point 2). Second, the optimal model is identified and individuals are classified into the groups at each time point. The transition between groups across time is also characterized and reflects movement from one latent group to another over time. Third, the relationship between the groups and external variables are examined correcting for the classification error to prevent bias (Figure 4.2; Bolck, Croon and Hagenaars, 2004; Vermunt, 2010; Bakk, Tekle and Bermunt, 2013). These models assume unobserved heterogeneity reflected in the latent grouping variable. LTA is a hidden Markov model, meaning that group membership at time *t* is condition on group membership at the immediately previous time point (*t* -1).

Figure 4.3
Conceptual Diagram and the Analytic Model



Note. Step 1 Model estimation: Latent Class Analysis (LCA) identifying the optimal proportional reasoning problem-solving model for equivalent fraction problems and unit ratio problems. Order indicates the order of proportional reasoning problem-solving task (attempt instructed strategy, setup, link, accuracy). Step 3: Examination of the relationship between problem-solving groups and covariates.

Note that the data set contains missing values. We addressed these omissions via multiple imputation the method of predictive mean matching, implemented using the *mice* library in R (van Buuren & Groothuis-Oudshoorn, 2011). See Section S2 in Supplemental materials for details of the imputation procedure and Tables for alternative analyses that used only complete cases.

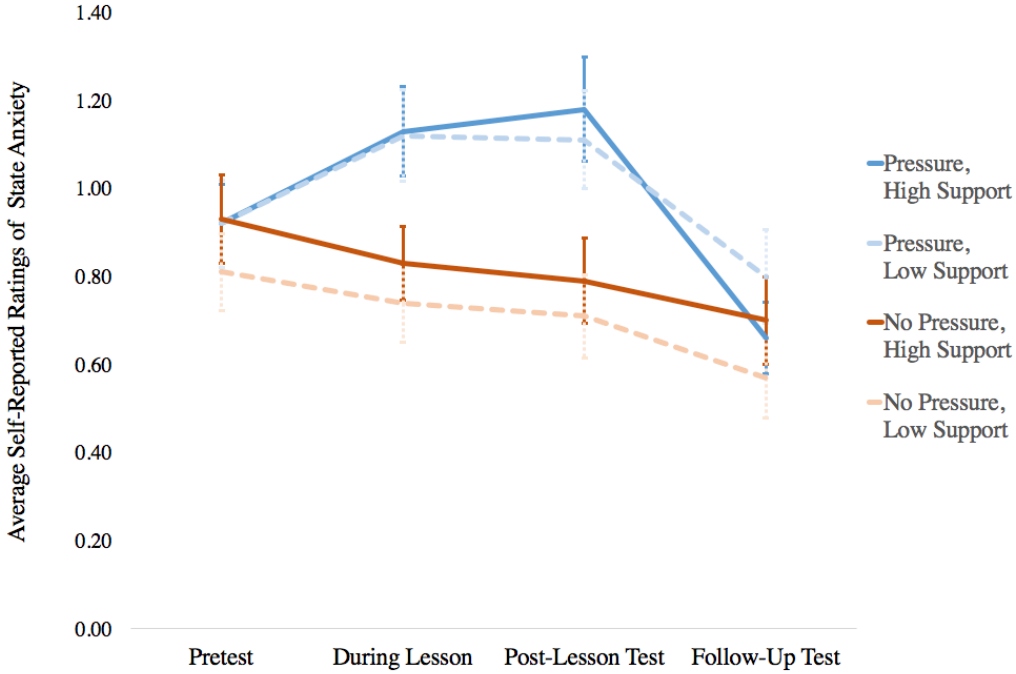
3. Results

3.1 Efficacy of the Imposed Pressure Manipulation During Learning

To evaluate the effectiveness of the imposed pressure manipulation, we examined if the pressure significantly impacted students state anxiety levels during the lesson. Students' self-reported feelings of state anxiety at pretest, during the lesson, post-lesson test, and follow-up test are summarized by condition in Figure 4.4. Students' self-reported state anxiety levels did not differ by condition at pretest $F(1,417)=0.23$, $p = 0.64$. The pressure manipulation, which included a social-evaluative, expressive writing and feedback feature, was successful in inducing pressure and anxiety during Day 2 (the lesson) of the experiment. On Day 2, students in the imposed pressure condition had higher ratings of self-reported anxiety immediately after the pressure was given $F(1,346) = 40.55$, $p<0.001$, during the entire lesson $F(1,434) = 12.65$, $p<0.001$, and during the post-lesson test $F(1,425)=13.96$ $p < 0.001$. As designed, after the pressure manipulation debrief at the end of Day 2, self-reported feelings of anxiety did not differ by condition on the delayed, follow-up test $F(1,407)=0.95$, $p = 0.33$ on Day 3. There were no differences in state anxiety by grade at the pretest $F(1,417) = 0.34$, $p = 0.56$. There were also no difference in state levels of anxiety by grade from the imposed pressure on Day 2 $F(1,425) = 1.27$, $p = 0.26$.

Figure 4.4

Self-Reported Feelings of State Anxiety at Pretest, During Lesson, Post-Lesson Test, and Follow-up Test



Note. Average reported ratings of anxiety: i) during pretest, ii) during lesson, iii) during post-lesson test, and iv) during follow-up test. Worry scale is from: 0 *not worried/anxious at all* to 4 *extremely worried/anxious*. High support indicates high visual support via simultaneous presentation, linking gestures, and worked examples. Low support indicates sequential presentation, no linking gestures, and no worked examples. Pressure was administered prior to the lesson on Day 2. The pressure debrief was administered at the end of Day 2 after the post-lesson test to ensure students did not feel pressure on Day 3 for the follow-up test.

3.2 During the Lesson

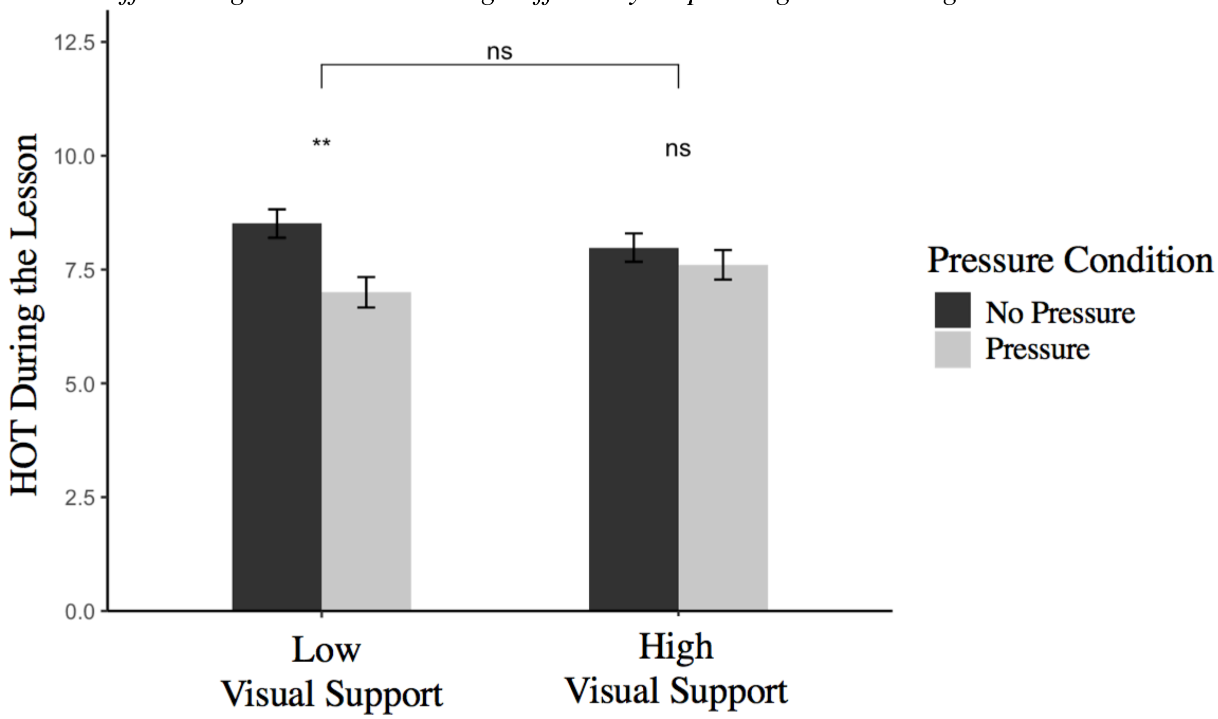
3.2.1 Higher Order Thinking During the Lesson

In the following sections we report how the conceptually demanding and cognitively taxing lesson on proportional reasoning and the associated intervention conditions impact different levels and types of learning and thinking during the lesson and during follow-up tests.

To begin we examined students' higher thinking during the lesson to determine if pressure and support influence critical thinking skills. We coded students' higher order thinking, that is their ability to compare, infer, abstract, and deduce hierarchies, during the lesson on free-response and multiple choice questions. Building from Study 2 during the lesson, first we examined if visual support enhances higher order thinking for those who are not under any forms of imposed pressure. A one-way ANOVA for individuals who did not receive pressure yielded no significant effects of high visual support on higher order thinking $F(1,206) = 1.42, p = 0.24$, controlling for fixed effects of classrooms.

To determine the main effects of pressure and visual support on higher order thinking, we conducted a two-way analysis of variance with an interaction effect, controlling for classroom fixed effects. The two-way ANOVA yielded a main effect for the pressure condition $F(1,427) = 96.73, p < 0.01$, such that students in the no pressure condition had more instances of higher order thinking (M = 8.25, SD = 3.30) than for those in the pressure condition (M = 7.31, SD = 3.43) during the lesson. The main effect of visual support was not significant $F(1,427) = 0.09, p = 0.93$. The interaction effect of pressure and visual support was trending toward significance $F(1,427) = 38.51, p = 0.06$. A post hoc Tukey HSD (honest significant difference) test correcting for p-values shows that for students who received low visual support, pressure negatively impacted students' higher order thinking scores at $p < 0.01$ (see Figure 4.5). Interestingly, for students in the high visual support condition, there was no effect of pressure. These findings are overall evidence to suggest that pressure and visual support students' ability to think critically and engage in higher order thinking. See Supplementary Materials (S 2.1 and Table S 2.1) for alternative analysis using linear regression controlling for individual differences; results are consistent.

Figure 4.5
Pressure Effects Higher Order Thinking Differently Depending on Teaching Environment



Note. Error bars signify standard errors. ** $p < 0.01$. Higher order thinking during the lesson scores ranged from 0-17.

3.3 Student Problem-solving Accuracy

3.3.1 Summary Statistics of Accuracy: Pretest, Post-Lesson Test, and Follow-Up Test

Next, we examined the condition effects on proportional reasoning problem-solving accuracy. Accuracy means and standard deviations of the test scores at pretest, post-lesson test, and follow-up test are summarized in Table 4.2. Random assignment of students within conditions was successful. At pretest, there were no significant differences in proportional reasoning prior knowledge between the four conditions for 5th grade students $F(1,319)=0.16$, p=0.69, nor 6th grade students $F(1, 107) = 1.58$, p = 0.21.

Table 4.2
Accuracy Scores by Condition, Grade, and Testing Day

	Pretest Accuracy			Post-Lesson Test Accuracy			Follow-Up Test Accuracy		
	5th Grade	6th Grade	*Mean*	5th Grade	6th Grade	*Mean*	5th Grade	6th Grade	*Mean*
Low Pressure, **Low** Visual Support	.18 (.18)	.33 (.26)	*.25 (.22)*	.26 (.21)	.51 (.30)	*.33 (.26)*	.24 (.21)	.45 (.32)	*.33 (.26)*
Low Pressure, **High** Visual Suppport	.18 (.13)	.32 (.25)	*.22 (.21)*	.21 (.19)	.48 (.23)	*.31 (.25)*	.21 (.20)	.40 (.28)	*.29 (.26)*
High Pressure, **Low** Visual Support	.18 (.16)	.38 (.29)	*.22 (.18)*	.25 (.19)	.48 (.33)	*.28 (.23)*	.24 (.21)	.48 (.35)	*.24 (.23)*
High Pressure, **High** Visual Support	.19 (.16)	.41 (.27)	*.22 (.21)*	.28 (.21)	.48 (.31)	*.33 (.26)*	.27 (.21)	.49 (.31)	*.30 (.26)*
Mean	*.18 (.16)*	*.36 (.26)*	*.22 (.21)*	*.25 (.20)*	*.49 (.29)*	*.31 (.25)*	*.24 (.21)*	*.46 (.31)*	*.28 (.25)*

Note. Performance accuracy is an average out of 6 assessment problems. Standard deviations are in parentheses.

3.3.2 Pretest Knowledge: Accuracy

First we report on pretest performance before analyzing students' post-lesson learning gains (immediately following the lesson) and follow-up learning gains (one-week delay). As expected, average performance accuracy on the pretest was low. Students in the 6th grade had higher performance accuracy compared to students in the 5th grade at pretest $F(1{,}428) = 70.01$, $p < 0.001$. On the pretest, there was a trending effect of gender such that males had higher scores than females ($t(288) = 1.94$, $p = 0.05$).

Trends across the study suggest that students across all conditions learned from the lesson. Additionally, there is little evidence of fadeout as accuracy scores remained relatively consistent on the follow-up test. The trend in gains was consistent between grades on the post-lesson test and follow-up test. Students in the 5th grade gained an average proportion of .07 (SD = .19) in total accuracy immediately after the lesson, and sustained average gains of .06 (SD = .20) from their pretest after a week's delay. Students in the 6th grade gained an average proportion of .13 (SD = .28) in total accuracy immediately after the lesson, and sustained average gains .08 (SD = .28) from their pretest after a week's delay.

3.3.3 Post-Lesson Learning Gains: Post-Lesson Accuracy – Pretest Accuracy

Post-lesson accuracy gains were calculated by subtracting students' post-lesson accuracy test scores from pretest accuracy scores. The effects of visual support on problem-solving accuracy for students who were not under the imposed pressure was examined with a one-way ANOVA. Results revealed no significant effect of visual support on post-lesson accuracy gains for students in the no pressure condition $F(1,16) = 1.29$, p = 0.26 . To examine the research question if pressure and its added anxiety interacts with visual support and influences students' post-lesson performance gains, a two-way analysis of variance with an interaction effect, controlling for classroom fixed effects was conducted. The two-way ANOVA yielded no main effect for the pressure condition $F(1,402) = 0.00$, p = 0.99, no main effect of visual support $F(1,402) = 0.68$, p = 0.41, and no interaction $F(1,402) = 0.40$, p = 0.48. Findings overall suggest no relationship between condition manipulation and post-lesson accuracy scores.

Secondarily, we examined how individual difference factors are related to post-lesson gains in accuracy. Regression analysis with imputed missing data can be found in Table 4.3. Model 1 tests the independent effects of condition on accuracy gains, controlling for fixed effects of classroom; Model 2 adds in theoretically relevant covariates. We conducted a global test to compare these models and they do not significantly differ from other another $F(1,376) = 1.14$, p = 0.31. Model 3 tests whether higher order thinking acts as potential mediator to problem-solving gains. Model 2 and 3 provide suggestive evidence that students with more sophisticated arithmetic skills have greater performance gains in accuracy than students with less prior knowledge. Contrary to a priori hypotheses (see Table 4.3, Model 3), a mediation of students' higher order thinking during the lesson was not indicated as this variable is not a significant predictor to immediate learning gains (p>0.05). See Supplementary Materials (S 2.3, Table S 2.3) for an alternative analysis that used only complete cases.

Table 4.3
Model Specifications for Post-Lesson Accuracy Gains with Imputed Data

	Model 1				Model 2				Model 3			
	B	SE	*t*	*p*	B	SE	*t*	*p*	B	SE	*t*	*p*
Pressure	-.16	.20	-.81	.42	-.10	.19	-.49	.62	-.06	.20	-.31	.76
Visual Support	-.25	.20	-1.29	.20	-.17	.20	-.86	.39	-.16	.20	-.82	.41
Pressure*Visual Support	.24	.27	.90	.37	.17	.26	.67	.51	.15	.26	.57	.57
Female	-	-	-	-	.25	.17	1.51	.15	.22	.17	1.35	.19
Attentional Control	-	-	-	-	.00	.00	-.58	.57	.00	.00	-.48	.63
Working Memory	-	-	-	-	.00	.03	1.79	.09	.05	.03	1.72	.10
Arithmetic Skills	-	-	-	-	**.07**	**.03**	**2.36**	**.02**	**.06**	**.03**	**2.07**	**.04**
Mathematics Anxiety	-	-	-	-	-.01	.01	-1.24	.22	-.01	.01	-1.20	.23
Grade	-	-	-	-	.29	.19	1.51	.14	.29	.19	1.53	.14
HOT During Lesson	-	-	-	-	-	-	-	-	.03	.02	1.30	.20

Note. Model 1 tests the independent effect of condition, controlling for classroom fixed effects. Model 2 adds in relevant covariates. Model 3 tests whether higher order thinking acts as mediator to problem-solving accuracy gains.

3.3.4 Sustained Follow-Up Learning Gains: Follow-up Accuracy – Pretest Accuracy

To examine predictors of students sustained learning gains from pretest to a week's delay, we ran the same regressions. Regression analyses with imputed missing data can be found in Table 4.4. There was no main effect of pressure or visual support on students sustained gains in accuracy. Secondarily, we examine how individual difference factors are related to sustained gains in accuracy. Model 1 tests the effects of condition on performance accuracy gains, controlling for fixed effects of classroom; Model 2 adds in theoretically relevant covariates. We conducted a global test to compare these models and they did not significantly differ from other another $F(1{,}292) = 0.57$, $p = 0.90$. Model 2 and 3 provide suggestive evidence that students with more sophisticated arithmetic skills have more sustained performance gains in accuracy. See Supplementary Materials (S 2.4, Table S 2.4) for an alternative analysis that used only complete cases.

Table 4.4
Model Specifications for Follow-up Test Accuracy Gains with Imputed Data

	Model 1				Model 2				Model 3			
	B	SE	*t*	*p*	B	SE	*t*	*p*	B	SE	*t*	*p*
Pressure	-.08	.20	-.43	.67	-.02	.20	-.10	.92	.01	.20	.04	.97
Visual Support	-.32	.22	-1.44	.16	-.26	.21	-1.24	.22	-.25	.21	-1.21	.23
Pressure*Visual Support	.37	.34	1.07	.29	.31	.33	.94	.36	.29	.33	.88	.39
Female	-	-	-	-	.12	.18	.68	.50	.10	.18	.59	.56
Attentional Control	-	-	-	-	.00	.00	-.27	.79	.00	.00	-.20	.84
Working Memory	-	-	-	-	.01	.07	.16	.88	.01	.07	.13	.90
Arithmetic Skills	-	-	-	-	**.09**	**.03**	**2.67**	**.01**	.08	.03	2.51	.01
Mathematics Anxiety	-	-	-	-	-.01	.01	-.87	.39	-.01	.01	-.86	.40
Grade	-	-	-	-	.00	.27	.00	1.00	.00	.27	.01	.99
HOT During Lesson									.02	.02	.86	.39

Note. Model 1 tests the independent effect of condition, controlling for classroom fixed effects. Model 2 adds in relevant covariates. Model 3 tests whether higher order thinking acts as mediator to problem-solving sustained gains.

3.4 Problem-solving Steps Learning Gains using Latent Class Analysis

3.4.1 Step 1: Model Fit for Problem-solving

To capture a more fine-grained assessment of students' problem-solving gains rather than accuracy only, we explored the nature and sequence of students' problem-solving from their problem-solving approach (i.e., proportional thinking or non-proportional thinking) to their procedures and their final solution accuracy. Assessing the multi-step procedures students must employ to solve proportional reasoning problems required that we use a three-step latent transition analysis. Through this analysis we could identify and characterize i) students' problem-solving learning and retention and then ii) examine learning and retention as a function of condition, grade, attentional control, working memory, arithmetic skills, and mathematics anxiety. The LTA analysis was conducted for problem-solving steps on the pretest, post-lesson test, and follow-up test. The mathematics packets completed at these three time points included 6 questions in which students were instructed to solve the problems using the equivalent fraction or unit ratio strategy.

The first step in the LTA analysis is fitting the model; this information for the problem-solving steps can be found in Table 4.5. The fit criteria indicted the five-group model was the best fitting model based on the BIC and CAIC; the three-group model fit best according to the AWE and ILC-BIC. To determine if the three or five group model was best, we ran a conditional bootstrap procedure to compare the three- and four-group models (Vermunt & Magidson, 2016). The bootstrap likelihood difference analysis shows significantly more information was gained from the four-group model (Bootstrap 2LL Diff = 359.34, $p< 0.001$). Next, a conditional bootstrap procedure was conducted to compare the four- and five-group models (Vermunt & Magidson, 2016). The bootstrap likelihood difference analysis shows significantly more information was gained from the five-group model (Bootstrap 2LL Diff = 175.92, $p< 0.001$). Consequently, the five-group model was the best fitting model because of its statistical fit, parsimony, and theoretical interpretability in characterizing patterns of proportional reasoning problem-solving.

Table 4.5
Model Fit Indicators for the Latent Transition Analysis of Combined Problem-Solving Steps

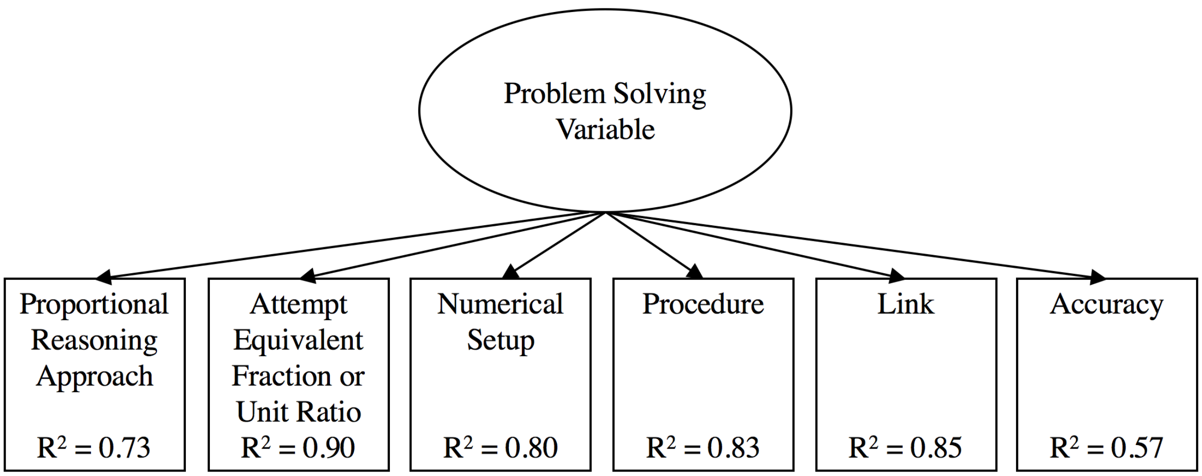
	LL	BIC	CAIC	AWE	ICL-BIC	Entropy R^2
1-Group	-14117.94	28272.90	28278.90	28327.92	28272.90	1.00
2-Group	-9022.69	18150.26	18167.26	18454.09	18298.20	0.92
3-Group	-8369.81	16937.05	16969.05	17493.64	17200.21	0.90
4-Group	-8190.14	16694.94	16745.94	17806.65	17339.00	0.83
5-Group	**-8102.19**	**16660.92**	**16734.92**	**18168.17**	**17489.62**	**0.81**
6-Group	-8074.77	16772.68	16873.68	18666.86	17740.73	0.79
7-Group	-8056.46	16927.31	17059.31	19191.64	17981.25	0.78

Note. Bold text indicates the selected model. LL, Log Likelihood; BIC, Bayesian Information Criterion; CAIC, Consistent Akaike Information Criteria; AWE, Average Weight of Evidence; ICL-BIC, a version of Integrated Classification Likelihood.

3.4.2 Step 2: Characterization of Problem-solving Groups

The cluster loadings for the problem-solving indicators, the distinct problem-solving steps, can be found in Figure 4.6. The problem-solving step indicators reported here merge key problem-solving steps for the equivalent fraction and unit ratio strategy. The loadings are obtained using a linear approximation of the class specific response probabilities (Vermunt & Magidson, 2005).

Figure 4.6
Cluster Loadings of the Problem-solving Indicators Estimated by the Latent Transition Analysis



3.4.2.1 Latent Classes for Problem-solving

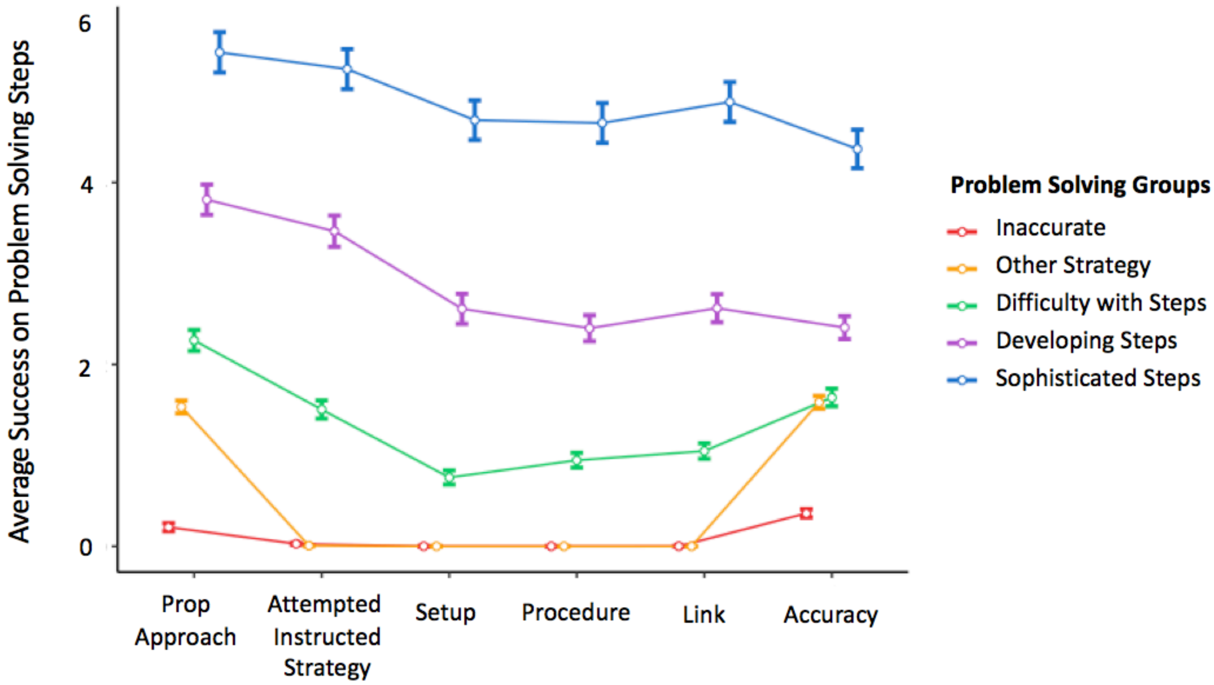
Next we looked at the specific latent classes that emerged from the item-response probabilities for the problem-solving steps. These latent classes represent groups of students and their distinct problem-solving abilities from approach to solution accuracy. The same latent groups exist at the pretest, the post-lesson test and the follow-up test; however, students move between groups across the time points as they gain (or lose) skills in problem-solving.

Five latent classes were identified and are summarized in Figure 4.7. Groups are characterized by increasing levels of sophistication and consistency in use for the equivalent fraction and unit ratio problem-solving steps. We present the least sophisticated to most sophisticated problem-solving groups in order.

One group was characterized by overall lack of proportional thinking; students in this group did not attempt the problems using a proportional reasoning strategy (i.e., they used incorrect subtraction or addition), they had incorrect procedures, and incorrect solution accuracy. This group was labeled “Inaccurate”. A second group, which we labeled “Other Strategy” was characterized by its use of other, correct proportional reasoning problem-solving approaches (e.g., build up, ratio, multiplicative thinking, etc.). This group had low rates of problem-solving step success as they did not attempt the equivalent fraction or unit ratio strategy. However, this group’s accuracy matched their approach level rates; this suggests that students who used other, less sophisticated problem-solving strategies were doing to successfully and arriving at the correct final solution. The next group had slightly higher levels of proportional reasoning problem-solving approaches and used the instructed strategies. This group had difficulty implementing the accurate problem-solving procedures, as shown by the decline in setup and procedures; it is called “Difficulty with Steps”. A fourth group attempted the instructed strategies more frequently but still had difficulty with the problem-solving steps; this group is called “Developing Steps”. A final group was characterized by exclusive attempt of the instructed strategies across the problems with success across all procedures; this group was labelled “Sophisticated Steps”.

Figure 4.7

Problem-Solving Groups as Identified by the Latent Transition Analysis



Note. Five problem-solving groups emerged from the LTA with increasing levels of frequency and sophistication of the equivalent fraction and unit ratio strategy: "Inaccurate", "Other Strategy", "Difficulty with Steps", "Developing Steps", and "Sophisticated Steps". Error bars indicate standard error. X-axis represent the sequence of problem-solving steps for the problems: approach using proportional reasoning strategy, attempt instructed equivalent fraction or unit ratio strategy, setup of the numbers, procedure, link, and accuracy. The mean of the y-axis represents mean for each of the problem-solving steps across six problems.

3.4.2.2 Transition of Group Membership by Testing Time Point

Next, we examined learning gains across the study from: pretest to the post-lesson test to the follow-up test. The nature of the LTA means that students can change group membership across each time point. A change from a less to a more sophisticate problem-solving group is interpreted as more frequent attempt of the instructed solution strategies and learning gains of the problem-solving steps from the instruction. A change from more to less sophisticated groups could be interpreted as a loss learning. We interpreted change across these groups in an ordinal nature from least to most sophisticated problem-solving groups. The overall trend shows that students are moving from less sophisticated groups to more sophisticated groups from the pretest to post-lesson test. From the post-lesson test to the follow-up test students mainly stayed in the same groups or moved to less sophisticated groups, suggesting that there is slight a fadeout effect on learning.

The probabilities in Table 4.6 show 1) the probability of belonging to each problem-solving group at pretest, 2) the probability of transitioning to a different group at the post-lesson test, and 3) the probability of transitioning to a different group at the follow-up test. Higher probabilities indicate greater likelihood of movement. At pretest, almost half of the students were in the least sophisticated, "Inaccurate" group and did not have any understanding of proportionality. Less than less than 1/5 of students were in the top three groups suggesting that a fraction of the students in the study solved problems on the pretest using the equivalent fraction or unit ratio strategy.

The LTA analysis shows that there was indeed learning from the lesson. On the post-lesson test, students moved to more sophisticated groups and or stayed in their pretest groups. For students who started in the "Inaccurate" group and had incorrect strategy approaches at pretest, over one-third of these students moved to groups that attempted a proportional reasoning strategy or the instructed strategies with success. This suggests that the lesson not only encouraged problem-solving for the instructed strategies but also encouraged proportional thinking more broadly (e.g., students who used buildup strategy). Students who understood proportional reasoning properties but not the instructed strategies at pretest, were grouped in the "Other Strategy" group at pretest. More than half of these students moved to problem-solving groups that used the instructed solution strategies and with success. This supports the learning sciences literature which posits that greater prior knowledge, in this case familiarity with multiplicative thinking, affords for greater learning gains. Less than 1/5 of the students were familiar with the instructed strategies prior to the lesson;

after the lesson, students in these groups were likely to have more refined problem-solving skills and were completing problem-solving steps with higher frequency and success.

Table 4.6
Membership Status Between Pretest, Post-Lesson Test and Follow-up Test

1) Initial probabilities at Pretest

	Pretest Problem Solving Groups				
	Inaccurate	Other Strategy	Difficulty with Steps	Developing Steps	Sophisticated Steps
	.46	.32	.16	.03	.03

2) Transition Probabilities from Pretest to Post-Lesson Test

	Post-Lesson Test Problem Solving Groups					
	Inaccurate	Other Strategy	Difficulty with Steps	Developing Steps	Sophisticated Steps	*Total*
Inaccurate	*.62*	.12	.12	.12	.02	*1.00*
Other Strategy	.02	*.41*	.20	.27	.10	*1.00*
Difficulty with Steps	.01	.17	*.33*	.28	.21	*1.00*
Developing Steps	0	0	0	*.35*	.65	*1.00*
Sophisticated Steps	0	0	0	0	*1.00*	*1.00*

3) Transition Probabilities from Post-Lesson to Follow-up Test

	Follow-up Test Problem Solving Groups					
	Inaccurate	Other Strategy	Difficulty with Steps	Developing Steps	Sophisticated Steps	*Total*
Inaccurate	*.97*	.01	0	.02	0	*1.00*
Other Strategy	0	*.85*	.10	.05	0	*1.00*
Difficulty with Steps	.18	.33	*.35*	.14	0	*1.00*
Developing Steps	.06	.17	.16	*.58*	.03	*1.00*
Sophisticated Steps	0	0	.04	.25	*.71*	*1.00*

Note. "Inaccurate" is interpreted as the least sophisticated, inaccurate problem-solving group. The problem-solving groups increase in an ordinal fashion based on sophistication, frequency, and success of problem-solving procedures. The italicized diagonal indicates probability of staying in the same group.

Examination of the post-lesson test to the follow-up test shows that students in the least and most sophisticated groups maintain their group membership at the follow-up test most consistently. Students who had the most instability in their knowledge (i.e., "Difficulty with Steps" group) were the most instable in their group membership. Half of these students dropped down to less sophisticated problem-solving groups and half of the students remained in this group or moved up. This is not surprising and these students were the students most susceptible to the lesson as their knowledge on proportionality and the solution strategies was elementary. Students in the most sophisticated groups (i.e., "Developing Steps" and "Sophisticated Steps") were most likely to stay in their post-lesson group. However, there is a proportion of students who dropped down to less sophisticated groups suggesting a fadeout effect.

3.4.3 Step 3: Covariates of Problem-solving Group Membership

3.4.3.1 Covariates at Pretest

Finally, we examined how condition and individual factors are related to group membership. First we examined individual difference factors that predicted students' problem-solving skills at pretest. Gender, grade, attentional control, working memory, arithmetic skills, and trait math anxiety were entered as covariates. Model outcomes are shown in Table 4.7. The coefficients indicate the probability of belonging to groups. Higher positive values indicate a higher probability of belonging and negative values indicate a lower probability of belonging.

The Wald's test indicated gender, grade, working memory, arithmetic skills, and math anxiety predicted pretest group membership. There was no effect of attentional control on group membership. Female students, 6^{th} grade students, students with better arithmetic skills, and lower anxiety were likely to be in the sophisticated groups.

Table 4.7
Covariates of Pretest Problem-Solving Group Membership

			Problem Solving Groups				
	Wald	p-value	Inaccurate	Other Strategy	Difficulty with Steps	Developing Steps	Sophisticated Steps
Female	**19.63**	**.00**	**-0.56 (.23)**	**-0.56 (.18)**	**-0.40 (.20)**	**-0.87 (.33)**	**2.39 (.56)**
Grade	**48.72**	**.00**	**-1.78 (.43)**	**-0.92 (.22)**	**-0.63 (.24)**	0.61 (.54)	**2.72 (.61)**
Attentional Control	4.64	.33	0.00 (.00)	-0.00 (.00)	-0.00 (.00)	-0.00 (.00)	0.01 (.01)
Working Memory	**21.03**	**.00**	-0.03 (.09)	**0.13 (.06)**	0.12 (.08)	**0.46 (.14)**	**-0.69 (.19)**
Arithmetic Skills	**68.41**	**.00**	**-0.71 (.11)**	**-0.26 (.07)**	-0.12 (.09)	-0.28 (.17)	**1.37 (.20)**
Math Anxiety	**13.02**	**.01**	**0.08 (.02)**	**0.05 (.02)**	0.04 (.02)	**0.06 (.03)**	**-0.23 (.07)**

Note. Bold indicates significance. Standard error values are in the parentheses.

3.4.3.2 Covariates at Post-Lesson Test

To examine the relationship between experimental conditions and problem-solving group membership on the post-lesson test we conducted a cross-classification analysis. This analysis showed LTA groups and condition were not significantly associated (χ^2 (12, n = 441) = 18.03, $p = 0.11$).

Next, we examined how condition and individual difference predict students' problem-solving gains from the pretest to the post-lesson test using a step-3 covariate analysis (see Table 4.8). Pretest group membership, experimental condition, grade, gender, attentional control, working memory, arithmetic skills, and trait math anxiety were entered as covariates. In this analysis, to account for pretest group membership, we entered the pretest group membership as a covariate (an option in the Latent GOLD Step 3 LTA).

The Wald's test indicated the interaction of condition (pressure * visual support) was significant. The step-3 covariate analysis suggested a trending effect such that students in the high pressure, high support condition were likely to be in the "Sophisticated Steps" group and were unlikely to be in the "Other Strategy" group. There was also evidence that gender, working memory, arithmetic skills, and grade predicted post-lesson group membership. Specifically, female students, students with higher working memory, students with higher arithmetic skills, and students in the 6th grade were most likely to be in the sophisticated problem-solving groups. However, these findings are suggestive and should be interpreted with caution.

Table 4.8
Condition Effects and Covariates of Pretest to Post-Lesson Problem-Solving Group Membership

			Inaccurate		Other Strategy		Difficulty with Steps		Developing Steps		Sophisticated Steps	
	Wald	*p* value	B	SE	B	SE	B	SE	B	SE	B	SE
Pretest Group Membership	**204.23**	**.001**										
Pressure	.65	.96	-.13	.33	.02	.16	-.05	.15	.01	.15	.16	.21
Visual Support	3.78	.44	-.41	.32	.01	.16	-.09	.15	.20	.16	.28	.21
Pressure * Visual Support	**11.27**	**.02**	-.31	.32	*.30*	*.16*	**-.33**	**.15**	.00	.15	*.34*	*.21*
Female	**12.77**	**.01**	*-.67*	*.35*	-.15	.16	.12	.16	*.32*	*.17*	*.38*	*.22*
Attentional Control	1.40	.84	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Working Memory	**10.35**	**.04**	-.17	.14	-.07	.08	**-.19**	**.08**	.11	.09	**.32**	**.16**
Arithmetic Skills	**32.09**	**.001**	**-.53**	**.15**	-.07	.07	.01	.07	**.18**	**.07**	**.42**	**.10**
Mathematics Anxiety	5.55	.24	-.07	.04	.03	.02	.01	.02	.03	.02	-.01	.03
Grade 6	**15.96**	**.001**	-.15	.40	-.19	.18	-.32	.20	-.08	.18	**.74**	**.21**

Note. Bold indicates significance p<.05; italicized represents trending towards significance p<.10.

3.4.3.3 Covariates at Follow-Up Test

Finally, to examine the factors that influence whether a student retains the learned strategies from the lesson or experiences fadeout, we conducted a step-3 covariate analysis for the differences between the post-lesson test and the follow-up test (see Table 4.9). Post-lesson group membership, experimental condition, gender, attentional control, working memory, arithmetic skills, trait math anxiety, and grade were entered as covariates. In this analysis, to account for post-lesson group membership, we entered the post-lesson group membership as a covariate (an option in the Latent GOLD Step 3 LTA).

The Wald's test indicated no significant effect of condition or the interaction on sustained learning from the post-lesson test to the follow-up test. There was suggestive evidence of individual difference factors such that students with less developed arithmetic skills and less anxiety were likely to be in the least sophisticated groups and students with more developed arithmetic skills and more math anxiety were likely to be in the more sophisticated groups at follow-up test. However, these findings are suggestive and should be interpreted with caution.

Table 4.9
Condition Effects and Covariates of Post-Lesson to Follow-up Problem-solving Group Membership

	Model 1											
			Inaccurate		Other Strategy		Difficulty with Steps		Developing Steps		Sophisticated Steps	
	Wald	*p* value	B	SE	B	SE	B	SE	B	SE	B	SE
Post-Lesson Group Membership	**19.50**	**.001**										
Pressure	7.28	.12	.30	.45	1.07	.45	-.18	.37	-.43	.43	-.76	.72
Visual Support	1.22	.87	.20	.41	-.20	.36	-.17	.33	-.02	.42	.19	.76
Pressure * Visual Support	*7.74*	*.10*	*1.03*	*.45*	.17	.38	*.95*	*.42*	*-1.17*	*.48*	-.98	.77
Female	*8.53*	*.07*	.07	.48	.23	.43	-.20	.42	*1.67*	*.60*	-1.77	1.31
Attentional Control	5.12	.28	*-.01*	*.01*	*-.01*	*.01*	-.01	.01	-.01	.01	.04	.02
Working Memory	3.19	.53	-.16	.19	.03	.18	.12	.17	-.01	.22	.02	.51
Arithmetic Skills	**13.91**	**.01**	**-.52**	**.24**	.23	.20	**-.42**	**.20**	**.81**	**.24**	-.10	.48
Mathematics Anxiety	**11.94**	**.02**	-.06	.06	**-.15**	**.05**	**-.13**	**.05**	.07	.06	**.26**	**.11**
Grade 6	4.51	.34	-.55	.53	-.03	.46	-1.06	.54	-.23	.50	1.87	1.31

Note. Bold indicates significance.

4. Discussion

This study is the first to investigate how pressure affects school-aged children's learning of a conceptually demanding mathematical context, proportional reasoning. We imposed pressure via false negative feedback, social-evaluative pressure (Beilock & Carr, 2001), and expressive free writing targeted to heighten the level and awareness of the pressure. This work adds to the literature in many ways. First, we provide a set of procedures that effectively induce anxiety and worry within ethical levels to middle school students. The findings show that pressure and visual support impact students engagement with higher order thinking but there are no significant effects of condition on problem-solving.

The literature on pressure during test performance finds that heighted feelings of anxiety or imposed pressure lead to decrements in test performance (e.g., Beilock and Carr, 2005; Beilock et al., 2010). The results of this study suggest that pressure and its associated anxiety impacts *learning* but varies based on the type of cognitive tasks. Pressure affects higher order thinking but does not significantly affect gains in problem-solving procedures.

The lesson included cognitively demanding opportunities for higher order thinking which were intended to promote lasting conceptual learning. The questions here were used to promote critical thinking of the solution strategies and problem types by engaging the students in comparison, inference and abstraction. This higher order thinking requires that the student engage with multiple solution simultaneously, map procedures and structural level features across the solution strategies and problem types, and manipulate features of the problem context; it is both cognitively complex and taxes the executive functioning system. Findings from this study suggest that imposed pressure which places a burden on cognitive resources, disrupt students' learning and engagement for opportunities in higher order thinking. However, these findings systematically affect students based on the learning environment. Pressure affects instances of higher order thinking but only for students in the low visual support condition. This suggests that high support learning contexts mitigate individual difference factors such as executive functioning and pressure. Our interpretation of these findings is that when students' cognitive resources are not strained, that is when there is high visual support from the worked examples, pressure is not having an effect because there are enough cognitive resources to spare. When cognitive resources are limited because students are holding two solution strategies in mind, mapping across the strategies, and manipulating them, the added effects of the pressure overload the resources system and negatively impact instances of higher order thinking. These findings suggest that the learning environment is crucially important for supporting students' cognitive load systems.

In regards to problem-solving, pressure does not affect problem-solving performance accuracy on the post-lesson test or the follow-up test. There are trending effects for the interaction of pressure and visual support on the nature and sequence of problem-solving. Trending results suggest that students who were in the high pressure, high visual support condition were most likely to be in the most sophisticated groups and least likely to be in the least sophisticated groups. This trending effects requires more research but suggests that when students are visually supported (i.e., simultaneous presentation, gestures, worked examples) while under imposed pressure they attempt new solution strategies and do so with success. This is in line with some literature which argues that pressure can increase motivation and effort (e.g., Angrist & Lavy, 2009; Vevy, List & Sadoff, 2016). These findings do suggest however that in order for the pressure to have a positive impact on problem-solving it must be paired with visual support suggesting that pressure does not act as a motivator unless learning environment provides necessary scaffolds and supports.

While there were limited gains in performance accuracy on the post-lesson and follow-up tests, students still learned from this lesson. Results from the LTA analysis suggest that students consistently transition from less sophisticated to more sophisticated groups. For students with no understanding of proportional reasoning strategies at pretest, nearly one-third of students from that group attempted proportional reasoning strategies and with success. More than half of the students who had some familiarity with the instructed strategies were completing the proportional reasoning problems using the instructed strategies with more success and accuracy. This supports the learning sciences literature which posits that greater prior knowledge, in this case familiarity with multiplicative thinking, affords for greater learning gains (Hailikari, Katajavouri, and Lindblom-Ylanne, 2008). These findings highlight that while students did not master the problem-solving procedures, students attempted the novel solution strategies and with success.

The step-3 covariate analysis also suggested a trend for condition predicting problem-solving groups. While the results were not significant, it is beginning evidence to suggest that students who are placed under pressure but supported are most likely to attempt novel solution strategies and with the support, do so with success. Consistent with other findings from this dissertation, gender, working memory, arithmetic skills, and grade predicted post-lesson group membership. These results suggest that individual difference factors affect if and how students engage with novel solution strategies in a short lesson. As these findings are suggestive future research should explicitly examine how these individual factors influence students' problem-solving and engagement with higher order thinking.

4.2 Limitations and Future Directions

There are a few key considerations from this study for future work in particular. The lesson on proportional reasoning was complex and challenging. It would be worthwhile to conduct a similar student with older students that have more prior knowledge to examine if these measures yield greater gains such that the students would be able to grasp the complexities of this lesson more fully.

Second, imposed pressure in an experimental context must be considered. Using imposed pressure in experimental contexts allow for experimental consistency and rigor. However, imposed pressure may not impact learning and performance in the same ways as real-life academic pressures. If imposed pressure does not have the same mechanisms as authentic mathematics anxiety or learning pressure, this could impact how findings are translated from research contexts to the classroom. More work is required to examine if the findings using imposed pressure are robust or if the mechanisms of real-life, authentic performance pressure differs in affecting learning.

Understanding the effects on authentic pressure and anxiety from sources outside of the classroom on learning is also necessary and especially relevant in the context of the global COVID-19 virus. It is important to consider whether the type and valence of the pressure systematically influences learning potential. The global COVID-19 pandemic, for example, has been a major source of stress and anxiety (Mesghina & Richland, *under review*). My current work examines how students' feelings of perceived anxiety from sources outside of the classroom may systematically influence their attention, learning, and performance. Moreover, we examine if worked examples have the potential to ameliorate the negative effects of pressure by freeing up cognitive load resources.

CHAPTER FIVE: DISCUSSION

Teaching new and complex mathematics content is highly challenging, especially in the middle school mathematics classroom where teacher's teachers must develop one lesson that fits the varying needs and skills of all their students. Findings from the three studies in this dissertation suggest that visual support in the form of worked examples can be used as a highly implementable tool to scaffold learning of novel and sophisticated solution strategies in the complex domain of proportional reasoning both on their own or paired with direct teacher instruction. Additionally, these findings suggest that worked examples combined with instruction can buffer against the detriments caused by pressure or lower cognitive functioning. These findings occur for conceptually demanding contexts like in higher order thinking contexts but pressure does not impact learning for procedures.

Worked examples are visual supports that are presented as a means to scaffold learning by providing an opportunity for students to map an expert strategy on near-transfer problems (Sweller & Cooper, 1985). Importantly, worked examples also reduce students' cognitive load freeing up cognitive resources for students to focus on structural, conceptual and procedural components of problem-solving, rather than requiring them to hold a full problem solving procedure active in working memory while adapting it to new target problem contexts (Tuovinen & Sweller, 1999).

Across three studies in this dissertation, we used the context of proportional reasoning to study the potential benefits of worked examples. Although conceptually challenging and cognitively demanding, proportional reasoning characterizes important structural relationships in mathematics and science as well as in everyday life (Cramer & Post, 1993; Lesh, Post, & Behr, 1988), it foregrounds other complex domains in mathematics (Epson, 1999; Fuson & Abrahamson, 2005), and it is deemed a crucial math skill for students to develop (National Council of Teachers of Mathematics, 1989). Despite the importance of mastering the fundamentals of proportionality, many students have pervasive difficulties. Additionally, teachers find this curriculum area to be quite challenging to teach (Sowder, 2007).

When learning to reason proportionally, there are many strategies students can employ that vary in their level of sophistication. Many students begin by using incorrect strategies that do not reflect the multiplicative nature of proportional thinking (e.g., additive thinking), and then progress to strategies that use addition (e.g., build up strategy). After this they may move to more sophisticated solution strategies such as those studied in this dissertation- equivalent fraction and unit ratio strategies. A shift from more elementary strategies to more sophisticated strategies not only require a high level of procedural fluency but are more conceptually and cognitively challenging (Cramer & Post, 1993; Boyer, Levine, & Huttenlocher, 2008). We sought out to explore the ways in which worked examples can support learning of these sophisticated strategies, build fluency in problem-solving skills, and engage students in higher thinking regarding the different strategies and their benefits.

Previously, research on worked examples has focused on developing students' problem-solving skills *after* instruction (see Renkl, 2014). In three studies, this dissertation i) examined worked examples as a pedagogical tool to *introduce* students to novel yet sophisticated solution strategies, ii) evaluated if and how worked examples can be incorporated during direct teacher instruction, and iii) explored if worked examples could improve outcomes for students who have temporarily low processing resources due to imposed pressure and heightened feelings of anxiety in both higher order thinking and problem-solving contexts. The research focuses on youth in low socio-economic, ethnically diverse schools at a crucial transition point, between upper elementary and middle school, at which academic motivation tends to reduce dramatically and where US student performance starts to decrease relative to other nations (e.g., Wigfield & Eccles, 2002).

5.1 Key Findings

Across all of these studies, we examined the intervention effects of worked examples on students' performance accuracy. But more uniquely and importantly, we captured how worked examples encourage students to fundamentally think and solve problems in new ways. To do so, we examined the approaches and procedures students employed to solve the proportional reasoning word problems. Students' qualitative written work was coded to capture their approach, setup, procedures, and final solution accuracy. In order to analyze these problem-solving steps, latent class analyses (LCA) was used to identify and characterize the full sequence of steps students took to solve proportional reasoning problems. The LCA allowed for the identification of distinct patterns of problem-solving, and grouped students by their problem-solving abilities. Finally, I examined how worked examples influenced the ways students engaged with higher thinking skills such as the ability to compare, abstract and infer. I will briefly review the results of each study in more detail.

Study 1 (Chapter 2) examined if worked examples can be used as a pedagogical tool to *introduce* solution strategies prior to direct teacher instruction. Within classrooms, students were assigned to the treatment condition of 1) fully worked examples or 2) partially worked examples and results were compared to a 3) control problems only condition. The study shows that in a single classroom session with only a mathematics booklet and one worked example for two solution strategies, both partially and fully worked examples can fundamentally change the way students think about and solve proportional reasoning word problems and improve performance accuracy. Providing exemplars not only encourage students to use novel solution strategies but provides the scaffolding for students to implement and map the problem-solving procedures with success. Moreover, there are benefits for all students, regardless of their prior knowledge and attentional control. While both fully and partially worked examples encourage use of novel solution strategies, findings suggest that the ideal pedagogical tool to introduce novel solution strategies is through fully worked examples (see Renkl, 2014) as students cannot intuit the missing procedures without sufficient prior knowledge. This study also validates our coding scheme and analytic tool, LCA, as a sensitive measure to capture variations in students' problem-solving.

Study 2 (Chapter 3) expanded on the findings from Study 1; we examined if a *double-dose* of worked example support for instruction, that is a high-quality teacher-led lesson paired with worked examples *during* instruction, allow students to learn of novel solution strategies more effectively than a high-quality lesson alone. In this study, all students watched a previously-recorded, conceptually challenging mathematics lesson on proportional reasoning. Within classrooms, students were assigned to the treatment condition of worked examples or no worked examples. Solution accuracy of problems during instruction did not vary between conditions. These findings were not particularly surprising as the arithmetic required for the link and final solution used non-whole numbers from non-divisible relationships; this has been shown to be a point of difficulty for students (Fuson & Abrahamson, 2005) and we did not specifically instruct on the arithmetic of problem-solving.

Examination of the nature and sequence of problem-solving in Study 2 provided a more nuanced perspective to understanding learning gains from worked examples. Findings from the LCA analysis suggested that students with *double-dose* of support (instruction + worked examples) were more likely to attempt the instructed strategy and had the greatest conceptual and procedural success on problem-solving steps for the more familiar, equivalent fraction strategy. For the more sophisticated and novel unit ratio strategy, results were trending in the same way but were not significant.

An interesting finding from this study was that over half of the students in the no worked example condition did not attempt the instructed strategies on the near-transfer problems. This suggests that without sufficient support via scaffolding, students do not have a strong enough representation of the problem-solving strategy after a single lesson, and that worked examples helped to support the representation as the students try and implement the novel procedures. Findings from this study highlight the benefits of worked examples *during* instruction as a way to provide the necessary scaffolds for students to attempt novel solution strategies on their own and do so with success.

While we did not anticipate that the mathematics packet or video instruction would lead to mastery of these sophisticated proportional reasoning strategies, findings from Study 1 and 2 highlight just how difficult this type of problem-solving is. Specifically, it highlights how difficult it is to change students' approach to the multiplicative nature of proportional reasoning. During instruction of the equivalent fraction strategy, a large proportion of the sample used additive instead of multiplicative procedures. During instruction of the unit ratio strategy, many students did not complete the two key aspects of problem-solving (i.e., division to find the unit, and multiplication to apply- multiplicative nature) or had difficulty deciding which numbers were to be compared. These key misconceptions suggest that one packet or one, 45-minute video lesson is not effective on its own for many students; reframing students' thinking from an additive perspective to understanding proportionality and its multiplicative nature requires more instruction for students to really master the topic. It may be that if we focused during instruction on these persistent difficulties we would observe more learning gains.

Findings across these two studies also highlight other factors that should be more fully explored in future studies to develop a full representation of the factors that impact learning from worked examples. Several characteristics of the instructional context seemed to affect learning, including: i) the format of the worked examples, ii) procedural instruction provided for the worked example, iii) the environmental context may effect learning gains, and iv) the complexity of numbers to be compared. In Study 1 we presented students of paired down, worked examples with explicit procedural steps. In Study 2, we used worked examples that were screenshots of the teacher's solution strategy but also highlighted the key conceptual features of the solution strategy. The effects on performance accuracy and problem-solving approaches were stronger in Study 1 than study 2. There are a few potential explanations for this. First, there may be differential effects for the format of the worked examples: typed solution strategies versus screenshots of a teacher's written strategy. Second, the amount of instruction of the worked example may affect learning. Findings from these studies suggest that highlighting key procedural steps influence problem-solving gains more than highlighting the key conceptual framework. This is in line with other work which suggest that labeling is an important feature for learning (Atkinson, Derry, Renkl, Wortham, 2000).

Third, the environment may affect gains. As we moved to a more ecologically valid learning environment the intermediate learning gains may be slower in more fully controlled format such as packets as in Study 1. Additionally, this ecologically valid video lesson may place more cognitive demands reducing the overall magnitude of learning gains. More research is necessary to determine the source of influence. Finally, in Study 1, the numbers to be compared were more easily divisible (while they still required decimal long division). Therefore, during a learning context, it is important to consider the arithmetic required for problem-solving and determine if complex arithmetic is a key feature of the lesson. This is something to consider in relation to the grade level, school district, prior levels of school success, and time of school year, which are complex to map out ahead of time but may be crucial for understanding variability in school-based studies.

Study 3 (Chapter 4) extended on Study 1 and 2 and examined the effects of worked examples paired with other visual supports during direct teacher instruction (e.g., simultaneous presentation, linking gestures, multiple strategy solutions) on students' 1) higher order thinking during instruction and 2) problem-solving gains on post-tests. Additionally, this study examined how pressure and its associated anxiety, which has been shown to strain students' cognitive resources (Eysenck et al., 2007) and effect test performance (Beilock, 2010) influenced learning potential. Finally, we examined whether reducing cognitive burden through visual support could improve outcomes for students with temporarily low processing resources due to an imposed pressure manipulation and associated anxiety.

The imposed pressure, consisting of social-evaluative, feedback, and expressive writing was successful in eliciting pressure in our sample. During the lesson, there was a negative main effect of pressure on students' higher order thinking. When the interaction between pressure and visual support was examined results suggested that visual support provides a buffer against the negative effects of pressure. That is, for those in the high visual support condition there were no observed effects of imposed pressure. For those in the low visual support condition, students who were in the imposed pressure condition had lower instances of higher order thinking than those in the no pressure condition. These findings reinforce how important it is to provide students with high quality support as it has the potential to reduce the negative impacts of pressure or low cognitive resources. Additionally, it highlights that pressure and its associated anxiety consumes cognitive resources that are required for conceptual higher order thinking, replicating previous work (Cho, Holyoak, Cannon, 2007).

In regards to problem-solving, there were no significant effects of pressure and visual support on students' problem-solving accuracy or problem-solving approaches and procedures (from the LTA). There were however, trending effects from the LTA which suggested that students in the high visual support and imposed pressure condition were likely to be in the most sophisticated problem groups. More research is necessary to examine these trending effects. While there were no significant effects by condition, there was overall learning gains from pretest to post-test and sustained learning from post-test to follow-up test suggesting again that even during a short- 45 minutes' lesson, high-quality instruction has the potential to fundamentally change the way students think about and approach problem-solving in a content area such as proportions.

The overall findings from Study 3 suggest that pressure and visual support may affect conceptual engagement and learning (i.e., higher order thinking) more than problem-solving gains. More research is necessary to examine these differential effects for higher order thinking and problem-solving.

5.2 Implications of the Research Findings

In today's 21st century world mathematics, problem-solving, and higher order thinking is increasingly important. Furthermore, in the mathematics classroom, real-world word problems are encouraged as part of the curriculum to promote motivation and "sense-making" of mathematical principles (Schoenfeld, 1991). Not only do word problems encourage application of symbolic mathematics and procedures to the real-world context, but it engages the student in higher order thinking, another crucial 21st century skill. Problem-solving in a word-problem context requires that the student infer the main goal of the problem, compare multiple strategies solutions, determine the optimal strategy solution, and then employ those procedures (Palm, 2008). Not only is this conceptually complex, it is cognitively effortful (Berends & van Lieshout, 2009). These multiple steps and multilevel reasoning make word problems notoriously difficult. Nationally, one study found that children performed 10-30% lower on arithmetic word problems than the same problems in their symbolic form (Carpenter, Corbitt, Kepner, Linquist, & Reys, 1980).

High quality instruction which places emphasis on comparison, inference, and abstraction has the potential to support problem-solving learning while also building other crucial higher thinking skills (Richland & Simms, 2015). The latent class analyses in these studies gave insight into the problem-solving procedures that consistently present difficulties for studies and therefore requires more attention and focus. The LCAs identify that both proportionality (e.g., the multiplicative link) and arithmetic fluency suggest that both teacher instruction and worked examples should scaffold and support learning more in these areas. Specifically, instruction which focuses students' attention on the structural features of problem-solving and encourages use of sophisticated solution strategies which highlight the multiple ways of approaching a problem (Richland, Zur & Holyoak, 2007), has the potential to scaffold learning and build students critical thinking skills.

The findings have theoretical implications for math education research and practical implications for classroom instruction. The research suggests that learning should not only be conceptualized as complete mastery of problem-solving (i.e., performance accuracy) but rather should examine learning as a process in which students employ different strategies solutions and procedures. This approach allows for a more detailed understanding of gains in a research setting and allows researchers to examine how pedagogical tools fundamentally influence the way students reason about and approach problem-solving.

Previous work has shown that worked examples can support learning *after* instruction (see Renkl, 2017). Findings from this work show that that worked examples can be used as a pedagogical tool to *introduce* and support skill development of novel solution strategies *during* instruction. This was shown in Study 1 where worked examples were paired down and clearly delineated the problem-solving steps and in Study 2 and 3 where worked examples were integrated with ecologically valid video lessons. Overall, findings suggest that worked examples should be used more frequently as a way to support and scaffold students' learning.

In line with current literatures, findings also suggest that pressure and its associated anxiety, which taxes individual's cognitive resources, can negatively influence students' higher order thinking skills (Maloney, Sattizahn, & Beilock, 2014). This was shown in Study 3 where pressure impacts students' higher order thinking but only for those in the low visual support condition. These findings highlight the importance of high quality instruction and visual support as it has the potential to reduce the negative impacts of pressure or low cognitive resources. Additionally, it highlights that pressure and its associated anxiety consumes cognitive resources that are required for conceptual higher order thinking, replicating previous work (Cho, Holyoak, Cannon, 2007). As expected, pressure does not have the same impact on procedural skills like problem-solving as the same cognitive resources are not required for this type of learning and engagement. Overall, evidence from these studies show that even with short, high quality instruction, students can engage with novel solution strategies and importantly in supported learning environments, engage in higher order thinking.

5.3 Future Research

There are many avenues for future research following this work. Based on the findings from Study 1, it is necessary to examine the utility of faded worked examples as a tool to introduce solution strategies in a broader time frame or in a longer mathematics packet. It may be that partial worked examples can be faded within one mathematics packet. For examples, mathematics packets can present fully worked examples of novel solution strategies and then faded worked examples with procedures eliminated over the course of the introductory packet. This instructional protocol may allow for students to see the full problem-solving approach and procedures while also providing an opportunity for exploratory learning. A packet like this may also prime students for teacher instruction on these solution strategies as students may be more motivated after engaging with the problem strategy in this capacity (DeCaro, DeCaro, Rittle-Johnson, 2015).

Interestingly, the magnitude of learning was greater in Study 1 than compared to Study 2. Study 1 used worked examples that had clearly delineated steps whereas Study 2 used worked example screenshots from the lesson. A necessary next step would be to replicate Study 2 but instead of using the worked example from the teacher, test whether a worked example that clearly delineates the problem-solving steps produces larger gains in learning, similar to the findings from Study 1. Findings from other research suggest that labeling or segmenting of procedures support learning (Atkinson, Derry, Renkl, Wortham, 2000). Worked examples that highlight key procedural steps may produce larger gains in learning compared to more ecologically valid worked examples that simply show the solution strategy and procedures (i.e., a screenshot from the teacher).

In Studies 2 and 3, we presented students with a short but conceptually challenging and cognitively taxing lesson on proportional reasoning and examined the effects of worked examples during instruction. To further our understanding, it is necessary to examine how worked examples can be used to reactivate learning on the follow-up posttest. Worked examples before the follow-up posttest may provide the necessary support for students to familiarize themselves with the solutions strategies after a week's delay and reduce any fadeout effects. This *double-dose* of visual support after time has passed may enhance the short video lesson and solidify learning gains.

Finally, future research should examine the impacts of imposed performance pressure, compared to authentic academic pressure, and also relative to pressure and anxiety that students experience outside of the classroom. It is necessary to explicitly examine how pressure/anxiety impacts learning across these different contexts or to determine if the mechanisms of impact differ based on the type and source of the pressure and anxiety. It is still unclear whether authentic academic pressure and anxiety impacts *learning* and whether the mechanisms are the same for authentic pressure compared to imposed pressure. Additionally, it is crucial to understand the mechanisms of how pressure and anxiety from sources outside of the classroom effect learning potential. These effects are especially relevant this year during the global Coronavirus pandemic. If anxiety and pressure consume cognitive resources that are necessary when learning complex mathematics material (Eysenck, Payne, and Derakshan, 2005) we would hypothesize that the mechanisms of pressure on learning are the same across these three contexts. Alternatively, the type and source of pressure may effect learning differently. More research is necessary to explicitly examine how various forms of pressure and their valence impact learning. My current work is examining just that. Specifically, I am exploring how feelings of anxiety from sources outside of the classroom (anxiety from Coronavirus pandemic) may systemically influence students' attention, learning, and performance inside the classroom. Additionally, I am examining if and how worked examples have the potential to ameliorate the negative effects of pressure by reducing the requirements for cognitive resources.

5.4 Conclusion

This thesis examined worked examples as a highly implementable tool to change the way students' think about and solve proportional reasoning word problems. The aim of this work was to determine the most effective form of worked examples and how to best implement them as a tool for introducing new solution strategies in the middle school mathematics classroom. Importantly, this work moves beyond accuracy and examines how worked examples impact the nature and sequence of the problem-solving process and the ways in which students' learnings scaffold students to fundamentally change their thinking and promote novel but more sophisticated problem-solving approaches with success. Characterization of the approach, setup, procedures, and final solution accuracy was utilized to examine problem-solving gains. Across the three studies, subgroups of students' problem-solving steps were identified, indicating heterogeneity in students' problem-solving abilities. Examination of intervention conditions and individual difference factors predicted these subgroups. The key findings suggest that worked examples not only scaffold students' learning of problem-solving steps, but that worked examples encourage and fundamentally change how students approach problem-solving. This supports the research on worked examples (see Renkl, 2014) but crucially adds to the literature and practical importance of worked examples as a key tool for introducing novel problem-solving approaches, both on their own and paired with classroom instruction. Finally, classroom contexts that incorporate high visual support have the potential to ameliorate the negative implications of pressure and its associated anxiety and for individuals with lower cognitive capacity during conceptually complex thinking (i.e., higher order thinking).

Overall, the research demonstrated the effects of worked examples as a pedagogical tool to introduce novel yet highly sophisticated solution strategies. The research integrated a range of theoretical perspectives and was a novel integration of the literatures on mathematics education instruction, relational reasoning, and cognitive science. The problem-solving coding scheme and novel analytic approach, latent class analysis, allowed for a unique but highly informative analysis of the impacts of worked examples as a tool to introduce complex and sophisticated solution strategies to students. The research suggests that for a more complete characterization of learning, future research should adopt an approach which examines not only problem-solving accuracy but the approaches and strategies students use to solve problems. Finally, this research highlights that executing problem-solving strategies requires distinct cognitive mechanisms for higher order thinking, including comparing and mapping between problems or solutions.

SUPPLEMENTAL MATERIALS

Section S1: Study 2

S 1.1 Multiple imputation procedure

To impute missing data, we used the mice (Multivariate Imputation by Chained Equations) package in R (van Buuren & Groothuis-Oudshoorn, 2011). The method uses predictive mean matching (Little, 1988; Rubin, 1986). Predictive mean matching calculates the predicted value of each variable in turn on the basis of all other variables. For each missing data point, the 5 complete cases whose predicted values are closest to the predicted value of the missing data point are selected as candidate 'donors'. One of these donors is then selected at random, and the observed value for the donor case is used to replace the missing data point (van Buuren, 2018). One complete data set was used to assess the model.

S 1.2 Accuracy Scores by Grade with Imputed Data

Table S 1.1 shows binary regression models for performance accuracy broken down by grade for the equivalent fraction and unit ratio problems. Model 1 and 2 predict equivalent fraction accuracy. Model 3 and 4 predict unit ratio accuracy. In Model 1 and 3 we examined accuracy for 5th grade students controlling for classroom fixed effects. in Model 2 and 4 we examined accuracy for 6th grade students controlling from classroom fixed effects. Across all of these models, condition was not a significant predictor nor were individual difference characteristics. There was suggestive evidence that working memory and arithmetic skills were trending towards significant for 5th grade students but not for 6th grade students for the equivalent fraction problem. There was suggestive evidence that arithmetic skills predict accuracy scores for the unit ratio problem for 5th grade students.

Table S 1.1

Equivalent Fraction and Unit Ratio Model Specifications for Accuracy with Imputed Data, by Grade

	Equivalent Fraction Strategy Accuracy							
	5th Grade				6th Grade			
	Model 1				Model 2			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.29	.52	-.55	.58	.25	.68	-.37	.71
Attentional Control	.00	.00	.52	.60	.00	.01	-.16	.87
Working Memory	*-.21*	*.12*	*-1.77*	*.08*	-.18	.14	-1.24	.22
Knowledge of Proportions	.10	.24	.42	.68	.23	.27	.84	.40
Arithmetic Skills	*.22*	*.13*	*1.70*	*.09*	-.04	.13	-.35	.73

	Unit Ratio Strategy Accuracy							
	5th Grade				6th Grade			
	Model 3				Model 4			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.29	.39	-.74	.46	.18	.61	-.30	.76
Attentional Control	.00	.00	1.30	.19	.00	.00	.55	.58
Working Memory	-.13	.09	-1.43	.15	-.16	.13	-1.18	.24
Knowledge of Proportions	.21	.18	1.21	.23	.20	.24	.82	.41
Arithmetic Skills	**.28**	**.10**	**2.84**	**.00**	.01	.11	.05	.96

Note. High visual support (worked example) condition is compared to the low visual support condition (no worked example, problem only). Covariates are standardized.

S 1. 3 Accuracy Scores with Complete Cases Only

The results for Study 2 using standardized data with only complete cases are reported in Table S 1.2. The model estimates are very similar to those we obtain when using imputation. One key difference emerges for the equivalent fraction problem accuracy regression: the effect of grade is no longer significant at $p < .05$ in the overall model. One key difference emerges for the unit ratio problem-solving accuracy regressions: arithmetic skills in the overall model and 5th grade model is no longer significant, it is trending toward significant $p < 0.07$. Since using only complete cases reduces the power of our analyses, these differences are not entirely unexpected, and we choose to report the main results from our more highly powered imputation-based analysis. However, it is reassuring that the results are broadly similar. Data used for these models is as following: in total 44 cases were deleted due to missingness: 29 cases for the 5th grade models and 15 cases for the 6th grade models.

Table S 1.2

Equivalent Fraction and Unit Ratio Model Specifications for Accuracy using Only Complete Cases

	Equivalent Fraction Strategy Accuracy							
	Total Sample							
	Model 1				Model 2			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.31	.38	-.81	.40	.25	.52	.48	.63
Attentional Control	-	-	-	-	.00	.00	.43	.67
Working Memory	-	-	-	-	-.16	.11	-1.49	.14
Knowledge of Proportions	-	-	-	-	.23	.22	1.04	.30
Arithmetic Skills	-	-	-	-	.06	.12	.53	.60
Grade	-	-	-	-	.66	.56	1.17	.24
	5th Grade				6th Grade			
	Model 3				Model 4			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.71	.71	1.00	.32	.61	.97	-.63	.53
Attentional Control	.13	.44	.30	.76	.16	.52	.30	.76
Working Memory	.06	.40	.16	.88	-.74	.48	-1.53	.13
Knowledge of Proportions	.19	.40	.48	.63	.51	.56	.92	.36
Arithmetic Skills	.33	.49	.68	.50	-.49	.52	-.95	.34

	Unit Ratio Strategy Accuracy							
	Total Sample							
	Model 5				Model 6			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.20	.30	-.68	.50	.04	.40	-.10	.92
Attentional Control	-	-	-	-	.00	.00	1.52	.13
Working Memory	-	-	-	-	-.12	.09	-1.43	.15
Knowledge of Proportions	-	-	-	-	.26	.18	1.44	.15
Arithmetic Skills	-	-	-	-	*.17*	*.09*	*1.80*	*.07*
Grade	-	-	-	-	-.12	.47	-.26	.80
	5th Grade				6th Grade			
	Model 7				Model 8			
	b (logits)	*SE*	Wald z	*p*	b (logits)	*SE*	Wald z	*p*
High Visual Support	.04	.55	-.07	.94	.50	.91	-.55	.58
Attentional Control	.14	.32	.45	.66	.51	.49	1.06	.29
Working Memory	-.09	.31	-.29	.77	-.09	.46	-.20	.84
Knowledge of Proportions	.36	.33	1.08	.28	-.26	.49	-.52	.60
Arithmetic Skills	*.73*	*.41*	*1.78*	*.07*	.19	.48	.39	.70

Note. High visual support (worked example) condition is compared to the low visual support condition (no worked example, problem only). Covariates are standardized. Model 1, 2, 5 and 6 use the full sample. Models are then broken down by grade (i.e., Models 3, 4, 7, and 8) and control for classroom fixed effects.

Section S 2: Study 3

S 2.1: Higher Order Thinking During Lesson

An alternative analysis was conducted using regression to control for individual differences. ANOVA models in the main text use only complete cases because missing data did not exist at the condition level. Missing data exists for covariates therefore these models include imputed data from mice package in R. The same findings exist. There is a main effect of pressure and a trending effect of the interaction between pressure and visual support on higher order thinking, controlling for fixed effects of classroom (see Table S 2.1, Model 1). When individual difference covariates are added to the model, there is no longer a trending effect of the interaction term (see Table S 2.1, Model 2). There is also suggestive evidence that females have more instances of higher order thinking than males and that arithmetic skills are positively related to higher order thinking during the lesson.

Table S 2.1

Model Specifications for Higher Order Thinking with Imputed Data

	Total Sample							
	Model 1				Model 2			
	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Pressure	**-1.50**	**.45**	**-3.36**	**.00**	**-1.31**	**.45**	**-2.93**	**.00**
Visual Support	-.55	.44	-1.24	.21	-.32	.45	-.72	.47
Pressure*Visual Support	*1.18*	*.63*	*1.89*	*.06*	.95	.64	1.49	.14
Female	-	-	-	-	**1.04**	**.38**	**2.70**	**.01**
Attentional Control	-	-	-	-	.00	.00	-1.00	.35
Working Memory	-	-	-	-	.11	.08	1.43	.17
Arithmetic Skills	-	-	-	-	**.25**	**.07**	**3.72**	**.00**
Mathematics Anxiety	-	-	-	-	-.01	.02	-.48	.63
Grade	-	-	-	-	-.14	.38	-.38	.71

Note. Models 1 control for classroom fixed effects. Missing data is imputed with predictive mean matching technique.

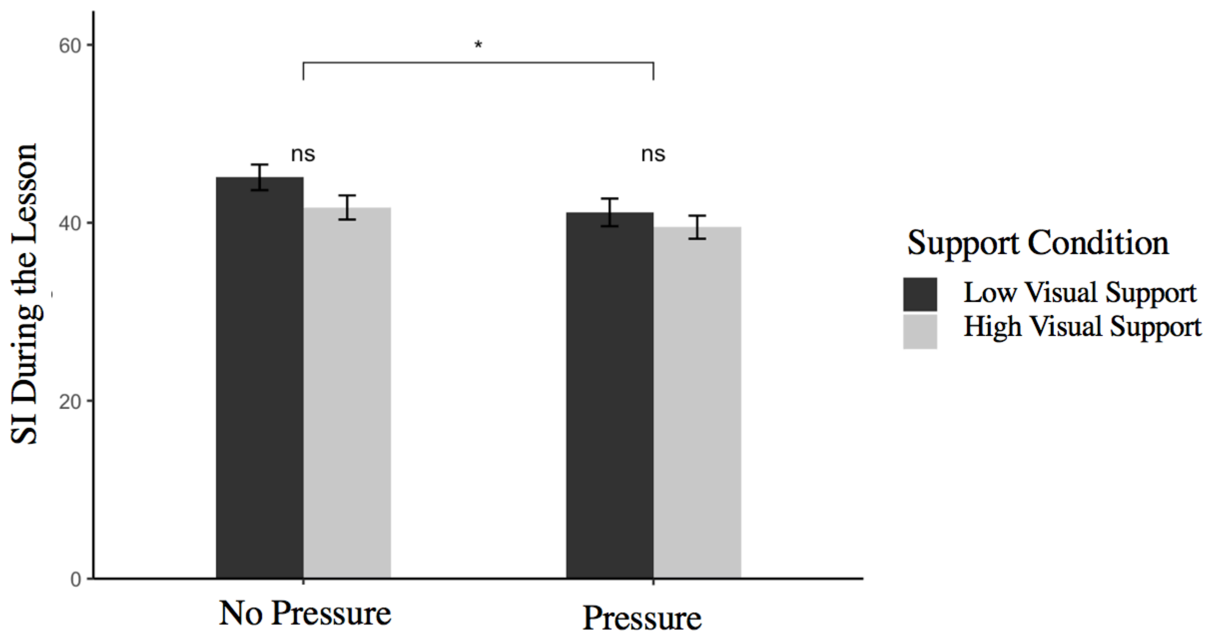
S 2.2: Situational Interest During Lesson with Imputed Data

To examine how students' engagement and interest are effected by visual support, we examined students situational interest (Chen et al., 1999) during learning for those who were not under any forms of imposed pressure. A one-way ANOVA for individuals who did not receive pressure yielded no significant effects of visual support on situational interest during the lesson $F(1,181) = 3.05$, $p = 0.08$, controlling for fixed effects of classrooms.

Our next key question was to examine the possibility that pressure and its added anxiety interacts with visual support and influences students' situational interest. To determine the main effects of pressure and visual support on situational interest, we conducted a two-way analysis of variance with an interaction effect, controlling for classroom fixed effects. The two-way ANOVA yielded a main effect for the pressure condition $F(1,369) = 4.89$, $p = 0.03$, such that students in the no pressure condition had higher reported ratings of situational interest ($M = 43.33$, $SD = 13.96$) than for those in the pressure condition ($M = 40.33$, $SD = 13.94$; see figure S 2.2.1). The main effect of visual support revealed a non-significant trend $F(1,369) = 3.45$, $p = 0.06$, and the interaction was not significant $F(1,369) = 0.29$, $p = 0.59$. A post hoc Tukey HSD (honest significant difference) test correcting for p-values shows that for students who received pressure and high visual support had higher ratings of situational interest than students who were in the no pressure condition and had low visual support, $p < 0.05$. This suggests that pressure paired with visual acted as motivator to enhance interest and engagement. See Supplementary Materials (S 2.2 and Table S 2.2) for alternative analysis using linear regression controlling for individual difference factors.

Figure S 2.1

Pressure Effects Students' Situational Interest During Lesson



Note. Error bars signify standard errors. ** $p < 0.01$. Situational interest scores during the lesson scores ranged from 0-64. Note the x-axis differs from the figure on higher order thinking.

An alternative analysis was conducted using regression to control for individual differences. ANOVA models in the main text use only complete cases. Missing data exists for covariates therefore these models include imputed data from mice package in R. With imputed data, there were no main effects of pressure nor visual support on situational interest (Table S 2.2.2, Model 1 and 2). When we add in the theoretically relevant covariates, there is still no significant effect of condition nor an effect of individual difference factors on students' situational interest (Table S 2.2.2, Model 2).

Table S 2.2
Model Specifications for Situational Interest with Imputed Data

	Total Sample							
	Model 1				Model 2			
	β	*SE*	*t*	Sig.	β	*SE*	*t*	Sig.
Pressure	-2.41	2.06	-1.17	.24	-2.69	2.11	-1.27	.20
Visual Support	-1.45	2.24	-.65	.52	-1.53	2.36	-.65	.52
Pressure*Visual Support	-.44	3.29	-.13	.89	-.41	3.39	-.12	.90
Female	-	-	-	-	.23	1.71	.14	.89
Attentional Control	-	-	-	-	.00	.02	.02	.98
Working Memory	-	-	-	-	.18	.43	.42	.68
Arithmetic Skills	-	-	-	-	-.11	.32	-.34	.73
Mathematics Anxiety	-	-	-	-	-.06	.09	-.62	.54
Grade	-	-	-	-	-3.77	3.14	-1.20	.27

Note. Models 1 control for classroom fixed effects. Missing data is imputed with predictive mean matching technique. Models 1 and 2 use the full sample of data.

Findings suggest a negative relationship between worry/anxiety and situational interest. Learner motivation is a key component to learning (Resnick & Klopfer, 1989) and interest has been a powerful and effective in engaging student learning (Dewey, 1913). We found that more anxious students have less interest in the lesson although these effects do not influence students' learning gains. Increased levels of anxiety and therefore decreases in situational interest are however important to consider as sustained or prolonged levels of anxiety during mathematics could have systematic, deleterious long-term effects on engagement with STEM (e.g., courses, careers).

S 2.3: Post-Lesson Learning Gains with Complete Cases Only

An alternative analysis to examine post-lesson gains in performance accuracy was conducted using linear regression. This model uses only complete cases. Model outcomes from imputed data and only complete cases were nearly identical. In the imputed cases models, working memory was trending toward significance, when the same analyses were run with only complete cases, this individual difference was significant (Table S 2.3, Model 2 and 3). In Model 1, 19 observations were due to missingness. In Model 2, 148 observations were due to missingness. In Model 3, 179 observations were due to missingness.

Table S 2.3

Model Specifications for Post-Lesson Accuracy Gains using Only Complete Cases

	Model 1				Model 2				Model 3			
	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Pressure	-.09	.18	-0.48	.63	-.08	.21	-.36	.72	-.08	.23	-.37	.71
Visual Support	-.19	.18	-1.07	.28	-.17	.21	-.79	.43	-.22	.22	-1.01	.32
Pressure*Visual Support	.18	.26	.69	.49	.19	.30	.64	.53	.11	.31	.35	.73
Female	-	-	-	-	*.27*	*.15*	*1.84*	*.07*	.20	.16	1.26	.21
Attentional Control	-	-	-	-	-.04	.08	-.47	.64	-.02	.08	-.21	.84
Working Memory	-	-	-	-	**.22**	**.08**	**2.82**	**.01**	**.22**	**.08**	**2.73**	**.01**
Arithmetic Skills	-	-	-	-	**.25**	**.08**	**3.03**	**.00**	**.24**	**.09**	**2.76**	**.01**
Mathematics Anxiety	-	-	-	-	*-.14*	*.08*	*-1.85*	*.07*	*-.14*	*.08*	*-1.83*	*.07*
Grade	-	-	-	-	.21	.18	1.20	.23	.11	.19	.55	.58
HOT During Lesson	-	-	-	-	-	-	-	-	-.01	.09	-.16	.87
Situational Interest	-	-	-	-	-	-	-	-	-.01	.08	-.11	.91

Note. Pressure and visual support conditions here signify an addition of the condition manipulation. Data is standardized.

S 2.4: Sustained Learning Gains (Pretest to Follow-up Test) with Complete Cases Only

An alternative analysis to examine post-lesson gains in performance accuracy was conducted using linear regression. This model uses only complete cases. Model outcomes from imputed data and only complete cases were identical (see Table S 2.4). In Model 1, 33 observations were due to missingness. In Model 2, 142 observations were due to missingness. In Model 3, 178 observations were due to missingness.

Table S 2.4

Model Specifications for Sustained Accuracy Gains using Only Complete Cases

	Total Sample				Total Sample				Total Sample			
	Model 1				Model 2				Model 2			
	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Pressure	-.04	.19	-.19	.85	-.16	.23	-.71	.48	-.18	.25	-.73	.46
Visual Support	-.29	.19	-1.51	.13	-.23	.23	-.98	.33	-.33	.25	-1.33	.19
Pressure*Visual Support	.34	.26	1.28	.20	.34	.32	1.07	.28	.31	.34	.91	.37
Female	-	-	-	-	.10	.16	.61	.54	.04	.17	.23	.82
Attentional Control	-	-	-	-	.00	.08	-.05	.96	-.01	.09	-.16	.88
Working Memory	-	-	-	-	*.14*	*.08*	*1.69*	*.09*	.14	.09	1.55	.12
Arithmetic Skills	-	-	-	-	**.28**	**.09**	**3.25**	**.00**	**.27**	**.09**	**2.84**	**.00**
Mathematics Anxiety	-	-	-	-	-.04	.08	-.51	.61	-.07	.09	-.81	.42
Grade	-	-	-	-	.02	.19	.12	.91	-.03	.21	-.13	.90
HOT During Lesson									.01	.09	.14	.89
Situational Interest									-.06	.09	-.63	.53

Note. Pressure and visual support conditions here signify an addition of the condition manipulation. Data is standardized.

REFERENCES

Ackerman, P. L., & Humphreys, L. G. (1990). Individual differences theory in industrial and organizational psychology. *Handbook of industrial and organizational psychology*, *1*, 223-282.

Anghileri, J. (2006). Scaffolding practices that enhance mathematics learning. *Journal of Mathematics Teacher Education*, *9*(1), 33-52.

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87-113.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of experimental psychology: General*, *130*(2), 224.

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

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of educational psychology*, *95*(4), 774.

Atkinson, R. K., & Renkl, A. (2007). Interactive example-based learning environments: Using interactive elements to encourage effective processing of worked examples. *Educational Psychology Review*, *19*(3), 375-386.

Attali, Y., Neeman, Z., & Schlosser, A. (2011). Rise to the challenge or not give a damn: differential performance in high vs. low stakes tests.

Bachman, H. J., Votruba-Drzal, E., El Nokali, N. E., & Castle Heatly, M. (2015). Opportunities for learning math in elementary school: Implications for SES disparities in procedural and conceptual math skills. *American Educational Research Journal*, *52*(5), 894-923.

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological methodology*, *43*(1), 272-311.

Baars, M., Visser, S., Van Gog, T., de Bruin, A., & Paas, F. (2013). Completion of partially worked-out examples as a generation strategy for improving monitoring accuracy. *Contemporary Educational Psychology*, *38*(4), 395-406.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology*, *46*(3), 610.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure?. *Journal of experimental psychology: General*, *130*(4), 701.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem-solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 983.

Bauwens, J., & Hourcade, J. J. (1992). School-based sources of stress among elementary and secondary at-risk students. *The School Counselor*, *40*(2), 97-102.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Bentley, B., & Yates, G. C. (2017). Facilitating proportional reasoning through worked examples: Two classroom-based experiments. *Cogent Education*, *4*(1), 1297213.

Berends, we. E., & van Lieshout, E. C. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction*, *19*(4), 345-353.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, *2*(59-68).

Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York: McKay*, *20*, 24.

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political analysis*, 3-27.

Booth, J. L., McGinn, K. M., Young, L. K., & Barbieri, C. (2015). Simple practice doesn't always make perfect: Evidence from the worked example effect. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 24-32.

Bourne, L. E., Goldstein, S., & Link, W. E. (1964). Concept learning as a function of availability of previously presented information. *Journal of Experimental Psychology*, *67*(5), 439.

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: where young children go wrong. *Developmental psychology*, *44*(5), 1478.

Bray, B. C., Lanza, S. T., & Collins, L. M. (2010). Modeling relations among discrete developmental processes: A general approach to associative latent transition analysis. *Structural Equation Modeling*, *17*(4), 541-569.

Briars, D. J., & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. *Cognition and instruction*, *1*(3), 245-296.

Brickenkam, R., & Zillmer, E. (1998). d2-Test of Attention (1st US edition ed.).(D. Emmans, Trans.) Cambridge/Toronto.

Broza, O., & Kolikant, Y. B. D. (2015). Contingent teaching to low-achieving students in mathematics: Challenges and potential for scaffolding meaningful learning. *ZDM*, *47*(7), 1093-1105.

Bruner, J. S., & Austin, G. A. (1986). *A study of thinking*. Transaction publishers.

Bruner, J. S., Goodnow, J. J., & George, A. (1956). Austin. A study of thinking. *New York: John Wiley & Sons, Inc*, *14*, 330.

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

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

Catrambone, R., & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction*, *16*(2), 139-153.

Clark, R. C., Nguyen, F., & Sweller, J. (2011). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. John Wiley & Sons.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive science*, *5*(2), 121-152.

Chi, M. T., & Ohlsson, S. (2005). *Complex Declarative Learning*. Cambridge University Press.

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of educational psychology*, *79*(4), 347.

Cramer, K., Post, T., & Graeber, A. O. (1993). Connecting research to teaching: Proportional reasoning. *The Mathematics Teacher*, *86*(5), 404-407.

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of experimental child psychology*, *113*(4), 552-568.

Desoete, A., Roeyers, H., & De Clercq, A. (2003). Can offline metacognition enhance mathematical problem-solving?. *Journal of Educational Psychology*, *95*(1), 188.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. *Principles of frontal lobe function*, 466-503.

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135-168.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, *53*(1), 109-132.

Empson, S. B. (1999). Equal sharing and shared meaning: The development of fraction concepts in a first-grade classroom. *Cognition and instruction*, *17*(3), 283-342.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control.

English, L. D., & Halford, G. S. (1995). Mathematics education. *Mahwah, NJ: LEA*.

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & emotion*, *6*(6), 409-434.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, *7*(2), 336.

Fuson, K. C., & Abrahamson, D. (2005). Understanding ratio and proportion as an example of the Apprehending Zone and Conceptual-Phase problem-solving models. *Handbook of mathematical cognition*, 213-234.

Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem-solving: Prior knowledge matters. *Journal of Educational Psychology*, *104*(4), 1094.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child development*, *67*(5), 2022-2044.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child development*, 47-59.

Gentner, D., & Jeziorski, M. (1989). 11. Historical shifts in the use of analogy in Science. *Psychology of science: Contributions to metascience*, 296-325.

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

Gerofsky, S. (2015). Digital mathematical performances: Creating a liminal space for participation. In *Educational Paths to Mathematics* (pp. 201-212). Springer, Cham.

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

Gonzales, P., Guzman, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., & Williams, T. (2003). Highlights from the Trends in International Mathematics and Science Study: TIMSS 2003. Washington, DC: National Center for Education Statistics.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). Highlights from TIMSS 2007: Mathematics and Science Achievement of US Fourth- and Eighth-Grade Students in an International Context. NCES 2009-001. *National Center for Education Statistics*.

Goswami, U. (2001). Analogical reasoning in children. *The analogical mind: Perspectives from cognitive science*, 437-470.

Hailikari, T., Katajavuori, N., & Lindblom-Ylanne, S. (2008). The relevance of prior knowledge in learning and instructional design. *American journal of pharmaceutical education*, *72*(5).

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and brain sciences*, *21*(6), 803-831.

Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of psychology*, *87*(4), 621-636.

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology*, *79*(2), 192-227.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for research in mathematics education*, *21*(1), 33-46.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, *1*(1), 77-89.

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (2018). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences*, *66*, 4-15.

Hiebert, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. DIaNe Publishing.

Hiebert, J., Gallimore, R., & Stigler, J. W. (2003). The new heroes of teaching. *Education Week*, *23*(10), 42-56.

Hilbert, T. S., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction*, *18*(1), 54-65.

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. (2009). A mandate for playful learning in preschool: Applying the scientific evidence.

Holyoak, K. J. (2012). Analogy and relational reasoning.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & cognition*, *15*(4), 332-340.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive science*, *13*(3), 295-355.

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, *40*(3), 366-371.

Jitendra, A. K., Star, J. R., Starosta, K., Leh, J. M., Sood, S., Caskie, G., ... & Mack, T. R. (2009). Improving seventh grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology*, *34*(3), 250-264.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human factors*, *40*(1), 1-17.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem-solving is superior to studying worked examples. *Journal of educational psychology*, *93*(3), 579.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of educational psychology*, *96*(3), 558.

Karplus, R., Pulos, S., & Stage, E. K. (1983). Early adolescents' proportional reasoning on 'rate'problems. *Educational studies in Mathematics*, *14*(3), 219-233.

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in cognitive science*, *2*(2), 285-305.

Kilpatrick, S., & Swafford, J. Findell.(2001). *Adding it up: Helping children learn mathematics*.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*(3), 239-264.

Krawczyk, D. C., Hanten, G., Wilde, E. A., Li, X., Schnelle, K. P., Merkley, T. L., ... & Levin, H. S. (2010). Deficits in analogical reasoning in adolescents with traumatic brain injury. *Frontiers in Human Neuroscience*, *4*, 62.

Lamon, S. J. (1993). Ratio and proportion: Connecting content and children's thinking. *Journal for research in mathematics education*, *24*(1), 41-61.

Lamon, S. J. (2020). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers*. Routledge.

Lesh, R., Post, T. R., & Behr, M. (1988). Proportional reasoning. In *Number concepts and operations in the middle grades* (pp. 93-118). National Council of Teachers of Mathematics, Lawrence Erlbaum Associates.

Levitt, S. D., List, J. A., Neckermann, S., & Sadoff, S. (2016). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy*, *8*(4), 183-219.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into practice*, *32*(3), 131-137.

Lyons, E. M., Simms, N., Begolli, K. N., & Richland, L. E. (2018). Stereotype threat effects on learning from a cognitively demanding mathematics lesson. *Cognitive science*, *42*(2), 678-690.

Maloney, E. A., Ansari, D., & Fugelsang, J. A. (2011). Rapid communication: The effect of mathematics anxiety on the processing of numerical magnitude. *Quarterly Journal of Experimental Psychology*, *64*(1), 10-16.

Maloney, E. A., Schaeffer, M. W., & Beilock, S. L. (2013). Mathematics anxiety and stereotype threat: shared mechanisms, negative consequences and promising interventions. *Research in Mathematics Education*, *15*(2), 115-128.

Maloney, E. A., & Retanal, F. (2020). Higher math anxious people have a lower need for cognition and are less reflective in their thinking. *Acta psychologica*, *202*, 102939.

Maloney, E. A., Sattizahn, J. R., & Beilock, S. L. (2014). Anxiety and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(4), 403-411.

Mayer, R. E. (1999). Designing instruction for constructivist learning. *Instructional-design theories and models: A new paradigm of instructional theory*, *2*, 141-159.

Maxwell, S. E., O'Callaghan, M. F., & Delaney, H. D. (1993). Analysis of covariance.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words?. *Journal of educational psychology*, *82*(4), 715.

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., ... & Shonkoff, J. P. (2017). Impacts of early childhood education on medium-and long-term educational outcomes. *Educational Researcher*, *46*(8), 474-487.

McLaren, B. M., Lim, S. J., Gagnon, F., Yaron, D., & Koedinger, K. R. (2006). Studying the effects of personalized language and worked examples in the context of a web-based intelligent tutor. In *International Conference on Intelligent Tutoring Systems* (pp. 318-328). Springer, Berlin, Heidelberg.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100.

McLaren, P. (2015). *Life in schools: An introduction to critical pedagogy in the foundations of education*. Routledge.

McGinn, K. M., Lange, K. E., & Booth, J. L. (2015). A worked example for creating worked examples. *Mathematics Teaching in the Middle School*, *21*(1), 26-33.

Moore, C. F., Dixon, J. A., & Haines, B. A. (1991). Components of understanding in proportional reasoning: A fuzzy set representation of developmental progressions. *Child Development*, *62*(3), 441-459.

Moreno, R. (2006). When worked examples don't work: Is cognitive load theory at an impasse? *Learning and Instruction*, *16*(2), 170-181.

Morrison, R. G., Doumas, L. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: balancing inhibitory control in working memory and relational representation. *Developmental science*, *14*(3), 516-529.

Morrison, R. G., Holyoak, K. J., & Troung, B. (2001). Working-memory modularity in analogical reasoning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 23, No. 23).

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and instruction*, *16*(2), 173-199.

Nabors, W. K. (2003). From fractions to proportional reasoning: a cognitive schemes of operation approach. *The Journal of Mathematical Behavior*, *22*(2), 133-179.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, *131*(1), 5.

National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: The Counc

National Governors Association. (2010). Common core state standards. *Washington, DC*.

Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education*, *28*(4), 1237-1263.

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

Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems?. *Journal of educational psychology*, *77*(3), 272.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. *National Center for Education Research*.

Parish, L. (2010). Facilitating the Development of Proportional Reasoning through Teaching Ratio. *Mathematics Education Research Group of Australasia*.

Palm, T. (2008). Impact of authenticity on sense making in word problem-solving. *Educational studies in Mathematics*, *67*(1), 37-58.

Pass, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 85*(4), 429-434.

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and motor skills*, *79*(1), 419-430.

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of experimental child psychology*, *141*, 83-100.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, *14*(2), 187-202.

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive science*, *38*(1), 1-37.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem-solving in cognitive skill acquisition: A cognitive load perspective. *Educational psychologist*, *38*(1), 15-22.

Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading worked solution steps works–a cognitive load perspective. *Instructional science*, *32*(1), 59-82.

Retnowati, E., Ayres, P., & Sweller, J. (2010). Worked example effects in individual and group work settings. *Educational Psychology*, *30*(3), 349-367.

Richland, L. E., & Begolli, K. N. (2016). Analogy and higher order thinking: Learning mathematics as an example. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 160-168.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science-New York then Washington-*, *316*(5828), 1128.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 177-192.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: psychometric data. *Journal of counseling Psychology*, *19*(6), 551.

Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological science*, *24*(7), 1301-1308.

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.

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

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing procedures?. *British Journal of Educational Psychology*, *82*(3), 436-455.

Ross, R. M. (2005). *The d2 test of attention: An examination of age, gender, and cross-cultural indices*. Argosy University.

Salden, R. J., Koedinger, K. R., Renkl, A., Aleven, V., & McLaren, B. M. (2010). Accounting for beneficial effects of worked examples in tutored problem-solving. *Educational Psychology Review*, *22*(4), 379-392.

Sapolsky, R. M. (2015). Stress and the brain: individual variability and the inverted-U. *Nature neuroscience*, *18*(10), 1344-1346.

Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broadsided by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behavior*, *17*(6), 639-655.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and instruction*, *16*(4), 475-5223.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and instruction*, *22*(2), 129-184.

Schwonke, R., Renkl, A., Krieg, C., Wittwer, J., Aleven, V., & Salden, R. (2009). The worked-example effect: Not an artefact of lousy control conditions. *Computers in human behavior*, *25*(2), 258-266.

Siegler, R. S., & Vago, S. (1978). The development of a proportionality concept: Judging relative fullness. *Journal of Experimental Child Psychology*, *25*(3), 371-395.

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in cognitive sciences*, *17*(1), 13-19.

Sowder, J. T. (2007). The mathematical education and development of teachers. *Second handbook of research on mathematics teaching and learning*, *1*(157-223).

Steele, C. M. (1998). Stereotyping and its threat are real.

Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psychological review*, *84*(4), 353.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational leadership*, *61*(5), 12-17.

Sullivan, P., Tobias, S., & McDonough, A. (2006). Perhaps the decision of some students not to engage in learning mathematics in school is deliberate. *Educational Studies in Mathematics*, *62*(1), 81-99.

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., Ayres, P. L., Kalyuga, S., & Chandler, P. (2003). The expertise reversal effect.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296.

Trafton, J. G., & Reiser, B. J. (1993). Studying examples and solving problems: Contributions to skill acquisition. In *Proceedings of the 15th conference of the Cognitive Science Society* (pp. 1017-1022).

Trezise, K., & Reeve, R. A. (2018). Patterns of anxiety in algebraic problem-solving: A three-step latent variable analysis. *Learning and Individual Differences*, *66*, 78-91.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of educational psychology*, *91*(2), 334.

van Loon-Hillen, N., Van Gog, T., & Brand-Gruwel, S. (2012). Effects of worked examples in a primary school mathematics curriculum. *Interactive Learning Environments*, *20*(1), 89-99.

Van Merrienboer, J. J., & Krammer, H. P. (1990). The" completion strategy" in programming instruction: Theoretical and empirical support. *Research on instruction: Design and effects*, 45-61.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, 450-469.

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, *24*(3), 335-359.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition*, *28*(7), 1205-1212.

Watson, A., & Mason, J. (2006). Seeing an exercise as a single mathematical object: Using variation to structure sense-making. *Mathematical thinking and learning*, *8*(2), 91-111.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*(7), 352-360.

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

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of child psychology and psychiatry*, *17*(2), 89-100.

Wood, T., Nelson, B. S., & Warfield, J. E. (Eds.). (2014). *Beyond classical pedagogy: Teaching elementary school mathematics*. Routledge.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Punishment: Issues and experiments*, 27-41.

Yeung, A. S., Jin, P., & Sweller, J. (1997). Cognitive load and effects in reading with explanatory notes. *Contemporary Ed ceps*.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and instruction*, *4*(3), 137-166.