Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

Let's be rational: worked examples supplemented textbooks improve conceptual and fraction knowledge

Christina Areizaga Barbieri[a], Julie L. Booth[b], and Kamal Chawla[a]

[a]College of Education and Human Development, University of Delaware, Newark, DE, USA;
[b]College of Education and Human Development, Temple University, Philadelphia, PA, USA

Contact
Christina Areizaga Barbieri, barbieri@udel.edu
College of Education and Human Development
University of Delaware, 113 Willard Hall Education Building, Newark, DE 19716, USA

© 2022 Informa UK Limited, trading as Taylor & Francis Group

Abstract
The current study assessed whether adding worked examples with self-explanation prompts focused on making connections between mathematical principles, procedures, and concepts of rational numbers to a curriculum focused on invented strategies improves pre-algebra students' fraction number line acuity, rational number concepts and procedures. Finally, the study assessed whether individual differences in prior knowledge of rational numbers moderate the effect of the textbooks. The experimental textbooks resulted in significantly greater improve-ments overall in fraction number line estimation skills than con-trol textbooks. Further, example-supplemented textbooks were more effective than control textbooks at improving understanding of rational number concepts for those with low prior conceptual knowledge. Both textbooks were equally effective at improving procedural skills.

Keywords
Rational number; numerical cognition; fraction magnitude; numerical magnitude; mathemat-ics education

Worked examples paired with self-explanation prompts improve learning, especially for students with low prior knowledge (Barbieri & Booth, 2020). The current study aims to develop rational number knowledge by supplementing textbooks with examples and self-explanation prompts. US teachers report frequent use of textbooks in their classrooms (Fulkerson, 2013). There is variability in this use, yet most report using them to guide the structure and content emphasis of the unit. One feasible way to implement worked examples is to embed them into textbooks.

Strong foundational knowledge of fractions is necessary for algebra success (NMAP, 2008) and predictive of mathematics achievement (Bailey et al., 2012). Understanding fractions as a numerical magnitude on a number line is especially important for later achievement (Resnick et al., 2016). Those who enter middle school lacking proficiency with fractions are at risk for mathematics failure (Mazzocco & Devlin, 2008). Students who leave sixth grade with weak fraction knowledge experience cascading math difficulties throughout middle school when algebra becomes a primary topic (Siegler et al., 2012). Resnick et al. (2016) demonstrated that 42% of their sample showed

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

minimal growth in fraction number line estimation acuity between the fourth and sixth grades. Thus, designing effective classroom materials targeting these skills is crucial.

We assess whether adding worked examples with self-explanation prompts to a research-based curriculum improves pre-algebra students' fraction number line acuity. We also aim to replicate prior research demonstrating that worked examples improve knowledge of mathematical concepts and procedures and extend this to rational number concepts. Finally, we examine whether individual differences in prior rational number knowledge moderate the effects of worked examples with self-explanation prompts as found in other target areas such as algebra (Barbieri & Booth, 2016).

The worked examples effect

Worked examples are worked-out problem solutions that learners study when practicing new material. They are often presented as being completed by a fictional student and paired with a practice problem similar in structure. Worked example-problem pairs are used in place of problems to be solved. The worked examples effect (Sweller & Cooper, 1985) is derived from cognitive load theory (CLT; Sweller, 1988). CLT was inspired by the lack of evidence that simple problem-solving practice fosters learning. Sweller suggests that humans' cognitive capacity is limited in what it can process and traditional problem-solving practice (i.e. the rote practice of a to-be-learned procedure) overloads that capacity, leaving little room for acquisition and development of new knowledge structures. Some distinctions are made between the types of load that learning tasks should manage (Paas et al., 2004). These types are additive and include intrinsic and extraneous load (Sweller et al., 2011). Intrinsic load is needed to perform the task and is crucial for learning. Extraneous load is *not* necessary for learning and should be reduced in whatever ways possible. Extraneous load is due to aspects of the instructional design, learner characteristics that can deter learning or make learning more difficult (e.g. distracting thoughts about learning or failure), or features of the learning environment (e.g. distracting information; Paas & van Merriënboer, 2020). Tasks should be designed in a way that reduces *extraneous* load or task demands that are irrelevant to schema acquisition (i.e. presentation formats that make learning difficult) to allow remaining cognitive resources to be dedicated to handling intrinsic load. This is particularly important in domains that are complex and highly interactive, such as mathematics (i.e. domains with high *intrinsic* load). Studying worked examples, or worked-out problem solutions, frees up learners' cognitive resources normally present in traditional problem-solving in that they are not required to generate random processes to solve a problem or attend to irrelevant features. Rather, they can devote effort to only the relevant aspects for solving the problem and developing schemata for the problem-type (Sweller, 2006). After studying multiple worked examples, the student can extract the features of or rules for a problem that are common to other problems of that type, leading to automation of problem-solving.

Recently, as work in mathematical cognition and related domains has emphasized the reciprocity of the development of procedural and conceptual skills (Rittle-Johnson, 2017), researchers have begun to design worked examples in ways that attend to

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

both conceptual and procedural features of problem-solving (Barbieri & Booth, 2020). This can be accomplished by supplementing examples with self-explanation prompts that encourage learners to make explicit connections between concepts and procedures (Berthold et al., 2009). Prompts ask students to explain how and why something occurred or to explain a strategy which reduces errors (Siegler & Chen, 2008) and boosts conceptual algebra knowledge (Booth et al., 2013). In the current study, we pair examples with self-explanation prompts to encourage the integration of new and prior knowledge.

Research on worked examples has revealed an *expertise-reversal effect* or the reversal of cognitive load effects after learners develop expertise. Instruction that is effective with inexperienced learners or learners with low prior knowledge is particularly *ineffective* with experienced learners (Kalyuga et al., 2003, for a review). Recent work points to the need to consider the correctness of examples along with learners'prior knowledge or expertise. Worked examples are most commonly presented as correct. However, incorrect and faded worked examples are also used. Incorrect examples draw attention to the underlying principle that makes a solution incorrect, reduce misconceptions (Durkin & Rittle-Johnson, 2012) and promote learning of concepts and procedures (Adams et al., 2014; Booth et al., 2015). Incorrect worked examples have particular benefits for those with low prior knowledge (Barbieri & Booth, 2020).

Faded worked examples, or examples that gradually reduce the number of steps shown so that the student can complete the missing steps independently, improve problem-solving proficiency (Atkinson et al., 2003) and efficiency (Renkl et al., 2004). We use correct, incorrect, and faded worked examples in the current study and do not aim to isolate effects by type.

Struggles with rational numbers

Fractions are introduced in the third grade and decimals in the fourth grade in U.S. Schools (CCSS-M, 2010). Students relate to-be-learned material to prior knowledge. Problems sometimes arise when learners overgeneralize rules and strategies to the new material (Stafylidou & Vosniadou, 2004). When students first expand their knowledge of numbers to include fractions and decimals students incorrectly apply natural number rules to fractions and decimals (DeWolf & Vosniadou, 2015). Difficulties revolve around conceptual components of rational number knowledge such as determining their size as well as procedures such as operating with rational numbers. This is understandable considering the differences between fractions, decimals, and integers. For example, when comparing the relative size of two fractions, the fraction with the larger numerator is not necessarily the larger number (e.g. 3/4 vs. 6/12). With decimals, the length of a number does not indicate its size (e.g. $.100 = .10$). Multiplication does not always lead to a product greater than an operand (e.g. $.2 \times .4 = .08$; $1/2 \times 3/4 = 3/8$) and division does not lead to a quotient smaller than the dividend (e.g. $.5 \div .25 = 2$; $1/6 \div 1/3 = 1/2$). A common error for fraction addition made by sixth graders is to add across numerators and denominators (Newton et al., 2014). Developing and assessing instructional supports appropriate for classroom use that advance fraction knowledge is critical.

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

Overview of the present study

In the present study, we assessed whether students in the early stages of skill acquisition when working on operations with fractions, decimals, and percents showed greater improvement in rational number knowledge after using problem sets in a mathematics textbook supplemented with worked examples than those who used the commercially available textbook without supplementation. The experimental manipulation was aimed at strengthening problem-solving skills through making meaningful connections between rational number concepts and operations, which improves higher-level mathematical thinking (Stein & Lane, 1996). We assessed whole number and fraction number line estimation skills as well as conceptual and procedural knowledge of rational numbers. We define conceptual knowledge as an understanding of general mathematical principles such as magnitude (Rittle-Johnson et al., 2001) or feature knowledge (e.g. the meaning of the equal sign; McNeil, 2014). These are abstract concepts or principles that do not rely on a series of steps to solve a problem, which is characteristic of procedural knowledge. Though they lie on a continuum, we use separate measures of conceptual and procedural knowledge to ensure that students are developing adequate skills on each end. Further, some work confirms that there are two different components of numerical understanding that overlap (Ye et al., 2016). For whole number understanding, these include numerical magnitude and calculation. We also predict fraction number line estimation, commonly thought of as a conceptual skill, separately due to its significance for student learning. Finally, we assess whether individual differences in prior knowledge of rational numbers moderate the effects of textbook type.

To be practically meaningful, worked examples with self-explanation prompts should be effective in classrooms with similar effects as found in more controlled laboratory experiments. Prior research on worked examples has mainly included controlled laboratory and classroom experiments in which students are randomly assigned to complete worked example tasks on paper or within a computerized tutor. Although still effective in the classroom, the worked examples effect is more pronounced in the laboratory (Booth et al., 2016). Considering this, it is unclear whether embedding examples with self-explanation prompts throughout textbooks would yield an effect strong enough to be practically meaningful. One possibility is that embedding worked examples with self-explanation prompts throughout a textbook may yield only negligible effects in less controlled environments. Alternatively, these supplements may work just as well when made a normal part of homework or classwork.

Research questions

The present study has three research questions: (1) Does having students use a textbook revised to replace some practice problems with worked examples paired with self-explanation prompts benefit fraction number line estimation skills? (2) Do these supplemented textbooks improve (a) conceptual and (b) procedural knowledge of fractions, decimals, and percents? (3) Does prior knowledge moderate the effects of the supplemented textbook?

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

Methods

Participants

Middle school students ($N = 248$; 52% female, 48% male) from 6th-grade pre-algebra classrooms of three different teachers in a public school district in the Midwest participated in the current study. The sample included White (55.5%), Black (25.5%), Hispanic/Latinx (11.7%), Asian (1.2%), and multi-ethnic or biracial (6.1%) participants. Participants were spread across 14 classrooms taught by three different teachers. 7.26% of the recruited sample ($n = 18$) did not complete assessments due to absences reducing the final analysis sample to 92.74% ($N = 230$) of its initial size. No differences were found between the attrition and analysis samples on demographic data, suggesting data were missing at random.

Materials

Content knowledge test

Pre- and posttests included 26 items on fractions, ratios, decimals, and percents. Items overlapped target skills but most items focused on decimals (~50%) and percents (~38%). Seven items were procedural in nature and 19 items were conceptual (Figure 1). Scores on each were computed as proportion correct (Table 1). Internal consistency was high for procedural and conceptual measures ($\alpha = .724$, $\alpha = .826$ respectively).

Number line estimation tasks

Whole number line estimation. Two paper-and-pencil number line estimation tasks measured the whole number and fraction magnitude representations (Booth et al., 2014). The whole number magnitude task used a 0-6257 scale on which students were asked to label 12 target numbers on individual number lines (Figure 2). An independent coder measured the distance between 0 and each student's mark for all target numbers. Percent absolute error (PAE, how inaccurate an estimate is) was calculated for each estimate by dividing the absolute difference between the location of the participant's estimate and the accurate target location by the scale of the estimates. PAEs for each target were averaged to determine the overall whole number PAE. This paper-and-pencil task has shown high reliability in prior work (Barbieri et al., 2021). We also report students' R^2_{Lin} for the whole number line task, as it has previously been shown to be important for math achievement (Siegler & Booth, 2004). This measure represents the linearity of students' magnitude estimates. Magnitudes are first represented by a compressed, logarithmic distribution in which children overestimate the magnitude of smaller numbers and underestimate the magnitude of larger numbers. However, they become more accurate in their representations over time resulting in a linear distribution. These transitions from logarithmic to linear representations happen over time with increasing larger scales. For example, a child could have progressed to a linear representation on a smaller numerical range (e.g. 0–100) while still holding a logarithmic representation on a larger scale (e.g. 0–1,000). However, these transitions don't typically apply to fractions. PAE has become increasingly more common as a

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

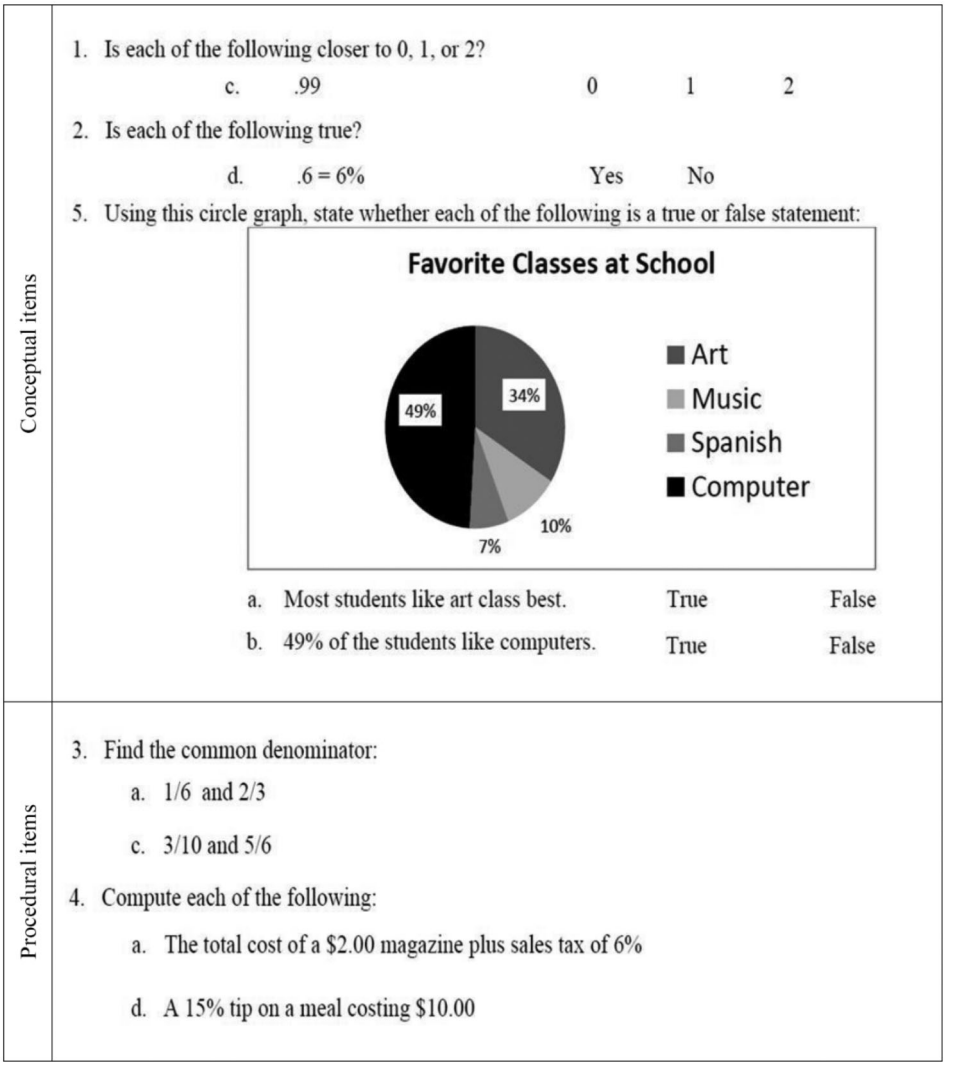

Figure 1. Sample assessment items.

measure of magnitude representation, we use this as our outcome measure for fraction magnitude knowledge.

Fraction number line estimation. The fraction magnitude task used a 0 to 1 scale. Students marked targets on 12 individual number lines.

Textbooks

We used one textbook within the *Connected Mathematics 2 (CMP)* curriculum titled *Bits and Pieces III*, which focuses on adding, subtracting, multiplying, and dividing decimals. Students build on their understanding of fraction operations to develop an understanding of decimal operations and then use this new knowledge to develop strategies to work with percentages. This unit focuses on decimal understanding using

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

Table 1. Equivalence between groups at pre-test on student performance and demographic variables.

	Analysis sample ($N = 230$)		Control ($n = 111$)		Worked examples ($n = 119$)	
	Mean	*SD*	Mean	*SD*	*Mean*	*SD*
Pretest conceptual	.80	.15	.78	.17	.82	.14
Posttest conceptual	.86	.13	.84	.15	.87	.11
Pretest procedural	.50	.30	.48	.27	.52	.31
Posttest procedural	.58	.30	.59	.30	.57	.31
Pretest whole number R^2_{Lin}	.90	.15	.90	.14	.90	.16
Posttest whole number R^2_{Lin}	.90	.14	.91	.10	.88	.16
Pretest whole number PAE	.09	.06	.09	.07	.08	.06
Posttest whole number PAE	.08	.05	.08	.04	.08	.05
Pretest fraction PAE	.09	.08	.09	.08	.09	.09
Posttest fraction PAE	.08	.08	.09	.09	.07	.06
Female	52.6%		48.6%		56.3%	
White	57.2%		50.5%		63.6%	
Black	24.0%		28.8%		19.5%	
Hispanic/Latinx	11.4%		13.5%		9.3%	
Asian	1.3%		0.9%		1.7%	
Multi-ethnic/Biracial	6.1%		6.3%		5.9%	

Note. Scale for all continuous measures is 0-1 representing proportion correct (or proportion of error for PAE).

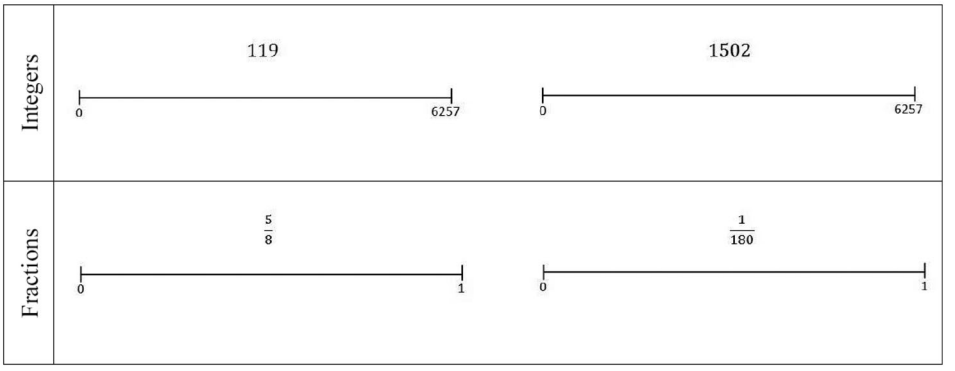

Figure 2. Sample items from number line estimation task.

Note. Whole number targets included 2, 44, 119, 644, 1502, 2165, 2741, 3141, 3836, 4511, 5225, and 6019. Fraction targets included $\frac{5}{8}$, $\frac{1}{180}$, $\frac{1}{3}$, $\frac{23}{25}$, $\frac{4}{29}$, $\frac{3}{10}$, $\frac{1}{5}$, $\frac{177}{352}$, $\frac{3}{5}$, $\frac{1}{12}$, $\frac{146}{149}$, and $\frac{5}{6}$.

the standard place-value system as well as interpreting decimals as fractions (e.g. \$1.97 = 197/100; Figure 3).

The CMP curriculum (Lappan et al., 2006) is a problem-centred curriculum that promotes inquiry-based instruction and aligns with CCSS-M. CMP aims to help students develop mathematical understanding by emphasising connections between mathematical ideas and real-world problem-solving applications. Each textbook includes multi-step problems, termed *Investigations*, and prompts to write out reflections on strategies. CMP lessons begin by ensuring students understand the mathematical nature of the problem. Students invent and share different strategies for finding solutions with teacher guidance. Teachers close the lesson with a summary discussion of strategies presented and connections made. Each *Investigation* ends with ACE (Applications-Connections-Extensions) problems assigned for homework which require estimation, computation, and explanation. ACE problems are similar, but *not* identical,

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

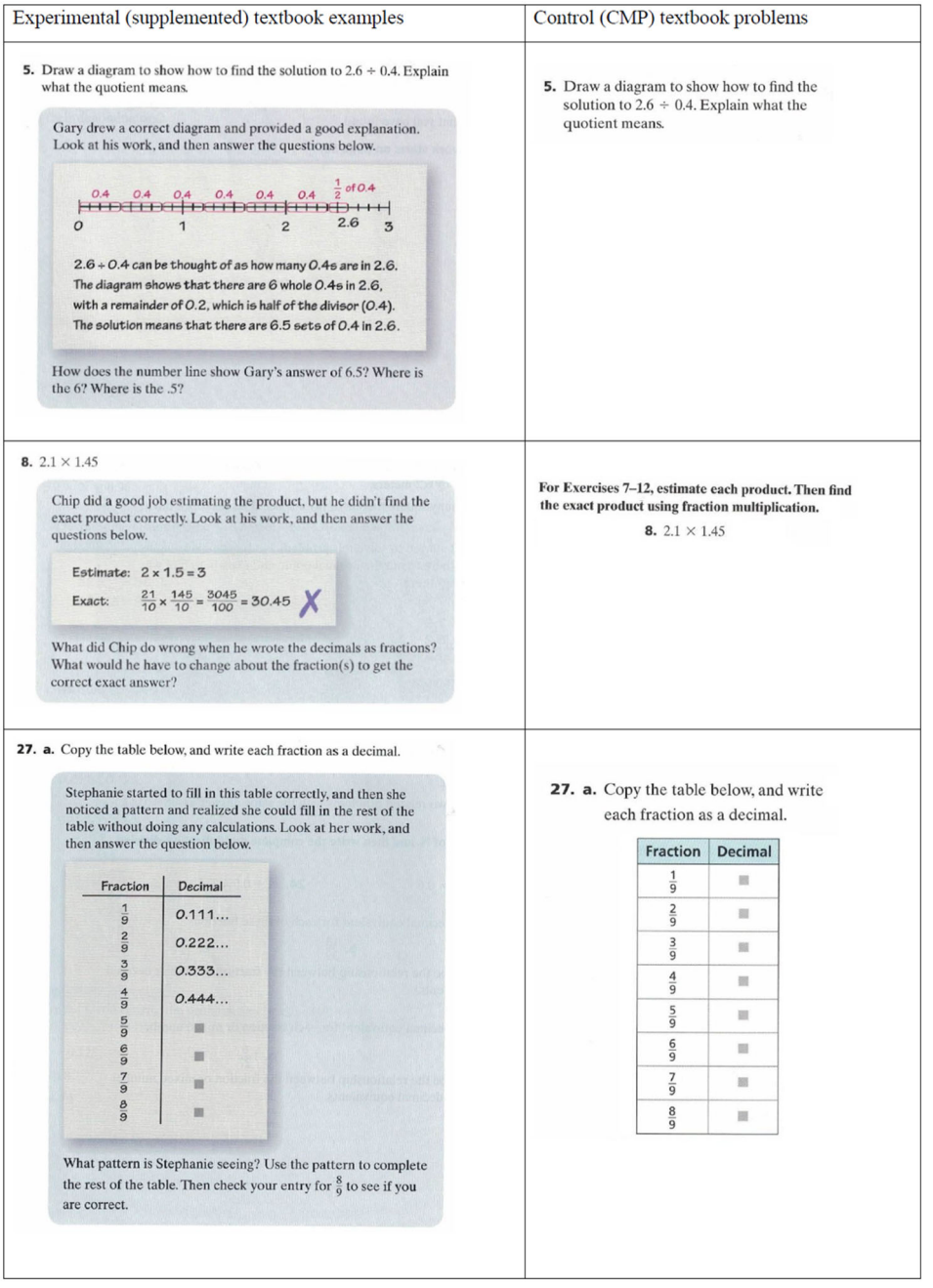

Figure 3. Sample practice problems and worked examples by condition.

to those in the lesson and require students to connect and extend their new knowledge to different problems and contexts.

The standard *Bits and Pieces III* textbook was used as the control textbook. The experimental modification was made solely to ACE homework problems, not to the *Investigations* themselves. ACE problems and examples were used after the lesson

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

ensuring instruction was held constant across conditions. Three types of worked examples replaced a portion of the practice ACE problems: (1) Correct examples, displaying a fictitious student's correct problem-solving, (2) Incorrect examples, displaying a fictitious student's incorrect problem-solving (clearly marked incorrect), and (3) Faded examples, displaying a fictitious student's correct but incomplete problem-solving and students were prompted to complete the problem. Each example was paired with written self-explanation prompts. We do not aim to isolate the effects of example type or self-explanation prompts as has been done in prior research (Barbieri et al., 2019; Barbieri & Booth, 2016). Rather, we leverage the benefits of each example type and explanations to provide students with multiple opportunities to learn. These materials were designed to increase students' opportunities to focus on the connections between procedures and underlying concepts so that both of these understandings improve simultaneously. Each of the examples and corresponding explanation prompts targeted specific concepts and procedures. For example, two of the examples shown in Figure 3 target decimal operations (i.e. division and multiplication), and all three emphasize conceptual connections between fractions and decimals that assist in solving these problems.

Forty-nine worked examples were added to the experimental textbook (replacing existing problems) split across five subunits or *Investigations*. Twenty-eight examples focused solely on decimals. These examples mainly addressed operations with a smaller focus on labeling number lines with decimals and using decimals to solve word problems. Seven examples focused solely on fractions, covering adding and multiplying fractions as well as finding equivalent fractions. Fourteen highlighted connections between fractions and decimals by targeting estimation and computation of decimal operations using fractions, converting fractions to decimals, and dividing decimals by using a number line partitioned in both decimal and fraction formats. Thus, although the unit focused primarily on decimal operations, 43% of the examples added to the experimental textbook included some focus on fractions (Figure 3). The solutions presented in the worked examples were consistent with those encouraged by CMP and CCSS-M.

Most teachers used ACE problems as homework and most classrooms completed this specific unit towards the end of the school year, after operations with fractions. Although there was variation in textbook use by the teacher, this didn't vary across classrooms within teachers. Teachers reported using the different versions of the textbooks in the same way with all of their students. Thus, use did not vary systematically by condition. Students were not made aware that their textbooks differed.

Procedure

All classrooms used the original *Bits and Pieces III* textbook from *CMP 2* for instruction as normally used, but half of the classrooms were randomly assigned to use a worked-example supplemented version of the textbook for homework (or occasional classwork). We assess fraction magnitude understanding, conceptual and procedural knowledge of rational numbers, and whole number line estimation, a precursor to fraction number line estimation (Resnick et al., 2016) and algebra achievement (Booth et al.,

2014). We expected that pre-algebra students' whole number line estimates would be highly accurate (Booth & Siegler, 2008) and thus would not show significant changes from pre- to posttest. Thus, whole number magnitude was primarily used as a control variable representing prior math ability when predicting fraction magnitude understanding.

Classrooms were assigned to use either a standard *CMP* text or the experimental text. Seven classrooms were assigned to the control condition ($n = 111$ students) and seven were assigned to the experimental condition ($n = 119$), stratified by teacher. Thus, participating teachers served as their own control in that they each taught at least one control classroom and at least one experimental classroom. Assignment was at the classroom level (as opposed to student level) to reduce the risk of contamination effects as students needed to use the textbooks over approximately 7–8 weeks. Prior to using the textbooks, participants completed pretest measures. Students then worked with either the control or experimental textbooks primarily as homework. After unit completion, students completed posttests identical to the pretests.

Teachers used the textbooks in their classroom for instruction, and then assigned problems as they normally would, for homework or in-class practice after instruction. The current study uses an intent-to-treat (ITT) framework (Burke-Miller, 2012). ITT analysis requires all participants recruited for the study to be included in predicting outcomes by their assigned and intended condition, regardless of whether they strictly adhered to the assigned protocol or not (e.g. dropped out prior to completion, completed a small portion of the assigned materials, etc.). ITT is particularly useful in less controlled settings, such as classrooms, when strict adherence to the study protocol is not possible. The logic behind ITT is that participants who do not adhere to the assigned condition may differ systematically from those who do. As such, removing them from the analyses invalidates random assignment and introduces bias which can lead to an overestimation of the treatment effect. We use this approach to establish ecological validity and get a truer estimate of the effects in the classroom.

Plan of analysis

We address three research questions: (1) Does using a textbook supplemented with worked examples paired with self-explanation prompts during problem solving practice benefit fraction number line estimation skills? (2) Do the same supplemented textbooks improve (a) conceptual and (b) procedural knowledge of fractions, decimals, and percents? and (3) Does student prior knowledge in these domains moderate the effects of the textbook on these outcomes? To answer research questions 1 and 2, several mixed ANOVAs were conducted. These 2 (Time) $\times$ 2 (Condition) mixed ANOVAs assess differential change in scores over time by condition. Our third research question was answered with multiple regression analyses that assess main effects and moderation to assess the impact of prior knowledge on effects of condition. More specifically, an interaction term was added to each main effects model between a condition dummy code (Example $=$ 1) and the corresponding pretest. A significant interaction signifies that the magnitude of the effect of condition on the predicted posttest significantly varies by pretest. Standard follow-up analyses involve testing

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

simple slopes or running split regressions and testing the predictor of interest (i.e. in this case pre-test) separately by condition along with all other predictors that were in the moderation model except for the dummy code and interaction term.

We did not have access to student-level data on completion rates and so did not control for this. If treatment is indeed effective, ITT analysis provides unbiased estimates of the effect of the treatment at the level of adherence present in the study (McCoy, 2017). Though it is likely that classrooms used varying amounts of their assigned textbook, removing students from the analyses that did not complete a substantial amount of the assignments would introduce bias that can lead to an overestimation of the effect of the textbook. It can be assumed that complete adherence to the study protocol and completion of the full text would have even greater effects (in favor of the outperforming group making ITT a conservative estimate).

Difference tests were conducted to assess baseline balance between conditions. Though classrooms were randomly assigned to experimental or control conditions stratified by teacher, the worked examples condition had higher conceptual scores at pretest ($t[228] = -2.025$, $p = .044$; $d = .26$). As such, pretest conceptual scores were added as a covariate in the mixed ANCOVAs that assessed condition effects on fraction PAE, whole number PAE, and procedural knowledge. The example condition also had a lower percentage of underrepresented racial and ethnic minority (URM) students ($\chi^2[1] = 5.143$, $p = .023$) than the control condition. Groups were comparable on all other pretest measures ($ps > .05$).

Assumptions of ANOVA and regression were met. Scores on continuous variables were normally distributed and variances were homogenous. Intraclass correlations (ICC) were calculated on posttest scores and were low, ranging from .05 for fraction PAE to .12 for procedural scores. Due to the low dependency of posttest scores on cluster and the small number of clusters, all further analyses were conducted at the student level. Our predictive models use maximum likelihood (ML) estimation with robust standard errors (SE). Robust SEs are cluster-adjusted and reduce bias. ML estimation is the gold standard for handling missing data (see Atkinson et al., 2003) and uses all information available to estimate parameters without requiring imputation or deleting participants with missing values (Acock, 2012).

Results

Main effects

To test whether there are benefits of supplementing textbooks with worked examples and self-explanation prompts on PAE for the whole number and fraction number line estimation tasks as well as procedural learning, we conducted three 2 (Time) $\times$ 2 (Condition) mixed analyses of covariance (ANCOVA), controlling for pretest conceptual scores. We also conducted a 2 (Time) $\times$ 2 (Condition) mixed ANOVA on conceptual scores. The interaction between Time and Condition on whole number PAE, conceptual scores, and procedural scores were not significant, indicating that improvements in each of these measures did not vary by condition ($F[1, 203] = 2.09$, $p = .149$; $F[1,228] = .594$, $p = .442$; $F[1, 227] = 2.438$, $p = .120$). However, the mixed ANCOVA

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

Table 2. 2 × 2 mixed ANCOVA on pre-posttest Fraction PAE by condition.

Source	*df*	*F*	*p*	η_p^2
Condition	1,220	0.345	.557	0.002
Time	1,220	4.014	.046	0.018
Pre-Conceptual	1,220	42.949	<.001	.163
Time × Pre-Conceptual	1,220	4.969	.027	
Time × Condition	1,220	4.796	.030	

Table 3. Moderating effect of prior estimation skills on effect of condition on post-test whole number and fraction PAE.

	Post-test whole number PAE				Post-test fraction PAE			
	Main effects ITT model		Moderation model		Main effects ITT model		Moderation model	
	β	*SE*	*β*	*SE*	*β*	*SE*	*β*	*SE*
Pretest whole number PAE	.63**	.07	.58**	.10	.13	.10	.17*	.08
Pretest fraction PAE	.09	.07	.08	.07	.56**	.10	.76**	.16
Preconceptual	−.13*	.06	−.13*	.06	.02	.05	.03	.05
Preprocedural	−.06	.05	−.07	.04	−.19**	.05	−.19**	.06
Example	.06	.04	.06	.05	−.10*	.04	−.11**	.04
Pre-whole number PAE×Example			.07	.08				
Pre-Fraction PAE×Example							−.30*	.12
R^2		.60		.61		.53	.57	

*$p < .05$ level. **$p < .01$ level. Estimator = ML with robust SE.

on fraction PAE scores yielded a significant Time × Condition interaction revealing differential improvements in PAE by condition over time (Table 2; $p = .030$).

Controlling for pretest conceptual knowledge, participants demonstrated greater improvements in fraction number line estimation (reduction in PAE) in the worked examples condition from pre- to posttest in comparison to the control condition, with a small effect size ($d = .218$).

Individual differences

Predicting Post-PAE for whole number and overall fractions

Next, to determine whether differences in whole number and overall fraction number line estimation skills at pretest moderated the effects of example on posttest whole number PAE and fraction PAE, we ran parallel multiple regressions predicting posttest whole number and fraction PAE (Table 3).

Although consistent with the mixed ANOVAs above, we present main effects models without the interaction terms to have a side-by-side comparison of standardized coefficients for the main and moderation models. To determine whether pretest whole number and overall fraction PAE scores moderated the effect of example on posttest whole number and fraction PAE, interaction terms between Pre-test Whole Number PAE and an Example dummy code and Pre-test Fraction PAE and the Example dummy code were added to the main effects models respectively. Continuous measures were centred prior to creating interaction terms for ease of interpretation.

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

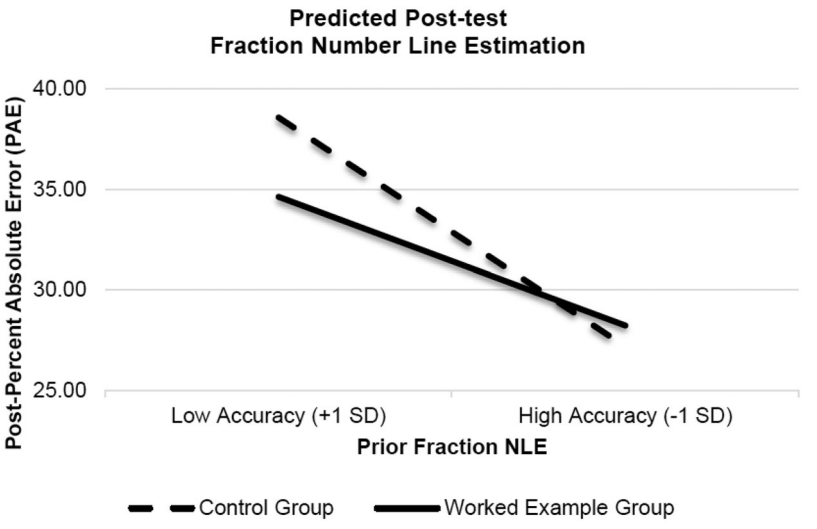

Figure 4. Moderation by pre-test fraction percent absolute error.

The effect of condition on posttest whole number PAE was not moderated by pretest whole number PAE ($\beta = .006$). Yet the effect of condition on posttest fraction PAE was significantly moderated by pretest fraction PAE ($\beta = -.259$; Table 3). To explicate the significant interaction between condition and pretest fraction PAE, simple slopes were estimated through a split regression by condition. Posttest overall fraction PAE was regressed onto pretest overall fraction PAE, pretest whole number PAE, and pretest conceptual and procedural scores for each condition. These split regressions revealed that, though pretest fraction PAE is a significant predictor of posttest fraction PAE in both conditions, the influence of pretest fraction PAE was smaller in magnitude for those in the example condition ($\beta = .483$, $p < .001$) in comparison to the control condition ($\beta = .637$, $p < .001$). Figure 4 displays the predicted posttest fraction number line PAE for participants at one standard deviation above and below the mean at pretest by condition.

Predicting post-procedural and post-conceptual

To determine whether differences in conceptual and procedural knowledge at pretest moderated the effects of example on posttest conceptual and procedural knowledge, we ran parallel multiple regressions (Table 4). As with number line measures, we established main effects models prior to moderation models in which we added interaction terms between each pretest and Example condition dummy code to the main effects models. Pretest procedural scores did not moderate the effect of condition on posttest procedural scores ($\beta = -.009$) but pretest conceptual scores did (Table 4).

Again, we estimated simple effects through split regressions regressing posttest conceptual score onto pretest, whole number PAE, and overall fractions PAE for each condition. Though pretest conceptual score is a significant predictor of conceptual learning in both conditions, using worked examples seems to buffer the influence of these pretest variables on posttest scores. This can be seen by a smaller beta (i.e. standardized effect size for regression interpreted as a proportion of a standard

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

Table 4. Moderating effect of prior knowledge on effect of condition on post-test conceptual and procedural scores.

	Post-test conceptual				Post-test procedural			
	Main effects ITT model		Moderation model		Main effects ITT model		Moderation model	
	β	*SE*	*β*	*SE*	*β*	*SE*	*B*	*SE*
Preconceptual	.40**	.10	.56**	.09	–		–	
Preprocedural	–		–		.55**	.05	.55**	.08
Whole number PAE	−.04	.08	−.04	.08	−.05	.08	−.05	.08
Fraction PAE	−.32**	.10	−.31**	.09	−.08	.07	−.08	.07
Example	.03	.05	.04	.06	−.07	.05	−.07	.05
Pre-conceptual × example			−.21*	.10				
Pre-procedural × Example							−.01	.08
R^2	.40		.42		.36		.36	

*$p < .05$ level. **$p < .01$ level. Estimator = Maximum likelihood with robust standard errors.

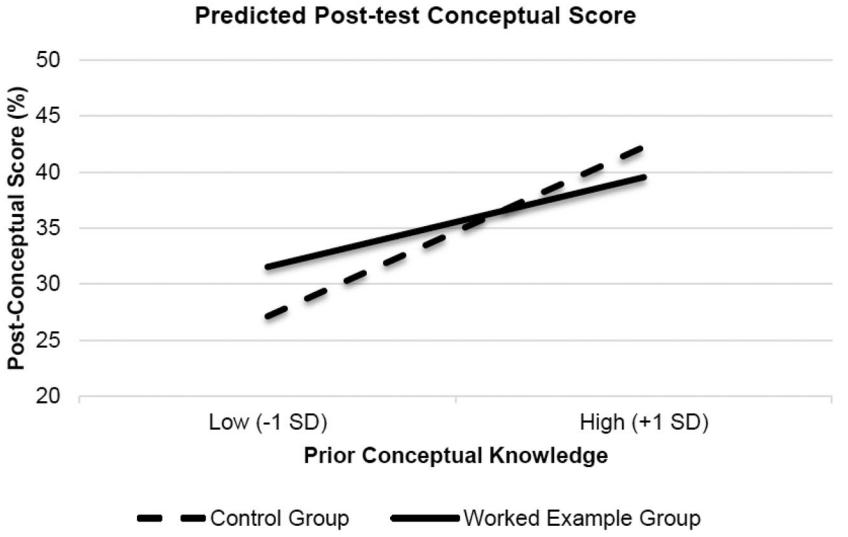

Figure 5. Moderation by prior rational number conceptual knowledge.

deviation change) for the influence of pre-conceptual scores ($\beta = .331$, $p < .001$) in the example condition in comparison to the control condition ($\beta = .499$, $p < .001$). Figure 5 displays the predicted posttest conceptual score for participants at one standard deviation above and below the mean at pretest by condition.

Discussion

We demonstrate that replacing practice problems with worked examples within a textbook (that highlight connections between decimals and fractions) is an effective method for improving pre-algebra students' fraction magnitude understanding. This is the first published study to extend the worked examples effect to the development of fraction magnitude, a critical understanding for learning higher-level mathematics. This effect was found despite the *noise* that exists in naturalistic settings like a classroom or home environment and thus has promising implications for instruction considering the ease with which these supplemented textbooks can be used in classrooms.

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

What is most interesting about this finding is that worked examples were not solely designed to target fraction magnitude estimation. That is, examples did not focus solely on number line representations of fractions but rather emphasized decimal magnitude and the interpretation of decimals as fractions. Worked examples that were added to the experimental textbook were feature-based problems and designed to focus students' attention on rational number principles and common conceptual errors. A possible explanation for this finding is that examples refined students' understanding of the features and principles of fractions and how they relate to decimals. This may have encouraged understanding of the meaning of the structure of a fraction. That is, when focusing on the fraction interpretation of decimals, understanding the relationship between the numerator and denominator would likely improve students' number line estimations of fractions. A recent study found that, although fractions are typically introduced in third grade (i.e. one grade level earlier than formal decimal instruction), fourth graders' level of decimal magnitude understanding in the middle of the academic year (i.e. winter) was a strong predictor of end-of-year estimation of fraction magnitudes, even after controlling for students' start-of-year fraction magnitude understanding (Resnick et al., 2019). Findings from this prior study suggest that improvements in decimal magnitude understanding may foster further understandings of fractions. Considering these findings and the shared numerical properties of rational numbers, it is not surprising that a unit on decimal operations that emphasizes connections between fractions and decimals and encourages the interpretation of decimals as fractions would yield improvements in fraction magnitude understanding. Unfortunately, we did not administer a decimal magnitude estimation task. Future work should include a decimal number line estimation task to determine whether emphasizing connections between fractions and decimals in conceptually-based worked examples also improves decimal magnitude understanding.

Differential effects by learners' prior knowledge

The benefits of worked examples are stronger for students with low prior knowledge of rational number concepts and those who struggle with fraction magnitudes. Students who began the unit with lower levels of fraction magnitude understanding showed more improvement in their fraction number line estimates when working with the supplemented textbook. Similarly, though pretest conceptual score was a significant predictor of conceptual learning overall in both conditions, worked examples lessened the influence that low prior knowledge would have on learning, yielding a buffering effect of the normally negative influence of low prior knowledge on the development of fraction magnitude understanding and rational number concepts more generally. Worked examples that highlight connections between the structure and magnitude representations of fractions and decimals seem to provide the support students need when in the early stages of skill acquisition. This finding is consistent with Booth et al.'s (2015) finding that a combination of worked examples is particularly effective for students with low prior algebra knowledge. However, this extends these findings to rational number concepts and fraction magnitude in particular which has important practical implications. Considering the predictive power of fraction

magnitude understanding for later mathematics achievement (Resnick et al., 2016) and algebra in particular (Booth et al., 2014), there is a need for evidence-based practices that improve students' fraction understanding.

Strengths and limitations

We establish high ecological validity by conducting this study in a real-world setting with a diverse sample. As laboratory findings do not always translate to classrooms, findings must be tested in real-world settings before generalizing and making recommendations to practitioners. The current study demonstrates that the worked examples effect indeed translates to the classroom and is powerful enough to yield effects even in minimally controlled settings such as the use of textbooks for homework or in classroom practice. We are confident in the level of control that we were able to retain while still maintaining ecological validity. Teachers taught at least one experimental class and one control class, textbooks were used mainly as homework, and minimal differences were found between classes on key study variables. Any differences in pretest were controlled for in the corresponding analyses. These features give us greater confidence that the improvements seen are due to the experimental textbooks.

Despite these strengths, the study is limited by several factors including the lack of a delayed post-test measure and a measure that assesses the influence of these improvements on students' overall mathematics achievement or later algebra knowledge. Future research might address these issues by including a delayed post-test several weeks after completion of the unit to ensure that improvements are retained. Collecting data from the school or district on standardized end-of-the-year measures as well as following students longitudinally as they progress through algebra will help to elucidate the relationship between improvements in understanding of fraction magnitude and algebra learning. Also, we did minimize the chance that effects would be influenced by differences in teachers or instruction by having teachers teach at least one control and one experimental condition. However, future research may randomize at the student level within classroom to fully control for any potential systematic effects on the influence of conditions due to differences by classroom or teacher. Lastly, though we do know that teachers assigned the manipulated problems as homework in all classes over the course of the study, we currently lack detailed information on dosage or frequency of use that is required to benefit from the experimental textbooks. There was likely great variability in the amount of textbook use by teacher. Future studies should include collection of data on how frequently the textbook was used, how many of the worked-examples problems were assigned, and how frequently practice was assigned as classwork versus homework to better understand the conditions under which examples are helpful in the classroom. Process data can also be collected, for example, by administering materials in an online cognitive tutor that has the capacity to capture usage automatically (e.g., ASSISTments; Roschelle et al., 2016). Additionally, although teachers reported using the same instructional strategies for both experimental and control conditions and Intent-to-Treat analysis was used, future studies should include classroom observations to measure fidelity,

Accepted Manuscript
Version of record at: https://doi.org/10.1080/01443410.2022.2144142

ensure no diffusion of treatment, and gain a better understanding of textbook use to make more specific recommendations for classroom practices.

Conceptual and procedural knowledge of rational numbers

Although worked examples improved students' fraction magnitude understanding as expected overall, we did not find improvements in rational number concepts, with the exception of benefits for students with low prior conceptual knowledge. This may be due to the already high level of conceptual knowledge that students revealed on the rational number concepts pretest. The overall sample had relatively high conceptual scores at 80% correct at pretest. When calculated by condition, this was 78% correct for the control group and 82% correct for the experimental group. This may be explained by the *expertise reversal effect*, which is the phenomenon that instruction that is effective with inexperienced learners is particularly *ineffective* with experienced learners (Kalyuga et al., 2003 for a review). Our finding that students with low prior conceptual knowledge benefitted more from the example textbooks than the control textbooks lends support to this interpretation.

One unexpected finding concerned the lack of differential improvement by condition in procedural skills. Although we expected the experimental textbooks to show greater improvements in procedural skills, the control textbooks were just as effective. Worked examples designed to highlight underlying principles and important conceptual features within a problem may be limited in their influence on procedures. Although these examples improve students' understanding of fraction magnitude and refine their understanding of other rational number concepts of those in the early stages of skill acquisition, direct instruction on procedures may be needed before students can fully benefit from these examples. Alternatively, it is also possible that the procedural skills enhanced through the example textbooks were not directly assessed on our procedural measure which focused mainly on finding common denominators and calculating percentages presented through word problems. Future research in this area should use a more comprehensive measure of procedures with rational numbers including skills such as fraction and decimal arithmetic.

Educational implications

The current study compared the effects of supplementing a textbook from *Connected Mathematics* – a Standards-based curriculum focused on inquiry and making connections between mathematical ideas to solve real-world problems with little to no reliance on worked examples – with worked examples during problem-solving practice (typically assigned after instruction as homework). If teachers followed the curriculum as intended, the only 'worked examples' that students would be exposed to would be those possible strategies presented by students and open to class discussion. Although the design of the current study does not allow one to test hypotheses regarding the role that CMP in particular played in the current findings, we propose several conjectures that may be considered in future work. It is possible that the effect of worked examples was more pronounced here when compared to CMP than it

would be to other textbooks, as CMP is a curriculum with minimal inclusion of worked examples that expects students to invent strategies based on their understanding of mathematics. This may be especially relevant when considering the finding that the experimental textbooks were helpful in improving conceptual understanding of those who began the study with low prior knowledge of rational number concepts. In a CMP classroom, these students may struggle to invent their own strategies based on the limited conceptual understanding that they have and may benefit from having access to several examples to work from when completing practice problems. This, in turn, may boost students' ability to engage in the classroom discussion as intended by the CMP curriculum.

On a related note, considering that the CMP textbooks provide very few examples, the experimental textbook's inclusion of examples may increase caregivers' ability to assist students with homework, as reported by some participating teachers in our study. The combination of examples and a conceptually-based textbook may have been a particularly strong pairing. It is possible that supplementing a more traditional curriculum or textbook that already includes many examples with even more worked examples would yield different or lesser effects. However, prior work demonstrates that supplementing a traditional curriculum that includes strictly procedural examples with conceptually-based examples and explanations improves students' conceptual understanding as well (Booth et al., 2013; Booth et al., 2015). Thus, further work is needed to determine whether there is a differential benefit to adding worked examples to different types of curricula or if the effect is comparable across curricula.

Supplementing textbooks or classroom learning materials with a range of worked example types (i.e. correct, incorrect, and faded) focused on the connections between procedures and underlying rational number concepts is an effective way to improve pre-algebra students' understanding of fraction magnitude. This specific worked example approach, which also makes explicit connections between fractions and decimals, is particularly beneficial for improving the conceptual understanding of those with low prior conceptual knowledge of rational numbers. The NMAP (2008) emphasizes the importance of fraction knowledge for future success. Prior research demonstrates the great difficulty that students have with fractions (DeWolf & Vosniadou, 2015). Fraction magnitude, in particular, is an area that typically shows little growth between fourth grade when fractions are often introduced and sixth grade when we expect proficiency in fractions (Resnick et al., 2016). Our findings provide an efficient and practitioner-friendly tool to implement in classrooms to help students overcome the hurdle that is conceptual fraction understanding.

Author note

The opinions expressed are those of the authors and do not represent views of the Institute or the United States Department of Education. The data are not publicly available due to privacy or ethical restrictions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Funding for this research was provided by the Institute of Education Sciences, the United States Department of Education, through Grant R305C100024 to West Ed. The data that support the findings of this study are available on request from the corresponding author.

ORCID

Christina Areizaga Barbieri http://orcid.org/0000-0002-4338-5807
Julie L. Booth http://orcid.org/0000-0003-2300-2116
Kamal Chawla http://orcid.org/0000-0003-0125-817X

References

Acock, A.C. (2012). What to do about missing values. In H. Cooper (Ed.), APA handbook of research methods in psychology. Washington, DC: American Psychological Association. https://doi.org/10.1037/13621-002

Adams, D. M., McLaren, B. M., Durkin, K., Mayer, R. E., Rittle-Johnson, B., Isotani, S., & van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior*, *36*, 401–411. https://doi.org/10.1016/j.chb.2014.03.053

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*(4), 774–783. https://doi.org/10.1037/0022-0663.95.4.774

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology*, *113*(3), 447–455. https://doi.org/10.1016/j.jecp.2012.06.004

Barbieri, C. A., & Booth, J. L. (2020). Mistakes on display: Incorrect examples refine equation solving and algebraic feature knowledge. *Applied Cognitive Psychology*, *34*(4), 862–878. https://doi.org/10.1002/acp.3663

Barbieri, C. A., Miller-Cotto, D., & Booth, J. L. (2019). Lessening the load of misconceptions: Design-based principles for algebra learning. *Journal of the Learning Sciences*, *28*(3), 381–417. https://doi.org/10.1080/10508406.2019.1573428

Barbieri, C. A., Young, L. K., Newton, K. J., & Booth, J. L. (2021). Predicting middle school profiles of algebra performance using fraction knowledge. *Child Development*, *92*(5), 1984–2005. https://doi.org/10.1111/cdev.13568

Barbieri, C., & Booth, J. L. (2016). Support for struggling students in algebra: Contributions of incorrect worked examples. *Learning and Individual Differences*, *48*, 36–44. https://doi.org/10.1016/j.lindif.2016.04.001

Berthold, K., Eysink, T. H. S., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science*, *37*(4), 345–363. https://doi.org/10.1007/s11251-008-9051-z

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, *79*(4), 1016–1031. https://doi.org/10.1111/j.1467-8624.2008.01173.x

Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect

examples. *Learning and Instruction*, *25*, 24–34. https://doi.org/10.1016/j.learninstruc.2012.11.002

Booth, J. L., McGinn, K. M., Young, L. K., & Barbieri, C. (2016). Simple practice doesn't necessarily make perfect: Evidence from the worked example effect. *Policy Insights from Behavioral and Brain Sciences*, *2*(1), 24–32. https://doi.org/10.1177/2372732215601691

Booth, J. L., Newton, K. J., & Twiss-Garrity, L. K. (2014). The impact of fraction magnitude knowledge on algebra performance and learning. *Journal of Experimental Child Psychology*, *118*, 110–118. https://doi.org/10.1016/j.jecp.2013.09.001

Booth, J. L., Oyer, M. H., Paré-Blagoev, E. J., Elliot, A., Barbieri, C., Augustine, A. A., & Koedinger, K. R. (2015). Learning algebra by example in real-world classrooms. *Journal of Research on Educational Effectiveness*, *8*(4), 530–551. https://doi.org/10.1080/19345747.2015.1055636

Burke-Miller, S. (2012). Intent-to-treat analysis. In S. Boslaugh (Ed.), *Encyclopedia of epidemiology* (p. 546). SAGE Publications.

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction*, *37*, 39–49. https://doi.org/10.1016/j.learninstruc.2014.07.002

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, *22*(3), 206–214. https://doi.org/10.1016/j.learninstruc.2011.11.001

Fulkerson, W. O. (2013). *2012 National survey of science and mathematics education: Status of middle school mathematics*. Horizon Research. http://www.horizon-research.com/2012-national-survey-of-science-and-mathematics-education-status-of-middle-school-science

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Lappan, G., Fey, J., Fitzgerald, W., Friel, S., & Philips, E. (2006). *Connected mathematics 2*. Prentice Hall (School Division).

Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and holes: Gaps in rational numbersense in children with vs. without mathematical learning disability. *Developmental Science*, *11*(5), 681–691. https://doi.org/10.1111/j.1467-7687.2008.00717.x

McCoy, C. E. (2017). Understanding the intention-to-treat principle in randomized controlled trials. *The Western Journal of Emergency Medicine*, *18*(6), 1075–1078. https://doi.org/10.5811/westjem.2017.8.35985

McNeil, N. M. (2014). A change-resistance account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, *8*(1), 42–47. https://doi.org/10.1111/cdep.12062

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards (CCSS-M)*. http://www.corestandards.org/

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. US Department of Education. https://files.eric.ed.gov/fulltext/ED500486.pdf

Newton, K. J., Willard, C., & Teufel, C. (2014). An examination of the ways that students with learning disabilities solve fraction computation problems. *The Elementary School Journal*, *115*(1), 1–21. https://doi.org/10.1086/676949

Paas, F., & van Merriënboer, J. J. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, *29*(4), 394–398. https://doi.org/10.1177/0963721420922183

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, *32*(1/2), 1–8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0

Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading worked solution steps works –A cognitive load perspective. *Instructional Science*, *32*(1/2), 59–82. https://doi.org/10.1023/B:TRUC.0000021815.74806.f6

Resnick, I., Jordan, N. C., Hansen, N., Rajan, V., Rodrigues, J., Siegler, R. S., & Fuchs, L. S. (2016). Developmental growth trajectories in understanding of fraction magnitude from fourth

through sixth grade. *Developmental Psychology*, *52*(5), 746–757. https://doi.org/10.1037/dev0000102

Resnick, I., Rinne, L., Barbieri, C. A., & Jordan, N. C. (2019). Children's reasoning about decimals and its relation to fraction learning and mathematics achievement. *Journal of Educational Psychology*, *111*(4), 604–618. https://doi.org/10.1037/edu0000309

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives*, *11*(3), 184–190. https://doi.org/10.1111/cdep.12229

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362. https://doi.org/10.1037/0022-0663.93.2.346

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open*, *2*(4), 233285841667396–233285841667312. https://doi.org/10.1177/2332858416673968

Siegler, R. S., & Chen, Z. (2008). Differentiation and integration: Guiding principles for analyzing cognitive change. *Developmental Science*, *11*(4), 433–448. https://doi.org/10.1111/j.1467-7687.2008.00689.x

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. https://doi.org/10.1177/0956797612440101

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction*, *14*(5), 503–518. https://doi.org/10.1016/j.learninstruc.2004.06.015

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, *2*(1), 50–80. https://doi.org/10.1080/1380361960020103

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction*, *16*(2), 165–169. https://doi.org/10.1016/j.learninstruc.2006.02.005

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. https://doi.org/10.1207/s1532690xci0201_3

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Ye, A., Resnick, I., Hansen, N., Rodrigues, J., Rinne, L., & Jordan, N. C. (2016). Pathways to fraction learning: Numerical abilities mediate the relation between early cognitive competencies and later fraction knowledge. *Journal of Experimental Child Psychology*, *152*, 242–263. https://doi.org/10.1016/j.jecp.2016.08.001