International Journal of Mathematical Education in Science and Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tmes20

Cognitive instructional principles in elementary mathematics classrooms: a case of teaching inverse relations

Meixia Ding, Ryan Hassler & Xiaobao Li

To cite this article: Meixia Ding, Ryan Hassler & Xiaobao Li (2021) Cognitive instructional principles in elementary mathematics classrooms: a case of teaching inverse relations, International Journal of Mathematical Education in Science and Technology, 52:8, 1195-1224, DOI: 10.1080/0020739X.2020.1749319

To link to this article: https://doi.org/10.1080/0020739X.2020.1749319

Published online: 16 Apr 2020.

INTERNATIONAL JOURNAL OF MATHEMATICAL EDUCATION IN SCIENCE AND TECHNOLOGY
2021, VOL. 52, NO. 8, 1195–1224
https://doi.org/10.1080/0020739X.2020.1749319

Check for updates

Cognitive instructional principles in elementary mathematics classrooms: a case of teaching inverse relations

Meixia Ding [a], Ryan Hassler[b] and Xiaobao Li[c]

[a]Department of Teaching and Learning, Temple University, Philadelphia, PA, USA; [b]Mathematics Department, Pennsylvania State University - Berks, Reading, PA, USA; [c]Center for Education, Widener University, Chester, PA, USA

ABSTRACT

Instructional principles gleaned from cognitive science play a critical role in improving classroom teaching. This study examines how three cognitive instructional principles including worked examples, representations, and deep questions are used in eight experienced elementary teachers' early algebra lessons in the U.S. Based on the analysis of 32 videotaped lessons of inverse relations, we found that most teachers spent sufficient class time on worked examples; however, some lessons included repetitive examples that also included irrelevant practice problems. Most teachers also situated new teaching in concrete contexts, which were faded into abstract representations. However, connections between concrete and abstract were not always made. The largest challenge was rooted in teachers' inability to ask deep questions that elicited students' deep explanations. Some teachers focused on key words and provided students with direct explanations. Implications are discussed.

ARTICLE HISTORY
Received 6 February 2020

KEYWORDS
Worked example; representation; deep question; inverse relation; classroom instruction; video analysis

1. Introduction

Although enhancing mathematics classroom teaching continues to be a pressing goal in the U.S., the quality of classroom instruction continues to be unchanged. 'The core of teaching – the way teachers and students interact about content – has remained the same for a century or more' (Hiebert & Morris, 2012, p. 97). To address this issue, researchers call for a science of improvement (Bryk, 2009; Cai et al., 2019) focusing on teaching rather than teachers (Hiebert & Morris, 2012; Hiebert & Stigler, 2017). Instructional principles gleaned from cognitive science can be expected to play a critical role in this endeavour (Anderson, 2010; Kirschner et al., 2006; Mayer, 2010). In fact, the Institute of Education Sciences (IES) in the U.S. recommended teachers organize instruction based on cognitive principles (e.g. using worked examples to enhance problem solving, making connections between concrete and abstract representations, and asking deep questions to elicit student self-explanations) to improve student learning (Pashler et al., 2007). In this study, we refer to 'cognitive instructional principles' as the aforementioned recommendations on teachers'

CONTACT Meixia Ding meixia.ding@temple.edu Ritter Hall 436, 1301 Cecil B. Moore Avenue, Philadelphia, PA 19122-6091, USA

use of worked examples, representations, and deep questions. According to Pashler et al., these recommendations were based on cognitive research findings, which intended to provide teachers across subjects with strategies for organizing teaching to facilitate students learning. While cognitive research findings have a potential to improve classroom instruction, the applications of broad instructional principles into subject-specific teaching like mathematics is often not straightforward (Anderson et al., 2000; Kirschner et al., 2006). Ding and Carlson (2013) found that elementary teachers had challenges in implementing the aforementioned IES recommendations when planning their mathematics lessons; with support, those teachers utilized the cognitive principles better in lesson planning. Despite the previous findings on lesson planning, we know little about how the targeted cognitive instructional principles recommended by the IES (Pashler et al., 2007) are implemented in actual elementary mathematics classrooms. The lack of relevant information about the successes and challenges in existing classroom teaching may hinder the subsequent intervention that aims to improve instructional quality. As such, it is necessary to obtain detailed understanding of the strength and limitation of current mathematics teaching in alignment with cognitive instructional principles in natural classroom settings.

The purpose of this study is to examine experienced elementary teachers' classroom teaching on inverse relations, which serve as a case to understand teachers' instructional successes and challenges in applying the targeted cognitive principles in elementary mathematics classrooms. Inverse relations are an important topic of early algebra, which has been recognized as an important gatekeeper to students' success in mathematics (Carpenter et al., 2003; National Mathematics Advisory Panel [NMAP], 2008). Even though the development of students' algebraic thinking in elementary school has received increasing attention (Carpenter et al., 2003; Carraher & Schliemann, 2007; Common Core State Standards Initiatives [CCSSI], 2010; Kieran, 2018), the teaching of early algebra, including inverse relations, continues to be a challenge for elementary teachers (Kieran, 2018). As such, an examination of classroom teaching on inverse relations may bring insights into ways to better support student learning. To examine these lessons, we focus on the three aforementioned cognitive instructional principles, the use of worked examples, representations, and deep questions (Pashler et al., 2007, elaborate upon later). In particular, we ask: (1) How do sampled elementary teachers use worked examples when teaching inverse relations in elementary classrooms? (2) How do sampled elementary teachers use representations when teaching inverse relations in elementary classrooms? And (3) how do sampled elementary teachers use deep questions when teaching inverse relations in elementary classrooms? It is expected that this study will inform the fields of mathematics education and cognitive research about how the targeted instructional principles are used in elementary mathematics classrooms, which may further contribute to a science of improvement (Berwick, 2008; Bryk, 2009; Cai et al., 2019; Lewis, 2015) of mathematics teaching.

2. Literature review

Any successful learning entails integrating new knowledge to existing knowledge, which further leads to the changes of long-term memory resulting in more coherent mental models of concepts (Anderson, 2010; Mayer, 2010). In other words, unless students are cognitively engaged in classroom instruction, the learning outcome would be limited (Chi & Wylie, 2014; Chi et al., 2018). In this study, the three targeted instructional principles –

the use of worked examples to enhance problem solving, making connections between concrete and abstract representations, and asking deep questions to elicit deep explanations – were aligned with the above overachieving learning theory. These three instructional principles were modified from the IES recommendations (Pashler et al., 2007), which contains seven instructional principles drawn from best available evidence in cognitive research and classroom experiments on various subjects. We collapsed the recommendation of 'combine graphics with verbal descriptions' (p. 9) into 'connect and integrate abstract and concrete representations' (p. 15) because the former is related to concrete representations. Our priority in selecting the above principles is due to the consideration that the rest principles – 'space learning over time' (p. 5), 'use quizzing to promote learning' (p. 19), and 'help students allocate study time efficiently' (p. 23) – are relatively distant from classroom teaching. Below, we review the powerfulness of the targeted instructional principles and associated challenges for implementation in the classroom setting, followed by a review of our prior findings on teachers' use of these principles in lesson planning. Since our investigation is situated in the topic of inverse relations, we also provide a brief analysis of this research topic.

2.1. Using worked examples to enhance problem solving: powerfulness and challenges

Worked examples refer to problems with solutions given and researchers have been examining the use of worked examples for decades. Prior studies have found that worked examples were effective in helping students acquire necessary schemas to solve new problems (Catrambone & Yuasa, 2006; Sweller & Cooper, 1985; van Gog et al., 2011; Zhu & Simon, 1987). Schemas are mental constructs that allow for classifying a new problem into a previous category and selecting an appropriate solution mode for that specific category (Sweller & Cooper, 1985). Research on expert-novice found that domain specific knowledge in the form of schema was a main factor that distinguishes expert and novice in problem solving (Chi et al., 1981; Sweller, 1988). As such, worked examples were recommended as an effective means for new learning in comparison with conventional problem-solving (Kirschner et al., 2006; Pashler et al., 2007). This is because problem solving contains a search process used by novice to obtain a solution, which increases students' cognitive load (Sweller, 1988) and slows the schema acquisition rate, resulting in low effectiveness of learning (Sweller & Cooper, 1985). The effect of worked examples is relevant and critical for learning early algebra topics like inverse relations because well-understood worked examples have the potential to develop students' structural understanding (Kieran, 2018). In fact, Mason (2018) recommended that teachers discuss worked examples in classrooms to develop students' algebraic thinking.

Despite the well-known worked example effect (Catrambone & Yuasa, 2006; Sweller & Cooper, 1985; van Gog et al., 2011; Zhu & Simon, 1987), many mathematics classrooms have neglected this cognitive principle with limited classroom time spent on worked examples (Stigler & Hiebert, 1999). This is more of a concern in current reformed classrooms where teachers often step back with minimum instruction on an example task (Kirschner et al., 2006). For instance, teachers in reformed classes often moved quickly to problem solving where students were encouraged to work with peers through cooperative, or discovery, or project-based learning (Kirschner et al., 2006). Consequently, the worked examples often received little attention during classroom instruction. However, as explained above, without a relevant schema established through worked examples, the subsequent problem solving process would be questionable and ineffective (Ding et al., 2007).

To apply what was learned from a worked example, students ought to practice problems relevant to the given example. In fact, to enhance the worked example effect, some researchers suggest fading examples into practice (Renkl et al., 2004) by asking students to solve some steps of a worked example task. This indicates the necessity of linking worked example and practice problems. Pashler et al. (2007) also recommended alternating between worked examples and practice problems, which, however, was found unnecessary in van Gog et al. (2011). In this study, we examine whether teachers in our study appear to pay attention to the worked examples and whether their subsequent practice problems are related to the worked examples. Of course, any learning process should actively engage students. When teachers simply 'show and tell' students the full example solutions, teachers might have replaced students' knowledge-construction activities (Wittwer & Renkl, 2008), which does not promote the worked example effect (Catrambone & Yuasa, 2006). This calls for attention to the other two instructional principles (representations, questions) that promote active learning.

2.2. Linking concrete and abstract representations: powerfulness and challenges

Active learning demands students' engagement in the process of modelling and sense-making (Mayer, 2010). Situating classroom teaching of abstract concepts in concrete contexts supports students' initial learning because familiar situations may facilitate sense-making (Resnick et al., 1987). In mathematical learning, concrete contexts often refer to real-world situations presented in a format of word problems. Researchers of Realistic Mathematics Education (RME) has emphasized that mathematics must be taught in ways that are close to children and which are relevant to every day life. Gerofsky (2009) argued that real-world situations 'have the potential to offer memorable imagery that can act as a touchstone for teachers and learners in building and discussing abstract concepts' (p. 36). In comparison with real-world situations embedded in word problems, physical manipulatives seem to be less 'concrete' in the sense of activating students' imagery of every day life. In addition, students' learning of real-world problems in the written text can also be enhanced by adding relevant pictures and graphs (Pashler et al., 2007). Mathematics textbooks in high achieving countries (e.g. China, Singapore, and Japan) demonstrated this feature (Cai & Moyer, 2008; Ding & Li, 2014; Murata, 2008).

However, overexposing students to concrete representations may hinder students' transfer of the learned knowledge because these representations often contain irrelevant and distracting information (Kaminski et al., 2008). To overcome this representational difficulty, it is necessary to make connections between concrete and abstract representations of the targeted concept (Pashler et al., 2007). This is because only abstract representations of mathematical knowledge have the power to transfer over contexts to solve new problems (Goldstone & Son, 2005; Kaminski et al., 2008). To facilitate the transition from concrete to abstract, recent research suggests using a concreteness fading approach by fading from concrete (e.g. real objects and pictures) into semi-concrete (e.g. dots, lines, schematic diagrams) and then finally into abstract representations (e.g. number sentences; Goldstone & Son, 2005). According to Fyfe and Nathan (2019), the general notion of concreteness fading is aligned with Bruner's (1966) three stages of representation shifting from enactive (action based) to iconic (image-based) and eventually to symbolic (notation-based). Experimental studies have shown that the concreteness fading approach, in comparison with concrete only, abstract only, and from abstract to concrete, is most effective in supporting students' learning and transfer of mathematical concepts (Fyfe et al., 2015; McNeil & Fyfe, 2012). However, Duval (2006) noted that the translation between representations is challenging, serving as a source of mathematics incomprehension because students had difficulties seeing structural commonalities and thus struggled to map the structures of concrete and abstract representations. Classroom research has reported teachers' missed opportunities to promote such translation. For instance, Kazemi and Stipek (2001) found that when students presented two strategies using different concrete representations but with the same mathematical idea, the teacher did not facilitate connection making between concrete and abstract and thus treated these strategies as mathematically different. These findings indicate significant challenges for teachers when facilitating representational connections.

2.3. Asking deep questions to elicit self-explanations: powerfulness and challenges

Active learning also demands students' articulation of their mathematical thinking (Mayer, 2010). Research on self-explanations found that this process enables students to integrate new concepts and ideas to prior ones, thus generating new inference which produces more coherent mental models (Chi, 2000). Worked example research also found that it is important to promote students' self-explanation of worked examples to enhance the worked example effect (Catrambone & Yuasa, 2006; Chi et al., 1989). However, students may lack the motivation or the ability to spontaneously provide explanations, which calls for teachers' deep questions to elicit such explanations (Pashler et al., 2007). Deep questions refer to those targeting the underlying principles, relationships, and structures (Craig et al., 2006). Examples of deep questions include why, why-not, what caused X, how, how did X occur, what-if, what-if-not, how does X compare to Y, what is the evidence for X, and why is X important? (Pashler et al., 2007). Through deep questioning, teachers can invite students into classroom dialogues, providing explanations and justifications and modifying interpretations when needed.

Research on teacher questioning is not at all new. A century ago, Stevens (1912) already viewed teacher questioning as a measure for instructional efficiency. Gall (1970) conducted a thorough review of teacher questions in classrooms. Recently, there was a call for fundamental change from teaching by telling to teaching by questioning (Chi et al., 2018). However, asking deep questions has been consistently a challenge for classroom teachers especially in mathematics classrooms (Franke et al., 2009; Martino & Maher, 1999). Franke et al. (2009) reported that teachers in cognitive-guided classrooms were able to ask initial deep questions such as 'How did you get that?' However, when deep explanations were not elicited, they struggled with asking follow-up questions. This may be due to teachers' limited knowledge of concept connections because asking effective follow-up questions to facilitate students' sense-making demands teachers' onsite recognition of the relevant prior concepts on which the new concept was built on (Hiebert et al., 1997). In fact, some teachers tended to provide instructional explanations themselves, which may be due to their teaching habits or beliefs. However, such habits or beliefs may deprive students of reasoning and sense-making opportunities (Wittwer & Renkl, 2008), resulting in superficial learning.

2.4. Teacher implementation of the cognitive instructional principles in a prior study

The above three instructional principles – the use of worked examples, representations, and deep questions – form a conceptual framework for this current study. To our best knowledge, very few classroom studies have explored how elementary teachers implement this set of instructional principles to teach early algebra. Our prior study on elementary teachers' lesson planning explored these aspects (Ding & Carlson, 2013). Since lesson plans are closely related to classroom teaching and reflect teachers' thinking about how a lesson should be taught (Stein et al., 2007), we review our prior findings that may inform the current study.

In our prior study, we provided an intervention based on the above three instructional principles through a summer course involving 35 grades K-3 teachers who were selected to a research project. Prior to intervention, we asked teachers to read the IES recommendations (Pashler et al., 2007), which was expected to inform their initial lesson planning on early algebra topics such as inverse relations. After intervention, we asked teachers to revise their lesson plans incorporating the instructional principles discussed in class. At the end of the course, we also asked teachers to develop an independent lesson plan using a different early algebra topic. Teachers' three lesson plans were compared both quantitatively and qualitatively.

It was found that when reading the IES recommendations themselves, teachers had tremendous difficulties with translating these instructional principles to their lesson plans. For instance, with regard to worked examples, some teachers planned a series of repetitive worked examples with only minimal discussion devoted to each example. We also found that teachers who involved concrete representations in their lesson plans did not link them well to abstract representation; when multiple representations were used, the connections between these representations were often lacking. Moreover, we found that some teachers who asked 'why' questions in their lesson plans only expected superficial explanations from students. After the course discussion of these instructional principles, teachers made significant improvement on their initial lesson plans. They also maintained their learned skills in their end-of-course independently prepared lesson plan. However, challenges remained with making connections from concrete to abstract representations and asking deep questions to elicit students' self-explanations. Given that lesson plans are only teachers' lesson images rather than actual classroom practices, it is necessary to examine the degree to which experienced teachers' classroom teaching in the current study aligns with these instructional principles and whether these teachers face similar challenges in the teaching context. This endeavour is extremely important because one of the purposes of cognitive science research is to make an impact on human practice; on the other hand, the current need to improve mathematics classroom teaching also calls for scientific guidance of cognitive research.

2.5. Inverse relations: a case of early algebra

To explore how cognitive instructional principles occur in experienced teachers' mathematics classrooms, we focus on the case of inverse relations. This narrowed focus allows in-depth and close comparisons among classroom instructions. As previously mentioned, inverse relations is an important early algebra topic, which calls for bringing out the algebraic characteristics of arithmetic (Russell et al., 2011; Schliemann et al., 2007). To do so, early algebra researchers (Carraher & Schliemann, 2007; Kaput, 2008; Kieran, 2018) stressed the importance of focusing on fundamental mathematical ideas such as principles, relations, and structures that govern both arithmetic and algebra. Further, inverse relations is a fundamental mathematical idea that is given much attention by the Common Core State Standards (CCSSI, 2010) throughout elementary grades.

'Inverse relations' in the current study refers to the *complement principle* (Baroody et al., 2009), which includes both additive inverses (e.g. if $a + b = c$, then $c - b = a$) and multiplicative inverses (if $a \times b = c$, then $c \div b = a$; Vergnaud, 1988). Such structural relationships may be initially learned through arithmetic tasks such as fact family (e.g. $7 + 5 = 12$, $5 + 7 = 12$, $12–7 = 5$, and $12–5 = 7$; Carpenter et al., 2003), inverse word problems (the solutions form a fact family; Carpenter et al., 2003), start-unknown problems (Nunes et al., 2009), and using inverse relations to compute or check (e.g. $81–79 = ?$ may be solved/checked by using $79 + 2 = 81$; Torbeyns et al., 2009). Students' understanding of inverse relations likely contributes to their full comprehension of the four basic operations and overall algebraic thinking (Carpenter et al., 2003; Nunes et al., 2009).

Despite the importance of inverse relations, prior research on this topic revealed at least two limitations of classroom instruction. First, instruction concentrates on number manipulations rather than sense-making. Our previous study (Ding & Carlson, 2013) indicates that very few teachers planned to make use of concrete contexts presented by textbooks to teach inverse relation. This is problematic because the meaning of symbols is distant from students' real-life experiences. Therefore, concrete contexts that are familiar to students should be used to provide a space for students' sense-making (Radford & Roth, 2011). Otherwise, students tend to make mistakes such as $7 \div 35 = 5$ and $5 \div 35 = 7$ (Ding & Carlson, 2013). Second, existing instruction focuses on inverse-based strategies rather than the underlying relations. Torbeyns et al. (2009) reported that to solve 81–79 using $79 + 2 = 81$, students were taught to draw a little arrow from the subtrahend to the minuend, yet the underlying inverse relation was not made explicit. Instructional explanations that focus on procedural strategies rather than the underlying concepts do not contribute to students' deep learning (Chi & VanLehn, 2012; Wittwer & Renkl, 2008). The above findings suggest an examination of current classroom practice regarding how those available cognitive instructional principles such as using worked examples, representations, and deep questions may be used in classrooms to develop students' understanding of inverse relations and thus algebraic thinking.

3. Methods

This study is part of a five-year National Science Foundation (NSF) project on early algebra. The large project aimed to identify, in alignment of the aforementioned conceptual framework, the necessary knowledge to teach early algebra based on cross-cultural videotaped lessons in U.S. and China. For the current study, we employ a case study method

Table 1. Teacher characteristics and textbook uses.

Teacher	Grade	Gender	Year of Teaching	Reputation	Textbook
Ann	G1	F	> 26	Principal Recommended	Go Math[b]
Bea	G1	F	16–20	Principal Recommended	Investigation[c]
Carla	G2	F	> 26	NBCT[a]	Investigation
Daria	G2	F	16–20	Principal Recommended	Go Math
Emily	G3	F	16–20	NBCT candidate	Investigation
Faith	G3	F	21–25	NBCT	Go Math
Gia	G4	F	> 26	Principal Recommended	Go Math
Henry	G4	M	6–10	NBCT	Investigation

Note: [a]NBCT refers to National Board Certified Teachers. [b,c]Names of the mathematics textbook series used by the teacher participants in this study.

(Creswell & Poth, 2018) exploring how the sampled U.S. teachers teach inverse relations in their classrooms without any input from project researchers.

3.1. Participants

Eight experienced teachers from grades 1–4 participated in the current study. These teachers were our year 1 teacher participants who were selected from a candidate pool. Except for one, all teachers had more than 15 years of experience and were scored above their peers in the recruitment pool for the written knowledge survey. Note that the teacher who taught less than 10 years was a National Board Certified Teacher (NBCT) at the time he joined the project. Overall, three of the eight teachers were NBCT and one was an NBCT candidate. The other teachers were highly recommended by the school district and their principals. The following pseudonyms were given to the teachers: Ann and Bea (Grade 1 or G1), Carla and Daria (G2), Emily and Faith (G3), and Gia and Henry (G4). Seven of these teachers were female, and only one (Henry) was male. Table 1 summarizes the teacher characteristics in this study. The teachers taught at four different elementary schools in the same large, high-needs urban school district on the east coast of the U.S. The school district profile indicates that 86.19% of its students are non-white, 85.00% are economically disadvantaged, 14.05% receive special education services, and 10.47% are English language learners. Regardless of student diversity, state standardized test scores in mathematics indicate that all of these schools were above the overall school district average.

3.2. Instructional tasks

Each of the eight teachers taught four lessons ($n = 32$) that were part of their existing textbooks. It happened to be the case that for each grade level, one teacher used *Investigations* while the other used *Go Math* (see Table 1). *Investigations* is an NSF-supported curriculum which focused on student explorations. By the time of videotaping, teachers had received supplemental lessons from the textbook publisher to align with the Common Core. *Go Math* was a new textbook series adopted by the school district. This textbook series was developed based on the Common Core and thus incorporated recent research assertions (e.g. the use of schematic diagrams/bar model). Regardless of textbook differences, all lessons selected for this study either explicitly or implicitly involved the concept of inverse relations. Table 2 summarizes the detailed structure that guided lesson selection.

Table 2. The structure that guided lesson selection.

Additive Inverse		Multiplicative Inverse	
G1	• Fact family (or related facts) • Find the missing number • Using addition to compute subtraction • Initial unknown problem (to find how many initially)	G3	• Fact family (1) • Fact family (2) • Using multiplication to compute division • A topic suggested by teachers
G2	• Comparison word problem (1) – find the difference • Comparison word problem (2) – find the large or small quantity • Using addition to check for subtraction • A topic suggested by teachers	G4	• Comparison word problem (1) – find how many times • Comparison word problem (2) – find the small or large quantity • Using multiplication to check for division • Two-step word problems

As indicated by Table 2, teachers in G1 and G2 taught additive inverse lessons, while teachers in G3 and G4 taught multiplicative inverse lessons. Both part-whole and comparison word problems were included because these are major problem structures (Ng & Lee, 2009) that can be used to facilitate inverse relations (Carpenter et al., 1999). Lesson topics included fact family, finding the missing number, using inverse operations to compute (or check), initial unknown problems, comparison word problems (e.g. find the difference, find the large/small quantity), and two-step word problems where the solutions steps indicate inverse relations. These topics were recommended by the literature (e.g. Baroody, 1999; Baroody et al., 2009; Carpenter et al., 2003; Ding 2016; Nunes et al., 2009; Resnick et al., 1987; Torbeyns et al., 2009) and available in both textbooks.

3.3. Data sources

All 32 lessons were videotaped using two different cameras. One camera followed the teacher during instruction, while the other was focused on the students. The lessons lasted between 34 and 84 min. All lessons were first transcribed by the trained student workers of the project. Next, the first author watched all videos and commented on all transcripts about teachers' use of worked examples, representations, and deep questions. To obtain a more accurate measure, we further analyzed each lesson using the coding framework detailed below.

3.4. Coding framework

The coding framework was first modified from our prior study on teachers' lesson planning (Ding & Carlson, 2013). Because a lesson plan is closely related to the actual teaching (Stein et al., 2007), we found it is feasible to use this rubric as a basis to code the enacted lessons. This coding framework (see Table 3) contains three large categories aligned with the conceptual framework of this study – worked examples, representations, and deep questions – each containing two subcategories (worked examples and practice problems, concrete representations and abstract representations, deep questions and deep explanations). A 0–2 scale was used to score each of the six subcategories for a lesson (the total full score is 12). A score of '0' denoted that the subcategory was not addressed well, with a '2' representing fully addressed. For a teacher who's lesson was scored the full scores (12), we would expect that teacher to spend sufficient time discussing worked examples with relevant practice problems connected to those worked examples. We also expected that discussions, especially those surrounding worked examples, would be situated in concrete contexts to support sense-making. Meanwhile, we expected concrete representations to be linked to abstract representations to promote explicit understanding. Finally, we expected that a teacher who received a full score, would ask deep questions to elicit students' deep explanations.

The above coding framework was further validated through two coders' independent scoring of a few lessons. A comparison of these codes informed further revisions. For instance, a score of 2 in the category of worked examples initially stated, 'Worked examples are sufficiently discussed with the underlying ideas made explicit'. However, one coder pointed out that when the underlying idea was stressed, the scores of some participants would be lowered twice for the same reason because the underlying idea was also emphasized in deep questions. Based on this discussion, we removed the requirement of making the underlying idea explicit for worked examples. As long as a teacher spent sufficient time discussing at least one worked example, we acknowledged that the teacher paid attention to the worked example effect. Our rationale of focusing on time is that, it is hard to code a teacher's 'attention' to a worked example. However, the overt behaviour, instructional time spent on a worked example(s), at least indicates a teacher's attempt. We acknowledge that the length of a worked example does not necessarily suggest the 'quality' of its discussion. However, rushing through a worked example is arguably not desirable. In addition to the revision of the worked example category, we added a bullet to the description of a score of 1 in the explanations category: 'Teacher directly provides deep explanations'. We believed this was better than 'No deep explanations or teachers provide little or surface explanation' (a score of 0) but worse than 'Students provided deep explanations or a teacher rephrased student explanation to make it deep' (a score of 2).

3.5. Data coding and analysis

The first author used the finalized coding framework to score each of the 32 lessons. Rationale for each score was documented through comments. Typical examples of coding is reported in the Results section. Two challenges occurred during the coding process which were resolved through discussions within the research team. First, when coding worked examples of a lesson, what it meant by 'sufficient' was subjective. After discussion, we decided to justify our decision based on our sense referring to the length, focus, and completeness of class discussion of a worked example(s). We further recorded the time spent on each worked example to confirm our 'sense'. As reported later, the average time on each worked example was about 8 min. We were aware that '8 min' was not a magical number to follow. However, our video observation confirms that when a teacher spent about 8 min or more, they were generally able to provide a complete discussion involving representation uses and student discussion of the worked example, which may arguably have potential to generate some worked example effect.

The other challenge that occurred during coding was that we found that the three categories cannot be completely separated in a lesson, (e.g. A teacher may ask a question about representations when teaching a worked example). For the purpose of this research, we decided to focus on different aspects of each category. Consider, for example, a teacher

Table 3. A coding framework for videotaped lessons.

Category	Subcategory	0	1	2
Worked Examples	Example	Examples and guided practice cannot be differentiated.	Worked examples are discussed in a brief manner.	Worked example is sufficiently discussed.
	Practice	Practice problems have no connection to the worked examples.	Practice problems have some connections to the worked example.	Practice problems have clear and explicit connection to the worked example.
Representations	Concrete	Discussions, especially of worked examples, are completely limited to the abstract. No manipulatives, pictures, or story situations are used.	• Concrete contexts (e.g. word problems) are involved but not utilized sufficiently for teaching the worked example; • Semi-concrete representations such as dots or cubes are used as a context for teaching the worked example	Discussions, especially of worked examples, are well situated in rich concrete contexts (e.g. pictures and word problems). Concrete materials are used to make sense of the target concepts.
	Abstract	Discussions are limited to the concrete and are not at all linked to the abstract representations of the target concept.	• Both concrete and abstract representations are involved but the link between both is lacked; • Since all discussions remain abstract, the link between the concrete and abstract is invisible; • Opposite: from abstract to concrete.	Concrete representations are used to purposefully link the abstract representations to the target concept.
Deep questions	Question	No deep questions are asked when discussing a worked example or guided practices.	Some deep questions are posed to elicit deep explanations.	Deep questions are sufficiently posed to elicit student explanation of the targeted concepts.
	Explanation	• No deep student explanations are elicited. • Teacher provides little or surface explanations.	• A few deep student responses are elicited. However, most of the student explanations still remain at a surface level. • Teacher rephrases students' explanations without promoting to a higher level. • Teacher directly provides deep explanations.	• Deep student explanations are elicited. In particular, these explanations are related to the target concepts. • Teacher rephrases student explanations to make them deep.

who situated the new teaching in a story context which was faded into numerical solutions. We would score both concrete and abstract representations as 2. However, if this teacher failed to ask deep questions to elicit students' deep explanations of these representations, we would deduct points from the subcategories of 'deep questions' and 'deep explanations' rather than from 'representations'.

To check for reliability, another author who was familiar with the coding rubric independently coded 20% of the lessons ($n = 6$). Results were compared with those of the first author. Among the 72 codes, 8 were different, resulting in an inter-rater reliability of 88.9%. In addition, the first author re-scored all the 32 lessons one year later and checked the latest codes against the initial records. Among 192 individual codes (8 teachers × 4 lessons/teacher × 6 codes/lesson), only 16 were changed, an intra-rater reliability of 92%. Disagreements were resolved before data analysis. The scores from the 32 coding sheets were then compiled using an Excel spreadsheet.

To obtain enriched understanding, the first author conducted systematic inspections of all lessons documenting the lesson features. An Excel spreadsheet was used to record the total length of each lesson, the number of worked examples, and the length of each worked example. A second Excel spreadsheet was used to document the types of representations and connections that were observed between concrete and abstract. Finally, a third Excel spreadsheet was used to list teachers' typical 'deep questions' and corresponding students' explanations. Comments on missed questioning opportunities were also recorded. After all of these fine-grained analyses were completed, we checked these observations against teachers' lesson scores to ensure consistency. We also identified typical instructional episodes that illustrated teachers' successes and challenges in terms of each cognitive instructional principle. Finally, we examined teachers' lesson scores and their textbooks uses (either *Go Math* or *Investigations*) to identify if there was a pattern between them.

4. Results

In this section, we first report the sampled teachers' overall performance based on their lesson scores. We then report features of teachers' use of worked examples, representations, and deep questions.

4.1. An overview of teaching performance

Table 4 indicates each teacher's average video scores across four lessons. As previously introduced, the full score for each sub-category was 2, totalling up to 12 for each lesson. Despite the fact that all participating teachers were considered as experienced teachers in this study, there was clear variation among teachers' video scores, ranging from 6.75–11.25. The average score was 8.9 out of 12. Three teachers (Ann, Bea, and Gia) were scored at 7 or lower while the rests scored at 9 or above; a third grade teacher (Emily) obtained the highest score. Across the six sub-categories, the lesson scores indicate that teachers performed best in using worked examples ($M_{\text{examples}} = 1.78$, $M_{\text{practice}} = 1.84$) but worst in asking deep questions ($M_{\text{questions}} = 1.16, M_{\text{explanations}} = 0.91$). Teachers' average performance on representation uses fell in between ($M_{\text{concrete}} = 1.63$, $M_{\text{abstract}} = 1.59$).

In comparison with our prior study (Ding & Carlson, 2013) where the same six subcategories were used to evaluated teachers' initial lesson plans, findings in our current study appeared to be more positive (see Figure 1). Except for the average score of 'question' was slightly lower, the other sub-categories in the current study were all higher. Despite this positive observation, the overall pattern in using worked example, representations, and deep questions indicate similar overall successes and challenges as in our prior study.

Our further comparison between teachers' lesson scores and their textbooks used, however, did not show a clear pattern. Even though the average lesson scores of teachers who used *Investigations* ($M = 9.4$) is higher than those who used *Go Math* ($M = 8.4$), each

Table 4. Individual teachers' average video acores across four lessons.

Teacher	Grade	Example	Practice	Concrete	Abstract	Question	Explanation	Total	Textbook
Ann	G1	1.75	2	1.25	1.25	0.5	0	6.75	Go Math
Bea	G1	1	1.5	1	1.25	1.25	1	7	Investigation
Carla	G2	2	1.75	1.75	1.5	1	1	9	Investigation
Daria	G2	2	1.75	2	2	1	1	9.75	Go Math
Emily	G3	1.75	2	2	2	1.75	1.75	11.25	Investigation
Faith	G3	2	2	1.75	1.75	1.5	1	10	Go Math
Gia	G4	1.75	1.75	1.5	1	0.75	0.25	7	Go Math
Henry	G4	2	2	1.75	2	1.5	1.25	10.5	Investigation
Average		**1.78**	**1.84**	**1.63**	**1.59**	**1.16**	**0.91**	**8.9 (out of 12)**	

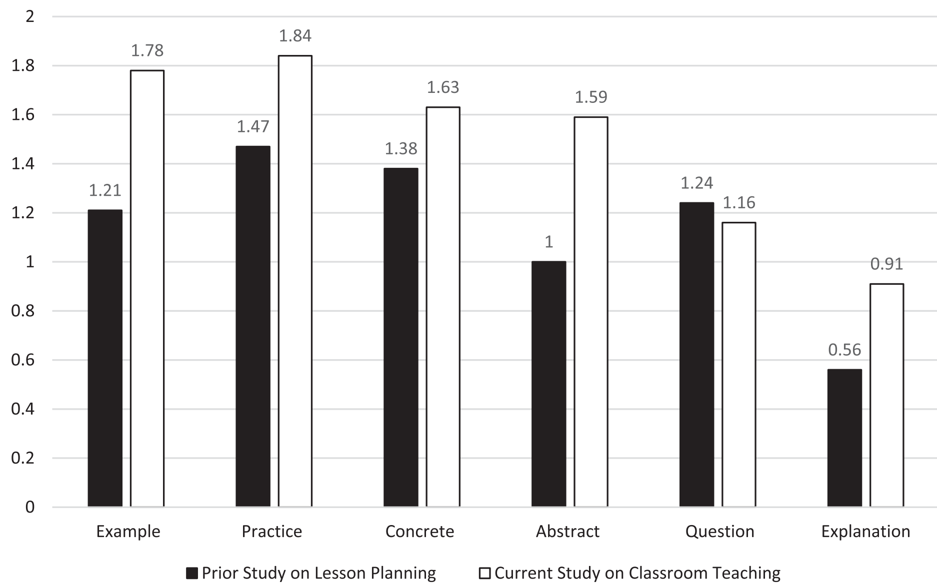


Figure 1. A comparison of teacher scores on six sub-categories in current and prior studies.

textbook was associated with both high and low scores (see Table 4, last two columns). This seems to make sense because there were diverse styles in textbook implementation (Nicol & Crespo, 2006) and an experienced teacher could enhance the cognitive demand of a textbook regardless of its quality (Hassler, 2016). Given that teachers' textbook use is beyond the scope of this study, we focus our report below on the instructional features in terms of the targeted cognitive principles.

4.2. *The use of worked examples*

4.2.1. *Attending to worked examples*

Results show that most teachers in this study attended to worked examples before asking students to practice ($M_{\text{examples}} = 1.78$), which is more positive than other classrooms where minimum guidance was provided (Kirschner et al., 2006). In this study, four teachers were judged as spending sufficient time on at least one worked example in all lessons, three teachers spent sufficient time in most lessons (75%), and only one (Bea) never discussed any worked example in a sufficient manner. Teacher Bea used the inquiry-based textbook, *Investigations,* which suggested games for each lesson. During the worked example time, Bea mainly explained how to play the games which often went quickly. However, 'textbook' seemed not necessarily to be the factor that hindered teacher's use of worked examples; in our study, the other three teachers (Daria, Emily, Henry) who also used *Investigations* were able to set up plenty of time for worked examples before students' own exploration.

Recall that our evaluation of teachers' use of worked examples focused mainly on 'quantity'. Taking the length of worked examples into consideration, we found that among 32 lessons (1842 total minutes) about 33% of class time was devoted to worked examples (603 min). Overall, there were 76 worked examples with an average of 8 min per example ($603/76 \approx 7.9$). As acknowledged, there was no magic number that indicated teachers' sufficient use of a worked example. However, in hindsight, our video observation confirmed that when a teacher spent about 8 min on one worked example, the process of unpacking an example task was generally complete. Otherwise, the discussion appeared to be rushed.

Taking the number of worked examples into consideration, we found there were variations across lessons, ranging from 1 to 6 worked examples. Specifically, 25% of the 32 lessons ($n = 8$) discussed one worked example. The rest of the lessons contained multiple examples, with 38% involving 2 examples, 35% involving 3–4 examples, and 2% involving 5–6 examples. For instance, while Emily used 14 min to discuss one example, Ann used 19 min to cover 6 examples, resulting in an average of 3 min per example (none reached 8 min). In this case, we scored Emily's use of worked example as 2 and Ann's as 1. Given that 75% of the lessons discussed multiple examples, although not factored into our scoring, we inspected the nature of these example tasks. We noticed that when multiple worked examples were used, the nature of the examples was often repetitive. For instance, Ann's lesson 1 included four worked examples of fact family, which involved different number pairs (8, 4, and 12; 6, 3, and 9; 7, 5, and 12; 2, 3 and 5) and she only guided students to find the answer for each task without involving additional conceptual aspects. The repetitive nature of worked examples within a lesson was indeed common in this study.

4.2.2. Linking to practice problems

Practice problems in most lessons were found to be relevant to worked examples ($M_{\text{practice}} = 1.84$). However, in a few lessons, the practice problems were irrelevant tasks, which may not help reinforce what was taught through the worked examples. For instance, the worked examples of Bea's lesson 2 focused on fact family. Using a ten-frame, the class generated four worked examples involving numbers (a) 7, 3, 10, (b) 5, 5, 10, (c) 10, 4, 14, and (d) 12, 2, 14, respectively. However, the practice problem was only to play the game of 'tens go fish' in pairs (see Figure 2, Bea). That is, if a student pulled an '8' card, the other student should give him/her a '2' because 8 plus 2 makes 10. Mathematically, the goal of the game was to find the missing number such as $8 + () = 10$, which is a different task from a 'fact family' (Ding, 2016). To solve $8 + () = 10$, a student may need to retrieve and apply their prior knowledge of '$10–8 = 2$', which is a harder process than seeing a given fact family. Another reason of difference is that to find missing numbers like $8 + () = 10$, a student does not necessarily need to use inverse relations. As seen from Bea's lesson 2, some students used strategies such as counting up from 8 to 10, which involved direct rather than inverse thinking (Ding & Auxter, 2017). In this sense, the worked example and practice problems in Bea's lesson 2 was not fully connected (scored a 1). Likewise, Carla's two worked examples in lesson 2 were initial unknown problems. However, none of her three practice problems was directly aligned with the worked examples.

4.3. The use of representations

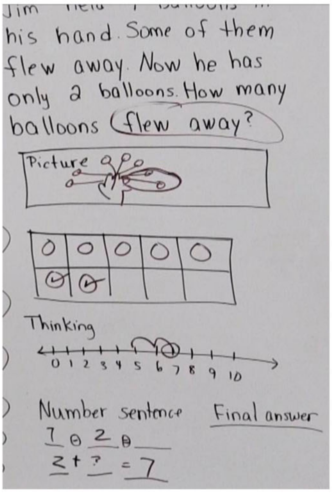
4.3.1. Using concrete representations

Teachers' use of concrete representations in this study was overall encouraging ($M_{\text{concrete}} = 1.63$). Concrete (or semi-concrete) representations in this study included using manipulatives (e.g. fingers, cubes, dominos, base-ten blocks), charts (e.g. fact triangles, ten-frame sheets, number grid), diagrams (e.g. number lines, tape diagrams, pictures), and story problems (see Figure 2). The use of multiple representations was a common feature across lessons. Typically, there were 3–4 different types of concrete representations used in each lesson. Students in all classrooms were allowed to choose their own manipulatives (e.g. cubes, blocks, fingers) to solve inverse relations problem. Some teachers (e.g. Carla) asked students to choose their favourite animals or food to create story problems based on a given bar model.

Encouragingly, the majority of the teachers situated worked examples in real world contexts (scored a 2 for concrete representations). This included all four of Daria and Emily's lessons (100%), three of the four lessons (75%) from Carla, Faith, and Henry's, two of Gia's lessons (50%) and one of Ann's lessons (25%). Figure 2 illustrates typical examples including Ann and Carla's balloon problem, Daria's soccer ball problem, Emily's robot-hands problem, Faith's bagel problem, and Henry's apple problem. As we can see from these pictures, some teachers guided students to understand the story situation by drawing corresponding pictures or diagrams, which was consistent with the IES recommendation for combining graphics with verbal descriptions. In fact, Daria also asked students to imagine the story context: 'Raise your hand if you can picture the locker room, and there is a big bag with 15 soccer balls in it' (See Figure 2, Daria). Note that the example of Ann in Figure 2 was the only case where the teacher situated her new teaching in a story context (scored a 2). However, her purpose for using these concrete representations, as indicated by her questions, was to help students seek computational answers rather than to understand the quantitative relationships. We captured this lack of depth in categories of 'deep questions' and 'deep explanations' (elaborated upon later).

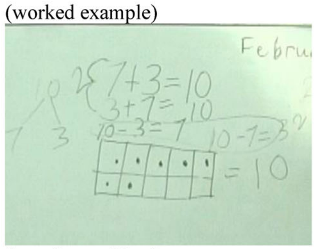
While the majority of lessons situated the teaching of worked examples in concrete contexts, the rest of the lessons limited the new teaching only to semi-concrete representations such as dominos, cubes, and dice. In fact, all four of Bea's lessons (100%), three of Ann's lessons (75%), two of Gia's lessons (50%), and one each of Carla, Faith, and Henry's lessons (25%) shared the same pattern. For these cases, 'concrete representation' were scored a 1. In Figure 2, Bea drew the aforementioned ten-frame with 7 boxes shaded, leading to a fact family ($7 + 3 = 10, 3 + 7 = 10, 10–3 = 7$, and $10–7 = 3$). However, the *Investigations* teacher's guide actually suggested a pair of story problems about pencils involving numbers of 7, 3, and 10 (the first problem is solved by $7 + 3 = 10$ and the second is solved by $10\text{-}3 = ?$). The teacher guide also reminded teachers, G1

Ann



Bea

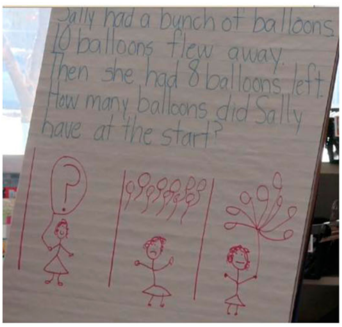
(worked example)



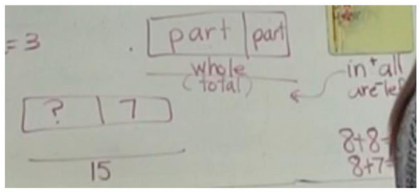
(practice problem)

G2

Carla

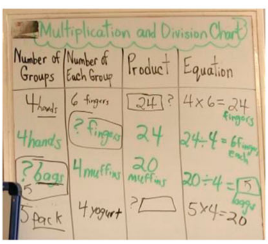


Daria



G3

Emily



Faith

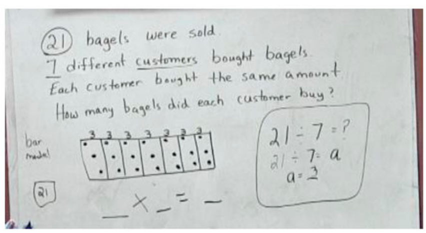


Figure 2. Teachers' use of concrete representations.

> After posing the second problem, ask students to consider how the problem they have already solved might help them solve the new problem. Keep in mind that for many first graders, the second problem will seem like a new, unrelated problem.

Unfortunately, this teacher did not include these story problems in her teaching. Bea could have started with the textbook suggested story context, which could have worked together

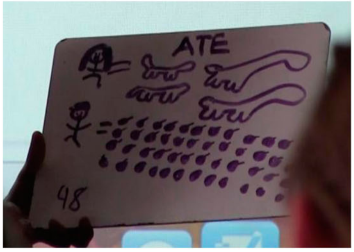
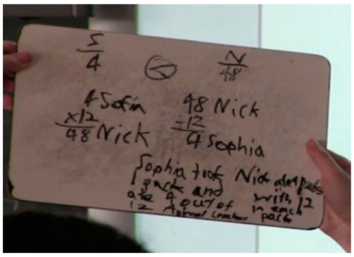
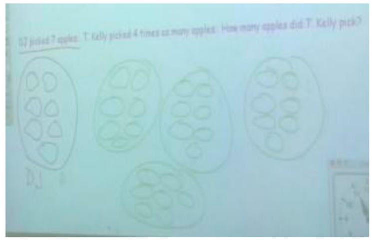
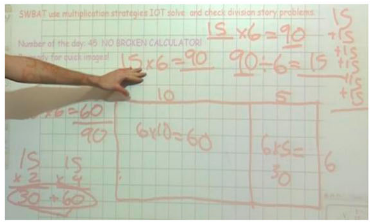
G4	Gia   	Henry Lesson 1  Lesson 4 

Figure 2. Continued.

with her ten-frame to help students make sense of the inverse relations. Likewise, Carla in her first lesson only asked students to use cubes to illustrate a fact family that served as a worked example. Also, Faith and Henry only used array models to teach multiplicative inverses in one of their lessons, which resulted in a lack of contextual support (see Figure 2, Henry for an example).

4.3.2. Making connections between concrete and abstract

In this study, teachers in the majority of their lessons made rich connections between concrete and abstract ($M_{\text{abstract}} = 1.59$), which often indicated concreteness fading (Fyfe et al., 2015; McNeil & Fyfe, 2012). This included all four (100%) of Daria, Emily, Faith, and Henry's all lessons, two (50%) of Bea and Carla's lessons, and one (25%) lesson from both Ann and Gia. Ideally, a teacher could ask questions to explicitly request connections between concrete and abstract. However, during our scoring, even if a teacher did not ask a question about the representational connections, we assumed that the same context (e.g. bagels, robots) would assist students to sense the connections. For instance, in Figure 2, Faith drew a bar model to transition from the word problems (21 bagels equally shared with 7 customers) and the corresponding number sentences ($\_\_ \times \_\_ = \_\_$, and $21 \div 7 = ?$). In this case, we scored 'linking to abstract' as 2. Similarly, Emily connected concrete and abstract representations using the multiplication and division chart (see Figure 2). Emily first discussed a multiplication problem: '*A robot has 4 hands. Each hand has 6 fingers. How many fingers does the robot have altogether?*' She guided the class to use the chart to identify the role of each quantity (e.g. '4 hands' indicated the number of groups), which led to the associated equation '$4 \times 6 = \square$'. Next, she requested that the students change this problem to a related division problem which was not proposed by her textbook, *Investigations*. Using the same chart, the class analyzed the roles of quantities in the division problem, which led to numerical solutions (See Figure 2). Emily's recording of this pair of inverse word problems using the chart indicated her structural awareness of inverse relations (number of groups $\times$ number in each group $=$ product/total, total $\div$ number of groups $=$ number in each group).

In contrast, lessons that did not receive a full score for making concrete-abstract connections demonstrated three issues. First, a teacher may have limited the instruction to semi-concrete representations. Bea's first lesson taught a game named five in a row with subtraction. Throughout this lesson, students only played game cards and there was not a single number sentence written. As such, we scored her linking to abstract as 0. The second case was due to an opposite representational sequence ranging from abstract to concrete. As mentioned above, Carla's first lesson started with a fact family ($7 + 1 =$, $8–1 =$, $8–7 =$, $1 + 7 =$). After discussion of the answers, she asked students to use cubes to demonstrate the fact family. In this case, the sequence went from abstract to semi-concrete. Therefore, we scored it a '1'. Finally, there were lessons that contained both concrete and abstract presentations simultaneously. However, the connection between concrete and abstract was too distant for students to discover. Figure 2 illustrated one example. In Gia's lesson, she requested that half of the class draw pictures (concrete) while the other half write numerical sentences (abstract) for the statement, 'Sophia had 4 crackers, and Nick had 12 times hers' (scored a '2' for concrete). She then asked four students to explain their drawings or their number sentences in front of the class. (See Figure 2, Gia). However, connections between the concrete and the abstract representations were never made in this class. Since there was not a process like 'concreteness fading' that may have focused students' attention to the connections, we scored the category of linking to abstract representation as a '1'.

4.4. The use of deep questions

4.4.1. Asking deep questions

Among the three instructional principles analyzed, teachers' performance on deep questioning demonstrated the most challenges ($M_{\text{questions}} = 1.16$). Among 32 lessons, only 9 lessons (28%) taught by four teachers (Bea, Emily, Faith, and Henry) received a full score of 2. In contrast, three lessons (9%) taught by two teachers (Ann and Gia) received a score of 0 due to a complete lack of deep questions. The other 20 lessons (62.5%) across all eight teachers received partial scores.

The main issue of teacher questioning was its focus on computational answers, which did not demand deep questions. For instance, in Ann's Lesson 4, one student suggested '7–2 = I don't know' and another student suggested '2 + I don't know = 7', providing a great opportunity to address inverse relations (See Figure 2, Ann). However, Ann only reminded the class that these two number sentences were a 'fact family' and proceeded with finding the answer for 'I don't know' rather than uncovering the underlying inverse relation. Another issue that diminished the need to ask deep questions was teachers' instruction to look for key words. For instance, Ann reminded students, 'There is a word in the story to help you decide if you add or subtract. Pick one word from this story'. Similarly, Daria guided students to look for the key words 'in all' for addition. Gia also reminded students, 'Here, look at the wording, look at the wording. What operation might you use? The wording should help you figure it out'. Even though the use of key words has been criticized for a long time (e.g. Nesher & Teubal, 1975), our findings show that this practice frequently occurs even in experienced teachers' classrooms.

Since asking deep questions appeared to be a challenge for many teachers, we inspected the features of those available deep questions in this study. First, we found that these questions stressed mathematical concepts (e.g. meaning of operations). One example was Henry's lesson that implicitly involved multiplicative inverses based on comparison problems. When discussing a multiplication problem – *DJ picks 7 apples. Teacher Kelly picked 4 times as many apples. How many apples did Teacher Kelly pick?* – Henry drew a diagram about apples (see Figure 2, Henry) and asked a series of questions focusing on the concept of 'times':

Excerpt 1:
T: What does 'times' mean?
S1: Uh, it means that when she has 7 more but 4 times.
T: What do you mean, she has 7 more but 4 times?
S1: Like, she has 7 more and then she has another 7 more, so it's like . . .
T: Call on someone to help you out to clarify your thinking.
S2: Can I give an example?
T: Please.
S2: Do you see how you have 7 apples?
T: I do see I have 7 apples. That's my favourite number.
S2: You just add on 4 more like. They saying like you're adding on 4 more bags of 7 apples.
T: Well, what does it mean that I have to add on 4 more bags of 7 apples?
S2: Because it says 4 times.
T: Okay, because it says 4 times, but why do the bags have to have 7?
S2: Because of the number that you already have, that's like the . . .
T: Ah, because of the number I have already. . . . Okay, because it's 4 times as many apples as I already picked. So she has to have 4 groups, with that same number inside of it. So now, who can tell us an equation that can represent how many apples teacher Kelly picked?

In Excerpt 1, Henry grasped the word 'times' to help student understand the concept of multiplicative comparison. According to the literature, comparison problems are challenging because they deal with relationships that are hard to manipulate (Nunes et al., 2009). Yet, multiplicative comparisons can be referred to the basic meaning of multiplication (the equal groups) to make them more understandable (Carpenter et al., 1999). In the above excerpt, with continuous prompts, students were able to explain '4 times' using their own words: 'Like, she has 7 more and then she has another 7 more, so it's like . . . ' 'Do you see how you have 7 apples? You just add on 4 more like. They saying like you're adding on 4 more bags of 7 apples'. These responses indicated students' linking between the concept of 'times' and their prior knowledge of the equal groups meaning.

The second feature of deep questions was related to stressing of quantitative relationships. As previously mentioned, Carla taught a lesson about an initial unknown problem: 'Sally had a bunch of balloons. 10 balloons flew away. Then she had 8 balloons left. How many balloons did Sally have at the start?' (see Figure 2, Carla.) This type of problem is challenging because the story describes a decreasing situation, yet the solution demands the use of addition (Nunes et al., 2009). Children who seek key words (e.g. flew away) would likely use subtraction (10–8 = 2) to answer this problem. To help students understand the quantitative relationship, Carla guided the class in drawing the problem's initial situation, change, and result (see Figure 2). During student group work, she spent about 5 min with a student who used cubes but suggested 10–8 as a solution for this problem. Pointing to the 10 cubes which represented the balloons that had flown away, Carla asked, 'So did these balloons fly away? (Mhmm). Okay, so did she have this many in the beginning?' This question prompted the student to focus on the relationship among 'initial', 'change', and 'result'. With continuous guidance, this student understood that the 10 balloons that flew away were part of the initial total amount. Thus, adding the balloons that had flown away and the ones remaining in Sally's hand would give him the answer to how many balloons Sally had at the beginning. Because Clara only asked deep questions to an individual student rather than the whole class, we scored this category a 1. Perhaps if the teacher had asked these questions to the whole class, this may have led to the following relationship being more explicit for more students: the flown-away + the left over = the initial total.

An interesting finding in this study was related to teachers' questions on representations, which revealed a dual purpose – targeting both quantitative relationships (or concepts) and computational answers. As seen in Figure 2, Daria began her lesson by drawing the textbook-suggested bar model on the board. She then highlighted the part-whole relation and indicated that her purpose for using the bar model was to understand the inverse relations. Later in the lesson, she read the class a pair of inverse word problems which were represented by the bar models. Next, she suggested the solutions 8 + 7 and 15–7, and shifted discussion towards computational strategies (e.g. double plus 1, using cubes). It was not until one student reported that she did not need to use the cubes to compute the subtraction problem that Daria oriented the class conversation back to the quantitative relationships: 'How many people didn't need to use the cubes either because they realized something?' One student referred to the two bar models. Daria then asked, 'How many people noticed that both bar models had the same numbers but they just had a different one missing?' This question elicited another student comment that this pair of bar models was just like the 'fact triangles' they learned before, indicating the student's realization of the inverse relation. The above dual purpose of questioning on representations was common across lessons.

4.4.2. Eliciting deep explanations

Among six sub-categories, 'deep explanation' was scored the lowest in this study ($M_{explanations} = 0.91$). Among 32 lessons, only 4 (12.5%) taught by two teachers (Emily and Henry) received a full score. In contrast, eight lessons (25%) taught by two teachers (Ann and Gia) completely missed deep explanations. The rest of the 20 lessons (62.5%) taught by six teachers received a partial score. In other words, deep explanations were largely missing in most classrooms. One of the reasons may be related to the lack of deep questioning in most lessons as reported above. Excerpt 2 demonstrates students' limited thinking when deep questions were missed (scored a 0 for 'explanation'). After presenting the set of representations on the board (see Figure 2, Ann), Ann had the following conversation with the students:

Excerpt 2:
T: All right. I need your final answer on your slate.
(Students raise their slates, and the teacher looks at them.)
T: (Goes back to the board and points to the unchecked circles on the ten frame.) Count for me!
S: 1, 2, 3, 4, 5.
T: That way gives me 5.
T: (Points to the uncircled balloons in the picture.)
S: (Students count.) 1, 2, 3, 4, 5.
T: That way gives me 5.
T: (Points to '5' on the number line.) I subtract 2 from 7, and the answer is?
S: 5!
T: So the answer must be 5. (Writes '5' under 'Final answer' and fills the number sentences with '5').

All these representations in Excerpt 2 indicate the part-whole relationship. However, Ann did not ask a single question about the part-whole structure. As a result, students merely counted from 1 to 5 and computed the answer for 7-2. Ann could have asked students why they could count in certain ways and why both addition ($2 + ? = 7$) and subtraction ($7–2 = ?$) solved this problem, as well as why/how these two number sentences were related to each other. Unfortunately, explanations to these deep questions were completely missed.

We also observed that teachers who did ask deep questions only anticipated surface explanations. For instance, Ann asked students how the number sentences in a fact family were related. She was satisfied by students' responding that they were the same three numbers and thus did not ask follow-up questions to orient students' attention to the part-whole relationships. Similarly, we observed that all eight teachers encouraged multiple solutions (often multiple computation strategies) by asking questions such as, 'who has another solution?' However, after different solutions were presented, the class usually proceeded without explicit comparisons between multiple solutions in terms of structural similarities and differences. Given that comparisons are an important strategy to promote structural thinking (Kotovsky & Gentner, 1996; Star & Rittle-Johnson, 2009), the teachers in the current study could have asked follow-up comparison questions to elicit deep explanations.

A slightly better, but still problematic, situation is that some teachers did not ask deep questions but provided deep explanations themselves. In the example of Carla (see Figure 2), she drew bar models and created corresponding stories. Later, when she guided students to see the connections between these two bar diagrams, she herself started by explaining the similarities and only asked how many students also noticed those similarities. Given that students should be part of the meaning-making process (Radford & Roth, 2011), teachers' giving out their own deep explanations likely deprives students of important thinking and reasoning opportunities.

5. Discussion

This study examines how cognitive instructional principles on the use of worked examples, representations, and deep questions (Pashler et al., 2007) are used by experienced elementary teachers when teaching a critical early algebra topic, inverse relations (Baroody et al., 2009; Torbeyns et al., 2009; Vergnaud, 1988). Results show differences in teachers' instructional quality. This observation echoes Hiebert and Stigler's (2017) insight that the variation of mathematics teaching quality in U.S. classrooms is unnecessarily large. This variation, however, has enabled us to identify teachers' successes and challenges in using worked examples, representations, and deep questions, which is informative for the fields of mathematics education and cognitive research for better supporting teachers' development of students' algebraic and overall mathematical thinking. Below, we discuss findings under each category. Due to the interconnected nature of these cognitive aspects, our discussions may occasionally relate all instructional principles together.

5.1. Unpacking one worked example sufficiently

The literature notes that worked examples are rarely utilized in US classrooms because teachers often present problem solving tasks for small group explorations with minimum teacher instruction (Kirschner et al., 2006; Stigler & Hiebert, 1999). In this study, the sampled teachers generally embraced the use of worked examples as measured by the time spent on them (on average, 33% of class time). This is encouraging due to the well-known worked example effect in developing students' mental schema to enhance problem solving (Sweller, 1988; Sweller & Cooper, 1985). However, it should be noted that our coding of worked examples in this study focused mainly on the 'quantity' rather than 'quality'. While some teachers spent a good amount of time discussing worked examples, some teachers rushed through several repetitive examples, resulting in limited time on each example. This is similar to our prior finding about teachers' lesson planning (Ding & Carlson, 2013) where teachers seemed to hold a common belief that if students don't understand the mathematical idea, they should see more examples until comprehension is reached. We argue that if a teacher does not effectively guide the class to unpack one worked example to establish the relevant mental schema (Sweller & Cooper, 1985), presenting many repetitive examples would still not help. Based on these findings, we propose that instead of presenting 2–3 extra short, repetitive worked examples, a teacher may cut off the extra examples to save time for the discussion of the main worked example. This way, the underlying concepts of the worked example may be adequately discussed to establish the targeted schema. As Stigler and Hiebert (1999) reported, Japanese teachers only discuss one problem-solving task in a lesson, which may even extend into follow-up lessons.

Now, the question may become, how can a teacher spend a long chunk of time discussing only one worked example? As reported in this study, only 25% of the 32 lessons discussed a single worked example. Effectively unpacking a single worked example may, therefore, be a common challenge faced by teachers. Findings from this study suggest that teachers may successfully use representations and deep questions to unpack worked examples. In other words, it is not enough to simply encourage teachers to teach worked examples. Rather, teachers need to know 'how' representations and deep questions may be incorporated in the process of unpacking a worked example. We will discuss these two aspects in later sections.

Additionally, our findings reveal a disconnect between worked example and practice problems in some classrooms. The purpose of teaching worked examples is to enhance the follow up problem solving that share structural connections with the worked examples (Catrambone & Yuasa, 2006; Sweller & Cooper, 1985; van Gog et al., 2011; Zhu & Simon, 1987). In this study, we notice that some teachers provided problems that share only surface but not structural connections, which may not advance the worked example effect. One of the reasons is that the teachers simply followed textbook presentation. This calls for attention of curriculum support in stream-lining the worked examples and practice problems to ensure students' deep learning of the core concepts. This also calls for knowledge support for teachers in enhancing their understanding of structural connections of relevant tasks.

5.2. Using concrete support with concreteness fading for meaning-making

In this study, many teachers situated new teaching of inverse relation in a real-world context, which is encouraging. This contextual support in learning early algebra is important because these concepts generally refer to mathematics structures that are distant from students' real-world experiences (Kaput, 2008; Kieran, 2018). Situating the new teaching of inverse relation in concrete situations may, therefore, help activate students' prior knowledge for sense-making (Gerofsky, 2009) and help address the limitation of focusing on number manipulation (Baroody, 1999; Torbeyns et al., 2009). In this study, concrete situations used by teachers may or may not have come from the textbooks. In fact, when the textbooks provided relevant real-word contexts, some teachers did not necessarily use those contexts but rather encouraged students to create story problems involving their own favourite objects. Moreover, there were teachers who asked students to create story situations that was not outlined by the textbooks. Likely, teachers' use of concrete contexts is related to their beliefs in the role of concrete contexts in supporting students' learning. Of course, we found that some teachers still limited the discussion of inverse relations to semi-abstract (e.g. dominos) and abstract (e.g. fact triangle) representations. This echoes the aforementioned instructional shortcoming about number manipulation (Baroody, 1999; Torbeyns et al., 2009).

For connection-making between concrete and abstract, some teachers employed the method of concreteness fadi2019ng, first starting with a word problem context which was modelled through drawings (circles, tapes, tallies) and further faded into number sentences. Such a sequence is recommended as a way to connect concrete and abstract representations (Pashler et al., 2007) and is found to be supportive of students' learning and transfer (Fyfe et al., 2015; Goldstone & Son, 2005; McNeil & Fyfe, 2012).

To facilitate connection making, teachers' use of bar model or tape diagrams (Murata, 2008) is promising. The tape diagram is a linear model that can effectively illustrate the embedded quantitative relationships and serve as a transition from concrete to abstract (Murata, 2008; Pashler et al., 2007). This type of schematic diagram has been widely used in East Asian countries (Cai & Moyer, 2008; Ding & Li, 2014; Murata, 2008) and has been found to be effective in supporting student learning (Ng & Lee, 2009). Therefore, the Common Core State Standards expect elementary students to learn to use them for problem solving (CCSSI, 2010). In this study, seven of eight (87.5%) teachers used tape diagrams during instruction. This is likely related to textbook influence as *Go Math* contained many more bar models than *Investigations*.

Despite these encouraging observations, our findings reveal that teachers used concrete representations and linear models (e.g. tape diagrams) with different purposes. Elsewhere (Ding et al., 2019), we reported detailed cross-cultural difference in teachers' representational purposes between U.S. and Chinese teachers. Our U.S. teachers often used linear models to find answers, which is indicated by their questions. As aforementioned, our coding framework aims to separate these categories, yet, actual classroom teaching shows that teachers' representation uses and questioning often go together. It seems that merely asking teachers to use real-world context, tape diagrams, and concreteness fading cannot guarantee the depth of teaching. Teachers ought to ask deep questions to elicit students' explanations of the deep meaning embodied by these representations. We discuss this point further in the next section.

5.3. Asking deep follow-up questions involving comparisons

To improve students' learning, deep questions are recommended for classroom instruction (Pashler et al., 2007). Research on teacher questioning has a long history (Gall, 1970; Stevens, 1912). Pomerance et al. (2016) also pointed out that asking deep questions was the only IES recommendation frequently stressed by current textbooks in teacher education programmes. Surprisingly, our classroom findings indicate that asking deep questions is the weakest area among the three aspects investigated. Consequently, both the quantity and the quality of students' deep explanations are discouraging. When teachers did ask deep questions, we found that they focused on either the concepts or the quantitative relationships. As reviewed, early algebra refers to fundamental mathematical principles, relations, and structures which are abstract in nature (Kaput, 2008; Kieran, 2018). Thus, it is important for teachers to ask deep questions to orient students' attention and elicit their explanations of the structural aspects of a representation. Interestingly, we found teachers in this study often asked questions on representations with a dual focus, targeting both quantitative relationships/concepts and computational answers, which is certainly better than a pure procedural focus on computational answers. However, the above dual focus on both quantitative relationships and computational answers is still different from our recent finding on Chinese teachers' sole focus on quantitative relationships (Ding et al., 2019). With the goal of developing algebraic thinking in mind, we argue for the importance of focusing on quantitative relationships/concepts over computational answers. This may help overcome another instructional limitation on inverse relations, that is, focusing on strategies rather than the undergirding ideas.

Our findings also indicate that a single deep question may not necessarily elicit students' deep explanations. Teachers' persistence in asking follow-up deep questions (e.g. Henry's class), especially comparison questions, helped elicit students' deep explanations. This finding adds insights to IES recommendations (Pashler et al., 2007) and prior literature on teacher questioning (Craig et al., 2006; Franke et al., 2009). Prior research suggests that comparisons facilitate critical cognitive processes, promote relational thinking (Kotovsky & Gentner, 1996), and are effective in learning mathematics (Star & Rittle-Johnson, 2009). Unfortunately, although we found that many teachers in the current study encouraged multiple solutions, they did not follow up with deep questions to elicit students' comparisons among these solutions. Instead, they were often satisfied with students' explanations of computational procedures and then moved toward other tasks. This finding echoes prior studies (Kazemi & Stipek, 2001; Richland et al., 2007) and calls for further research on how comparisons may be incorporated in teacher questioning. To date, we have no data to explain the exact reasons for teachers' lack of follow-up deep questions including comparison questions. Perhaps teachers' own limited understanding of structural similarities and differences is a factor; perhaps teachers' concerns about student motivation (not being over-challenged) affect their classroom decisions. Future studies should explore why the practice of asking deep questions appears particularly difficult and whether the strategy of asking follow-up comparison questions serves as a path to enhance this practice, resulting in increased 'depth' of teacher-student dialogues in classrooms.

5.4. Limitations, implications, and future directions

This study examines experienced teachers' classroom teaching of early algebra focusing on inverse relations. We acknowledge that only eight teachers' 32 lessons on one early algebra topic were analyzed. Thus, findings should not be generalized. In fact, the purpose of a case study is not for generalization but a deep understanding of the targeted case (Creswell & Poth, 2018). In addition, our coding framework only contains a 0–2 scale which cannot fully capture instructional features that may fall between the scores of 0–1 or 1–2. Possibly, a more precise scale in future studies may serve this purpose better. Finally, our findings are only based on videotaped lessons without linking to student performance.

Nevertheless, findings based on this case study provide insights into field, which have implications for both cognitive and instructional research. For cognitive researchers, our study shows that even though our participants are experienced teachers, they tend to present multiple, repetitive 'worked examples', which may not ensure the worked example effect. We suggest researchers design experimental studies to test whether an alternative approach – unpacking one example in depth, as opposed to presenting multiple, repetitive examples – can better enhance students' subsequent problem solving. To unpack one worked example in depth, we propose that representations (e.g. concreteness fading) and deep questions (e.g. follow-up comparison questions) can be incorporated in the process of unpacking a worked example. Future cognitive research design may take these components into consideration.

The above findings also have implications for mathematics educators and researchers. In this study, teachers' presentation of multiple, repetitive examples may be due to their misconception, 'more examples, the better' or due to their lack of knowledge regarding how to unpack an example in depth. Mathematics educators and researchers may start with the joint aspect, 'unpacking a worked example by asking deep questions on representations' to provide professional development (PD). In fact, our findings reveal that some experienced teachers have challenges in using representations for conceptual purposes and asking follow-up deep questions like comparisons. Future PD effort may target these areas to enhance classroom teaching. Of course, it is difficult to break teachers of their teaching habits which many times might have been the way they were also taught. As such, an alternative path for future PD and educational research is to target prospective and beginning teachers as they begin to develop their teaching knowledge. For instance, case studies may be conducted on novice teachers to understand how they learn to implement the targeted cognitive instructional principles in their teaching of mathematics.

In this study, our findings do not indicate for a direct connection between textbooks and classroom instruction. However, we have observed that when textbooks incorporate research assertations (e.g. bar models), teachers likely present these powerful models in classrooms, which provides learning opportunities for students. Of course, even with the same textbook series, teachers may use them differently (e.g. Bea and Emily's different use of *Investigations*), resulting in different qualities of classroom teaching. This seems to echo the pattern reported in Hassler (2016), that is, regardless of the textbook, teachers can increase or decrease the cognitive demand of a textbook lesson during their classroom teaching. Future studies may further explore how teachers can maximize their textbook potential in terms of using worked examples, representations, and deep questions to increase students' cognitive demand and to enhance early algebraic learning. With joint effort based on detailed classroom observation, we think a science of improvement in mathematics teaching (Cai et al., 2019; Lewis, 2015) is possible.

Disclosure statement

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

Funding

This work is supported by the National Science Foundation CAREER award (No. DRL-1350068) at Temple University. Any opinions, findings, and conclusions in this study are those of the authors and do not necessarily reflect the views of the funding agency.

ORCID

Meixia Ding http://orcid.org/0000-0002-5820-0684

References

Anderson, J. R. (2010). *Cognitive psychology and its implications* (7th ed.). Worth Publishing.
Anderson, J. R., Reder, L. M., & Simon, H. A. (2000, Summer). Applications and misapplications of cognitive psychology to mathematics education. *Texas Educational Review*, 29–49. http://act-r.psy.cmu.edu/papers/misapplied-abs-ja.html
Baroody, A. J. (1999). Children's relational knowledge of addition and subtraction. *Cognition and Instruction*, *17*(2), 137–175. https://doi.org/10.1207/S1532690XCI170201
Baroody, A. J., Torbeyns, J., & Verschaffel, L. (2009). Young children's understanding and application of subtraction related principles: Introduction. *Mathematical Thinking and Learning*, *11*(1–2), 2–9. https://doi.org/10.1080/10986060802583873
Berwick, D. M. (2008). The science of improvement. *Journal of American Medical Association*, *299*(10), 1182–1184. https://doi.org/10.1001/jama.299.10.1182
Bruner, J. (1966). *Toward a theory of instruction*. Belknap.
Bryk, A. (2009). Support a science of performance improvement. *Phi Delta Kappan*, *90*(8), 597–600. https://doi.org/10.1177/003172170909000815
Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., & Hiebert, J. (2019). Research pathways that connect research and practice. *Journal for Research in Mathematics Education*, *50*(1), 2–10. https://doi.org/10.5951/jresematheduc.50.1.0002
Cai, J., & Moyer, J. C. (2008). Developing algebraic thinking in earlier grades: Some insights from international comparative studies. In C. E. Greene & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics* (pp. 169–182). NCTM.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Heinemann.
Carpenter, T. P., Franke, L. P., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic & algebra in elementary school*. Heinemann.
Carraher, D. W., & Schliemann, A. D. (2007). Early algebra and algebraic reasoning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 669–705). Information Age.
Catrambone, R., & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction*, *16*(2), 139–153. https://doi.org/10.1016/j.learninstruc.2006.02.002
Chi, M. T. H. (2000). Self-explaining: The dual processes of generating and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Erlbaum.
Chi, M. T. H., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., Levy, R., Li, N., McEldoon, K. L., Stump, G. S., Wylie, R., Xu, D., & Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, *42*(6), 1777–1832. https://doi.org/10.1111/cogs.12626
Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. https://doi.org/10.1207/s15516709cog1302_1
Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2
Chi, M. T. H., & VanLehn, K. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177–188. https://doi.org/10.1080/00461520.2012.695709
Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823
Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. http://www.corestandards.org/the-standards
Craig, S. D., Sullins, J., Witherspoon, A., & Gholson, B. (2006). The deep-level-reasoning-question effect: The role of dialogue and deep-level-reasoning questions during vicarious learning. *Cognition and Instruction*, *24*(4), 565–591. https://doi.org/10.1207/s1532690xci2404_4
Creswell, J. W., & Poth, C. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.
Ding, M. (2016). Opportunities to learn: Inverse operations in U.S. and Chinese elementary mathematics textbooks. *Mathematical Thinking and Learning*, *18*(1), 45–68. https://doi.org/10.1080/10986065.2016.1107819
Ding, M., & Auxter, A. (2017). Children's strategies to solving additive inverse problems: A preliminary analysis. *Mathematics Education Research Journal*, *29*(1), 73–92. https://doi.org/10.1007/s13394-017-0188-4
Ding, M., & Carlson, M. A. (2013). Elementary teachers' learning to construct high quality mathematics lesson plans: A use of IES recommendations. *The Elementary School Journal*, *113*(3), 359–385. https://doi.org/10.1086/668505
Ding, M., Chen, W., & Hassler, R. (2019). Linear quantity models in the US and Chinese elementary mathematics classrooms. *Mathematical Thinking and Learning*, *21*, 105–130. https://doi.org/10.1080/10986065.2019.1570834
Ding, M., & Li, X. (2014). Transition from concrete to abstract representations: The distributive property in a Chinese textbook series. *Educational Studies in Mathematics*, *87*, 103–121. https://doi.org/10.1007/s10649-014-9558-y
Ding, M., Li, X., Piccolo, D., & Kulm, G. (2007). Teacher interventions in cooperative-learning mathematics classes. *Journal of Educational Research*, *100*(3), 162–175. https://doi.org/10.3200/JOER.100.3.162-175
Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Education Studies in Mathematics*, *61*(1–2), 103–131. https://doi.org/10.1007/s10649-006-0400-z

Franke, M. L., Webb, N. M., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education*, *60*(4), 380–392. https://doi.org/10.1177/0022487109339906

Fyfe, E. R., McNeil, N. M., & Borias, S. (2015). Benefits of "concreteness fading" for children's mathematics understanding. *Learning and Instruction*, *35*, 104–120. https://doi.org/10.1016/j.learninstruc.2014.10.004

Fyfe, E. R., & Nathan, M. J. (2019). Making "concreteness fading" more concrete as a theory of instruction for promoting transfer. *Educational Review*, *71*(4), 403–422. https://doi.org/10.1080/00131911.2018.1424116

Gall, M. D. (1970). The use of questions in teaching. *Review of Educational Research*, *40*(5), 707–721. https://doi.org/10.3102/00346543040005707

Gerofsky, S. (2009). Genre, simulacra, impossible exchange, and the real: How postmodern theory problematises word problems. In L. Verschaffel, B. Greer, & W. V. Dooren (Eds.), *Words and worlds: Modeling verbal descriptions of situations* (pp. 21–38). Sense Publishing.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences*, *14*(1), 69–110. https://doi.org/10.1207/s15327809jls1401_4

Hassler, R. (2016). *Mathematical comprehension facilitated by situation models: Learning opportunities for inverse relations in elementary school* [PhD dissertation]. Temple University, Philadelphia, PA.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., Olivier , A., & Human , P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Heinemann.

Hiebert, J., & Morris, A. K. (2012). Teaching, rather than teachers, as a path toward improving classroom instruction. *Journal of Teacher Education*, *63*(2), 92–102. https://doi.org/10.1177/0022487111428328

Hiebert, J., & Stigler, J. W. (2017). Teaching versus teachers as a lever for change: Comparing a Japanese and a U.S. Perspective on improving instruction. *Educational Researcher*, *46*(4), 169–176. https://doi.org/10.3102/0013189X17711899

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). LEARNING THEORY: The Advantage of abstract examples in learning Math. *Science*, *320*(5875), 454–455. https://doi.org/10.1126/science.1154659

Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5–17). Lawrence Erlbaum Associates.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *The Elementary School Journal*, *102*(1), 59–80. https://doi.org/10.1086/499693

Kieran, C. (2018). Seeking, using, and expressing structure in numbers and numerical operations: A fundamental path to developing early algebraic thinking. In C. Kieran (Ed.), *Teaching and learning algebraic thinking with 5- to 12-year-olds: The global evolution of an emerging field of research and practice* (pp. 79–105). Springer International Publishing AG.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologists*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*(6), 2797–2822. https://doi.org/10.2307/1131753

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher*, *44*(1), 54–61. https://doi.org/10.3102/0013189X15570388

Martino, A. M., & Maher, C. (1999). Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us? *Journal of Mathematical Behavior*, *18*(1), 53–78. https://doi.org/10.1016/S0732-3123(99)00017-6

Mason, J. (2018). How early is too early for thinking algebraically? In C. Kieran (Ed.), *Teaching and learning algebraic thinking with 5- to 12-year-olds: The global evolution of an emerging field of research and practice* (pp. 51–78). Springer International Publishing AG.

Mayer, R. E. (2010). *Applying the science of learning*. Pearson.

McNeil, N. M., & Fyfe, E. R. (2012). "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction*, *22*(6), 440–448. https://doi.org/10.1016/j.learninstruc.2012.05.001

Murata, A. (2008). Mathematics teaching and learning as a mediating process: The case of tape diagrams. *Mathematical Thinking and Learning*, *10*(4), 374–406. https://doi.org/10.1080/10986060802291642

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National mathematics Advisory Panel*. U.S. Department of Education.

Nesher, P., & Teubal, E. (1975). Verbal cues on an interfering factor in verbal problem solving. *Educational Studies in Mathematics*, *6*(1), 41–51. https://doi.org/10.1007/BF00590023

Ng, S. F., & Lee, K. (2009). The model method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education*, *40*, 282–313. https://www.jstor.org/stable/40539338

Nicol, C., & Crespo, S. (2006). Learning to teach with mathematics textbooks: How preservice teachers interpret and use curriculum materials. *Educational Studies in Mathematics*, *62*(3), 331–355. https://doi.org/10.1007/s10649-006-5423-y

Nunes, T., Bryant, P., & Watson, A. (2009). *Key understandings in mathematics learning: A report to the Nuffield Foundation*. Nuffield Foundation.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McGaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning (NCER 2007–2004)*. National Center for Education Research.

Pomerance, L., Greenberg, J., & Walsh, K. (2016). *Learning about learning: What every new teacher needs to know*. National Council on Teacher Quality.

Radford, L., & Roth, W. M. (2011). Intercorporeality and ethical commitment: An activity perspective on classroom interaction. *Education Studies in Mathematics*, *77*(2–3), 227–245. https://doi.org/10.1007/s10649-010-9282-1

Renkl, A., Atkinson, R. K., & Grobe, 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, L. B., Cauzinille-Marmeche, E., & Mathieu, J. (1987). Understanding algebra. In J. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 169–203). Clarendon.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). MATHEMATICS: Cognitive supports for Analogies in the mathematics classroom. *Science*, *316*(5828), 1128–1129. https://doi.org/10.1126/science.1142103

Russell, S., Schifter, D., & Bastable, V. (2011). *Connecting arithmetic to algebra: Strategies for building algebraic thinking in the elementary grades*. Heinemann.

Schliemann, A. D., Carraher, D. W., & Brizuela, B. M. (2007). *Bringing out the algebraic character of arithmetic: From children's ideas to classroom practice*. Lawrence Erlbaum.

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology*, *102*(4), 408–426. https://doi.org/10.1016/j.jecp.2008.11.004

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influence student learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 319–369). Information Age.

Stevens, R. (1912). *The question as a measure of efficiency in instruction: A critical study of classroom practice* (Contributions to Education No. 48). Teachers College, Columbia University.

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

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

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

Torbeyns, J., De Smedt, B., Ghesquière, P., & Verschaffel, L. (2009). Solving subtractions adaptively by means of indirect addition: Influence of task, subject, and instructional factors. *Mediterranean*

Journal for Research in Mathematics Education, *8*(2), 1–30. http://citeseerx.ist.psu.edu/viewdoc/download?doi = 10

van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology*, *36*(3), 212–218. https://doi.org/10.1016/j.cedpsych.2010.10.004

Vergnaud, G. (1988). Multiplicative structures. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 141–161). National Council of Teachers of Mathematics.

Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist*, *43*(1), 49–64. https://doi.org/10.1080/00461520701756420

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, *4*(3), 137–166. https://doi.org/10.1207/s1532690xci0403_1