Received: 22 April 2020 | Revised: 21 September 2020 | Accepted: 10 October 2020

DOI: 10.1111/ssm.12447

RESEARCH PAPER – MATHEMATICS EDUCATION

WILEY | SSMA

Comparing the frequency and variation of additive word problems in United States first-grade textbooks in the 1980s and the Common Core era

Robert C. Schoen | **Zachary Champagne** | **Ian Whitacre** | **Shelby McCrackin**

Florida State University, Tallahassee, FL, USA

Correspondence
Robert C. Schoen, Florida State University, Tallahassee, FL, USA.
Email: rschoen@lsi.fsu.edu

Funding information
Institute of Education Sciences, Grant/Award Number: R305A120781 and R305A180429

Abstract

Decades of research conducted worldwide has resulted in a taxonomy for word problems involving additive situations. Research published in the 1980s found that U.S. curricula only exposed first-grade students to a small subset of the easiest types of word problems. Widely adopted curriculum standards in the United States call for first-grade students to be exposed to the full set of 11 types of problems in the taxonomy, but research shows that substantive changes to the content of instructional materials can be elusive. We asked whether the U.S. curricula continue to expose students to the same small subset of problems. We coded four widely used first-grade textbooks from the Common Core era. Comparing our findings with those from the 1980s, we found many more word problems, a wider variety of word problems, and a consistent pattern in relative emphasis of various problems in the current textbooks. These results clearly show that U.S. students' exposure to word problems through curriculum materials has changed. This finding lays the groundwork for further research to determine whether such changes in exposure to different types of word problems have resulted in changes to relative problem difficulty.

KEYWORDS

curriculum, math/math education, standards, textbooks, word problems

1 | INTRODUCTION

Starting with a randomized controlled trial that was conducted in the 1980s, Cognitively Guided Instruction (CGI) has had a far-reaching impact on research, policy, and curriculum in mathematics education (Carpenter et al., 1989; Carpenter et al., 1996; Fennema et al., 1996; Jacobs et al., 2007; Schoen et al., 2018; Schoen et al., 2020). CGI is complex and multifaceted, but research-based frameworks for types of word problems and strategies students use to solve problems are two of its more salient features (Carpenter et al., 1999). The CGI frameworks for additive word problems (i.e., word problems that can be solved by adding or subtracting the two given quantities) are based on several decades of research in the learning sciences that crystalized during the 1980s into frameworks for types of word problems and associated strategies students use to solve those problems.

Compiled in the seminal book on CGI (Carpenter et al., 1999), the CGI framework for additive word problems provides a taxonomy of 11 types of problems. Stigler et al. (1986) found that, essentially, only three of those 11 types of problems were present in first-grade textbooks used in the United States during the mid-1980s, when CGI was first created. Of the 11 types of problems, the authors noted that the three predominant types of problems in first-grade textbooks were also the three easiest types, on average, for students to solve correctly.

During the 1980s and before, states, national-level organizations, and school districts created guidelines for mathematics curriculum, but textbook authors and publishers had a large influence on the content of school mathematics textbooks, including setting grade-level expectations for students. In 1989, the National Council of Teachers of Mathematics published *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), which ushered in the present-day era of standards-based curriculum that has been adopted across the United States. Over the next two decades, every state wrote and adopted curriculum standards for school mathematics, because the No Child Left Behind Act required it. Until 2010, each individual state wrote, adopted, and implemented its own set of curriculum standards, and variation in the substantive content of state standards was wide (Reys et al., 2010).

In 2010, the National Governors Association (NGA) and the Council of Chief State School Officers (CCSSO) published the Common Core State Standards for Mathematics (CCSSM; NGA & CCSSO, 2010). Over the next few years, 45 of the 50 states adopted and implemented those standards. The CCSSM have, without a doubt, altered the landscape of U.S. mathematics education. As we write, approximately a decade has passed since the finalization of the CCSSM. Textbook publishers have gone to considerable effort and expense to align their textbooks with their structure and content, and school districts have been using these books for several years. Publishing companies in the United States compete for contracts to provide curriculum materials to school districts, and a variety of elementary mathematics textbooks are available under the auspices of alignment with the CCSSM, which have provided the backbone of the accountability systems for most of the public schools in the United States during the 2010s. As a result, we consider the present period in U.S. education—at least in the subject areas of mathematics and language arts—to be the era of the Common Core.

First-grade standards in the CCSSM contain a taxonomy of additive (i.e., addition, subtraction) word problems that includes all 11 types of problems in the CGI framework/taxonomy (Carpenter et al., 1999; NGA & CCSSO, 2010). Comparison of the CCSSM with the findings of Stigler et al. (1986) seems to imply that many U.S. students are now expected to solve a wider variety of problem types than they were when Stigler et al. performed their analysis of first-grade textbooks. Although curriculum standards influence the textbooks and other instructional resources, and scholars agree that textbooks continue to play an important role in determining the enacted curriculum, the degree of alignment between the official curriculum and the instructional materials and other components of the enacted curriculum can vary (Hong et al., 2019; Polikoff, 2015; Remillard et al., 2014; Remillard & Heck, 2014; Tran, 2016). Comparing contemporary science textbooks with those of the 1980s, for example, Groves (2016) found that little had changed. The older science textbooks placed too much emphasis on vocabulary. The same was true of the contemporary texts, despite a body of research literature in the intervening period that had identified and raised warnings with regard to this issue. These recent shifts in curriculum standards in the United States prompted us to ask whether and how U.S. mathematics textbooks have changed with regard to the number and type of word problems that are present in first-grade textbooks.

The purpose of the study reported here was to determine whether and how the frequency and variation of word problems in first-grade U.S. mathematics textbooks differ from those in U.S. textbooks of the mid-1980s. We examined four textbooks that were widely used in the United States in the 2010s, focusing on two factors: (a) the frequency with which students may encounter word problems in their textbooks and (b) how the types of word problems presented in textbooks differ with respect to semantic structure and the position of the unknown value (Carpenter et al., 1999; Fuson, 1992; NGA & CCSSO, 2010; Verschaffel & De Corte, 1993). Our study was guided by the following research question: How do the frequency of word problems and variability of types of word problems in Common Core–era first-grade mathematics textbooks in the United States compare with those of the 1980s?

2 | LITERATURE REVIEW

2.1 | The rise of problem-type taxonomies

Many facets of word problems have been studied and categorized by scholars. Features of interest have included vocabulary and grammar, action versus nonaction, abstract versus concrete, factual versus hypothetical, consistent versus inconsistent language, additive versus multiplicative, number of steps to solve the problem, location of the unknown value, mode of presentation, number of words, parts of speech, verbal cues or keywords, classes of numbers involved, and presence of distractors (see, e.g., Bergeron & Herscovics, 1990; Caldwell & Goldin, 1979; Carpenter et al., 1999; De Corte & Verschaffel, 1987; Gibb, 1956; Hiebert, 1982; Jerman & Mirman, 1974; Lewis & Mayer, 1987; Stigler et al., 1986).

In the latter part of the 20th century, researchers began to focus on the influence on students' cognitive processes and problem-solving abilities of semantic structure and position of the unknown value in word problems involving additive situations. Research on this topic clearly shows that various combinations of semantic structure (e.g., Change, Combine, and Compare) and position of the unknown value affect cognitive strategies students use to solve the problems and the relative difficulty of types of word problems. This corpus of research has resulted in several different taxonomies of additive-type word problems that are conceptually similar but differ in their terminology and level of specificity (Carpenter et al., 1999; Carpenter & Moser, 1983; Fuson, 1992; Gibb, 1954; Nesher et al., 1982; Verschaffel et al., 2007).

Table 1 presents a widely acknowledged scheme organizing 11 types of one-step, additive word problems (Carpenter et al., 1999). It consists of two categories of problems involving sets of objects or values of quantities that change over the course of the story in the problem—Join and Separate—and two categories of problems wherein the quantities are static—Part-Part-Whole and Compare—together with the different possible cases for unknown values in each. This particular scheme emerged in the 1980s and 1990s. The taxonomy of problem types presented in Table 1 has influenced teacher professional-development programs in mathematics (see, e.g., Carpenter et al., 1996, 1999; Fennema et al., 1996; Schoen et al., 2018, 2020), further research on word-problem difficulty (see, e.g., Garcia et al., 2006; Olkun & Toluk, 2002; Stigler et al., 1986), and policy (see, e.g., NGA & CCSSO, 2010). Although it uses different terminology for the Join, Separate, and Part-Part-Whole semantic categories, the taxonomy of additive word problems in the CCSSM (see Table 2) is very similar to that in Table 1.

2.2 | Types of problems in mathematics curricula

In their analysis of four widely used U.S. textbook series for the early elementary grades in the mid-1980s, Stigler et al. (1986) found that U.S. curricula essentially only included three or four types of additive word problems and reported a distinct pattern in the types that were included; they observed that the problem types included tended to be the easiest ones for U.S. students to solve. Stigler and colleagues also reported that the U.S. textbooks of the early 1980s contrasted with the Soviet curriculum at the time. The latter included more variation in the types of word problems appearing in textbooks and placed greater emphasis on the more challenging types in the problem-type taxonomy. Stigler et al. (1986) recommended that the design of U.S. curricula be informed by research on how children learn mathematics in general and by problem-type taxonomies in particular.

2.3 | Standards and textbook alignment

Recent studies have investigated relationships between standards and curricula, including making comparisons between current curriculum materials and those from decades ago. Cady et al. (2015) compared traditional with standards-based curricula in light of relevant research literature, standards, and recommendations. They found significant differences related to features such as frequency of presentation

TABLE 1 Additive word-problem types

Join	*Result Unknown* Sofia saw a group of eight manatees in the river. Five more manatees joined the group. How many manatees are in the group now?	*Change Unknown* Sofia saw a group of eight manatees in the river. Some more manatees joined the group. Now there are 13 manatees in the group. How many manatees joined the group?	*Start Unknown* Sofia saw a group of manatees in the river. Five more manatees joined them. Now there are 13 manatees in the river. How many manatees were in the group to start with?
Separate	*Result Unknown* Sofia had 13 oranges. She ate five oranges. How many oranges does she have left?	*Change Unknown* Sofia had 13 oranges. She ate some oranges. Now she has eight oranges. How many oranges did she eat?	*Start Unknown* Sofia had some oranges. She ate five of the oranges. Now she has eight oranges. How many oranges did she have to start with?
Part-Part-Whole	*Whole Unknown* Sofia sees eight alligators on a log and five alligators in the river. How many alligators does Sofia see?	*Part Unknown* Sofia sees 13 alligators. Eight of the alligators are on a log, and the rest are in the river. How many alligators does Sofia see in the river?	
Compare	*Difference Unknown* Sofia has 13 oysters. Luka has eight oysters. How many more oysters does Sofia have than Luka?	*Compare Quantity Unknown* Luka has eight oysters. Sofia has five more than Luka. How many oysters does Sofia have?	*Referent Unknown* Sofia has 13 oysters. She has five more than Luka. How many oysters does Luka have?

Note: Problem types chart adapted from Carpenter et al. (1999).

TABLE 2 A Comparison of the word-problem types used by the study reported here, the Common Core State Standards for Mathematics (CCSSM; NGA & CCSSO, 2010), and Carpenter et al. (1999)

Type	NGA & CCSSO (2010)	Carpenter et al. (2015)
Change		
1	Add to result unknown	Join result unknown
2	Add to change unknown	Join change unknown
3	Add to start unknown	Join start unknown
4	Take from result unknown	Separate result unknown
5	Take from change unknown	Separate change unknown
6	Take from start unknown	Separate start unknown
Combine		
7	Put together/take apart total unknown	Part-part-whole whole unknown
8	Put together/take apart addend unknown	Part-part-whole part unknown
Compare		
9	Compare difference unknown	Compare difference unknown
10	Compare bigger unknown	—[a]
11	Compare smaller unknown	—[a]

[a]The word-problem-type taxonomy used by Carpenter et al. (1999) includes compare-type problems with the compare quantity unknown or the referent unknown. These two variations are not listed here, because they are not identical to the categories presented by the CCSSM (NGA & CCSSO, 2010) or to the word-problem types used in our investigation.

of representations and use of real-world contexts—thus, different degrees of (mis)alignment with research-based recommendations. Hong et al. (2019) examined math textbooks identified as being aligned with the CCSSM in comparison to the research literature for a particular topic, area measurement. They reported substantial misalignments, especially in the early grades. Similarly, Tran (2016) investigated alignment between high-school math textbooks and the CCSSM with regard to the presentation of the topic of statistical association. Tran found that some aspects of this topic (especially relationships between two numerical variables) were given substantial attention in the three textbooks analyzed, whereas other aspects (relationships between numerical and categorical or categorical and categorical variables) were given little attention. A common theme in the findings of the above studies is misalignment between curriculum materials and standards or research-based recommendations. In addition, Tran raises important question about what alignment means and what it should look like:

> The results of the current study call for further investigation of alignment studies. The challenge is that curriculum standards do suggest important topics to include in teaching and learning, but do not specify the level of treatment for the topics. Should a balance of task types be present for alignment or are particular problem types more challenging and thus more worthy of textbook terrain? (pp. 294–295)

Tran's question is particularly relevant to our study of word-problem types: If a wider variety of word problem types is represented in Common Core–era textbooks, how frequently do the less traditional problem types appear? [Correction added on February 13, 2021, after initial publication; the displayed quote in the section entitled “**Standards and textbook alignment**” was initially omitted due to a production error and has been reinstated.]

3 | CONCEPTUAL FRAMEWORK

3.1 | Curriculum

Within the domain of curriculum, Remillard and Heck (2014) draw a distinction among official curriculum, instructional materials, and operational curriculum. In practice, none of these three components of curriculum are isomorphic with either of the other two parts. Rather, each involves an interpretation of the other. Textbooks are one aspect of the instructional materials, whereas the parts of the textbook that are implemented or used by teachers and students contribute to the constellation of parts that comprise the enacted curriculum.

The CCSSM provided the keystone for the official curriculum in most of the United States during the 2010s. We hypothesize that the CCSSM, representing an important facet of the official curriculum, probably had a substantive impact on the instructional materials that were available to schools and teachers. Our study focused on the instructional-materials component of curriculum—specifically textbooks.

3.2 | Inherence or exposure

In the 1980s, problem-type taxonomies were created as a way to make sense of empirical findings generated by scholars in a variety of traditions, including the learning sciences, cognitive and developmental psychology, and mathematics education (Bergeron & Herscovics, 1990; Carpenter & Moser, 1983; Hiebert, 1982; Riley & Greeno, 1988; Riley et al., 1983). The taxonomy crystalized around an emerging theory that the difficulty of various types of problems could be explained by the combination of semantic structure and position of the unknown value in the word problem—we call this the inherence hypothesis: certain word-problem types are inherently more difficult (for students in the primary grades) to solve than other types of problems are.

Stigler et al. (1986) found that the three easiest types of problem comprised about 95% of the word problems found in first-grade textbooks in the mid-1980s, suggesting a correlation between the extent to which students were exposed to various types of word problems and the relative difficulty of the types of problems. More recent research—conducted in settings with a cultural tapestry different from those where the problem-type taxonomies were initially created—lends credibility to the claim that increased exposure to specific types of word problems is what makes those problems easier for students to solve, thereby challenging the inherence hypothesis and raising the possibility that exposure to certain problem types is the important factor that determines relative problem difficulty (Olkun & Toluk, 2002; Xin, 2007). We call this latter conjecture the exposure hypothesis.

3.3 | Roles of different types of additive word problem in aiding learning

The relation between addition and subtraction is a fundamentally important concept in mathematics, and the CCSSM draw explicit attention to the importance of students' understanding of this relation. A wider variety of problems may expose students to a fuller range of ways that the addition and subtraction operations can be applied to solve problems, and therefore, provide an expanded perspective on the conceptual interpretations of the two operations. Several of the types of problems in the 11-problem taxonomy provide students with opportunities to consider this relation (e.g., Join Change Unknown, Compare Difference Unknown, Part-Part-Whole Part Unknown), while some of the other types of word problems (e.g., Join Result Unknown, Separate Change Unknown, Part-Part-Whole Whole Unknown) do not lend themselves to such opportunities. The former types of problems may create more opportunity for learners to consider the relation between the addition and subtraction operations and how this relation can be used to solve arithmetic and algebraic mathematics problems.

4 | METHOD

4.1 | Selection of U.S. first-grade textbooks for analysis

We determined the frequency and variation of word problems in the first-grade student editions of *Investigations in Number, Data, and Space* (Akers et al., 2017), *Math Expressions* (Fuson, 2013), *enVisionmath 2.0* (Charles et al., 2016), and *Saxon Math* (Larson & Matthews, 2012) curriculum materials. We selected these four textbooks, because they were each widely used in U.S. classrooms in 2016, they differed in their approaches to teaching mathematics, and recent editions of each of these books were included in a randomized study involving a four-way comparison of the effects of first- and second-grade mathematics curricula on student achievement (Agodini & Harris, 2010).

We used the main student book for each publication as the basis of our investigation. We restricted our coding to the word problems presented in the student edition of the respective books—a decision made in similar analyses performed by Powell (2012) and McNeil et al. (2006) in their investigations of various presentations of the equals sign in mathematics textbooks. We included every word problem we encountered in each student edition, including those encountered in sections with titles such as homework, readiness, reteaching, assessments, grade-two ramp up, practice problems, review problems, and additional practice problems. The teacher edition and other ancillary materials would probably increase the overall quantity of word problems, so our counts probably underestimate the total number of problems that students potentially encountered, but we chose this narrow scope to permit a common baseline for comparison across the four current books and for consistency with the methods of Stigler et al. (1986).

4.2 | Working definition of word problem

When identifying word problems, we used the definition used by Stigler et al. (1986): "we defined a word problem as consisting of two premises (the given information) and a question. To be coded, a problem had to present two or more premises and a question" (p. 160). We used this definition as a primary guideline in aligning our coding system with the one used by Stigler et al. in their original paper to permit fair comparison. Here, we present a more complete list of inclusion criteria, and we believe they are all consistent with the definition and coding decisions made by Stigler et al. (1986):

- Problems must fit into one of the 11 categories presented in Table 2.
- Problems must include a question.
- Problems may include both words and pictures (see Stigler et al., 1986, p. 160).
- Problems include both constructed-response and selected-response formats.
- Problems may include extraneous information, so long as they fit the fundamental criteria of containing two or more premises and a question.
- Problems may require students to gather information from a table, graph, or chart, but only when those problems meet the other parts of the definition of word problem.
- Problems may require students to analyze the reasoning of other students.

- Problems may call for students to use specific strategies to solve them.

Stigler et al. (1986) coded multistep word problems with a sequential coding of each step of the word problem. In the U.S. first-grade textbooks coded in their study, only two of the textbooks included multistep problems, and those two only included three instances each. Because of the infrequency of these items and the absence of explicit reference to multistep word problems in the CCSSM for first grade, we decided not to code multistep word problems. The first-grade CCSSM do include addition problems with three addends. We extended our decision about multistep problems also to exclude instances of word problems with three addends.

4.3 | Categories of additive word problems in our analysis

Categorization schemes for one-step, additive word problems have been evolving over several decades, and many variations exist. One trend over the past three decades has been to reduce the number of distinct types from 20 or more in the taxonomies used in the mid-1980s (Carpenter & Moser, 1983; Fuson, 1992; Stigler et al., 1986) to somewhere between 11 and 14 in current use (Carpenter et al., 1999; NGA & CCSSO, 2010; Riley et al., 1983; Verschaffel & De Corte, 1993).

Table 2 contains a list of the 11 categories of problems types used in our study. Their numbers correspond directly those used by Stigler et al. (1986) with the exception that problem type 15 in their study is subsumed by problem type 2 in the present study. The categories we used are consistent with the 11 categories in the taxonomy of one-step, additive word problems in the CCSSM (NGA & CCSSO, 2010). The CCSSM scheme lists 12 problem types as a result of the addition of Both Addends Unknown problems to the scheme (Champagne et al., 2014). Because those problems have a different mathematical structure and were not included in the Stigler et al. (1986) study, we did not include them in our coding. Table 2 lists the 11 numbered problem types and their corresponding names as per the CCSSM (NGA & CCSSO, 2010) taxonomy and the Carpenter et al. (1999) taxonomy.

4.4 | Coding procedure

All four of the coauthors of the present article participated in the coding of the Common Core–era textbooks. Each person coded two books. Two people were assigned to each book, and the pairs were rotated so that the dyads did not consist of the same two people for more than one book.

First, the coder looked at each task in a given book and decided whether it met our working definition of word problem. After determining that a problem met the criteria, the coder determined which one of the 11 types of additive-type word problems in our coding scheme it represented. Working separately, the two coders completed this process for a whole textbook. After each had completed coding, the coding decisions were checked for agreement, and any discrepancies were resolved in a meeting of the two coders. The percentage agreement among pairs of reviewers was calculated with a formula offered by Miles et al. (2013), which corresponds to the number of agreements divided by the sum of the number of agreements and disagreements. The percentage agreement was calculated separately for identification of word problems and for assignment to problem type. The within-book agreement for identification ranged from 0.93 to 1.00. That for assignment ranged from 0.88 to 0.99.

4.5 | Secondary analysis of word problems in 1980s textbook data

Stigler et al. (1986) analyzed the frequency and variation of word problems in four U.S. textbook series. They split their results for frequency and variation of word problems by grade level, thereby allowing for secondary analysis at the individual grade level. Because our study focused on first grade, we used the data from the Stigler et al. article (Table 3, pp. 162–163 in the original source) to generate frequency data for each individual book in the form of the number of each of the 11 types of additive word problems, the total number of word problems in each of the four first-grade books from the 1980s, and the percentages of word problems for each problem type within each book. In addition, we computed the median count (and percentage) across the four books. We calculated the median as a measure of center, because the Common Core–era books included a large outlier, and the outlier would affect the mean more than the median, possibly resulting in an upward-biased estimate for the number of word problems in Common Core–era books. As explained previously, we reinterpreted equalize-type problems as Join Change Unknown problems to be consistent with the current perspective on those problems.

5 | FINDINGS

Table 3 shows the result of our secondary analysis of the 1980s books and our primary analysis of the Common Core–era books.

TABLE 3 Number (and percentage) of problems by type and textbook

1980s books					
Type	**Addison-Wesley (1983)[a]**	**Harper & Row (1985)[a]**	**Houghton Mifflin (1985)[a]**	**Scott, Foresman (1985)[a]**	**Median**
1	8 (0.13)	11 (0.12)	13 (0.16)	13 (0.15)	12.0 (0.14)
2	0 (0.00)	0 (0.00)	0 (0.00)	4 (0.05)	0.0 (0.00)
3	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.0 (0.00)
4	24 (0.38)	43 (0.45)	32 (0.39)	47 (0.53)	37.5 (0.42)
5	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.0 (0.00)
6	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.0 (0.00)
7	32 (0.50)	30 (0.32)	35 (0.42)	20 (0.23)	31.0 (0.37)
8	0 (0.00)	8 (0.08)	0 (0.00)	4 (0.05)	2.0 (0.02)
9	0 (0.00)	3 (0.03)	2 (0.02)	0 (0.00)	1.0 (0.01)
10	0 (0.00)	0 (0.00)	1 (0.01)	0 (0.00)	0.0 (0.00)
11	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0.0 (0.00)
Total	64	95	83	88	85.5
Common Core–era books					
Type	**enVision Math**	**Investigations**	**Math Expressions**	**Saxon Math**	**Median**
1	80 (0.18)	15 (0.16)	7 (0.06)	53 (0.28)	34.0 (0.17)
2	46 (0.10)	9 (0.10)	10 (0.09)	0 (0.00)	9.5 (0.09)
3	7 (0.02)	0 (0.00)	4 (0.03)	0 (0.00)	2.0 (0.01)
4	77 (0.17)	19 (0.20)	21 (0.18)	70 (0.36)	45.5 (0.19)
5	12 (0.03)	8 (0.09)	9 (0.08)	0 (0.00)	8.5 (0.05)
6	3 (0.01)	0 (0.00)	5 (0.04)	0 (0.00)	1.5 (0.00)
7	75 (0.17)	10 (0.11)	15 (0.13)	44 (0.23)	29.5 (0.15)
8	62 (0.14)	12 (0.13)	18 (0.16)	6 (0.03)	15.0 (0.13)
9	42 (0.09)	13 (0.14)	13 (0.11)	19 (0.10)	16.0 (0.11)
10	20 (0.05)	4 (0.04)	7 (0.06)	0 (0.00)	5.5 (0.04)
11	19 (0.04)	4 (0.04)	6 (0.05)	0 (0.00)	5.0 (0.04)
Total	443	94	115	192	153.5

[a]As cited by Stigler et al. (1986).

5.1 | Frequency of word problems

We find clear evidence that the Common Core–era books include many more word problems than did those from the 1980s. The median number of word problems in the 1980s books was 85.5; that in the Common Core–era books is 153.5, almost twice that observed in the 1980s books.

The variability in the total number of additive word problems is an order of magnitude greater in the Common Core–era books than in the 1980s books. The total number of word problems in the 1980s books ranged from 64 to 95, a range of only 31. That in the Common Core–era books ranged from 94 to 443, a range of 349. Even if we removed the outlier (*enVision Math*), the range covered by the other three books is 98, still much greater than the total number of word problems seen in any individual textbook in the 1980s. Moreover, the Common Core–era book containing the fewest word problems still contained only one fewer than did the 1980s book that contained the most.

5.2 | Distribution of types of word problems

The variation in the types of word problems in the Common Core–era books was also much greater than that in the 1980s books. The median counts and percentages of the various types of word problems clearly revealed that three types of problems predominated in the 1980s books. These correspond to the two change-type problems with the result unknown and the combine-type problem with the whole unknown. These three problem types comprised an average of 93% of all the additive word problems found in the first-grade books in the 1980s. Those three types comprise an average of only 51% of additive word problems in the Common Core–era textbooks. In sharp contrast with the distribution of types of problems in the 1980s, two of the four Common Core–era textbooks contain at least one example of every word problem type.

Although the relative emphasis on these three problem types has decreased, the average number of problems of those three types is 109 per book in the Common Core–era books. Comparing that number with the total number of word problems in the 1980s reveals that the Common Core–era books have more word problems of those three types than the 1980s books had total word problems.

5.3 | Patterns in the types of problems within and across textbooks

Figure 1 shows the percentage of each type of additive word problem observed in the four first-grade textbooks from the 1980s (Stigler et al., 1986). We converted the numbering system used by Stigler et al. to the numbering system we are using to permit a side-by-side comparison, but we retained the style of graph (i.e., points with line segments connecting the adjacent points) to be consistent with the data-visualization style in the original publication. A clear pattern emerges in the first-grade textbook data, and the pattern is somewhat different from the overall patterns reported across the primary-grades data in the original research. Although four types of problems dominated across the three grade levels analyzed by Stigler et al., only three types of word problems dominated the first-grade textbooks from the 1980s. The four major U.S. first-grade textbooks in use at that time followed a highly consistent pattern; all of them emphasized these three types of word problems.

The graph in Figure 2 shows the percentage of each type of additive word problem that we found in our analysis of the four Common Core–era, first-grade textbooks. The graph clearly shows that three of the Common Core–era books (*Math Expressions*, *Investigations*, *enVision Math*) show much more overall variation in problem types than did the 1980s books. At the same time, these same three textbooks appear to show a high level of internal consistency with respect to which problems are represented more frequently or less frequently (with the exception of the notably lower emphasis on Type 1 problems in *Math Expressions*). Also clear is that the Common Core–era edition of *Saxon Math* does not follow the same pattern as the other three Common Core–era textbooks but instead shows patterns nearly identical to those in the 1980s books, with the exception that compare-type problems with the difference unknown are more heavily emphasized in the current first-grade *Saxon Math* textbook than they were in the 1980s books.

6 | DISCUSSION

For decades, mathematics education researchers have been interested in ways of categorizing additive word problems, children's thinking about different types of word problems, and the distributions of various types of word problems in textbooks. A particular categorization scheme has become prominent and now appears in the CCSSM, whereas U.S. textbooks only focused on a small number of problem types (those with the unknown value in the traditional location) in the 1980s, we find that U.S. textbooks today present a much more balanced distribution of additive word problem types.

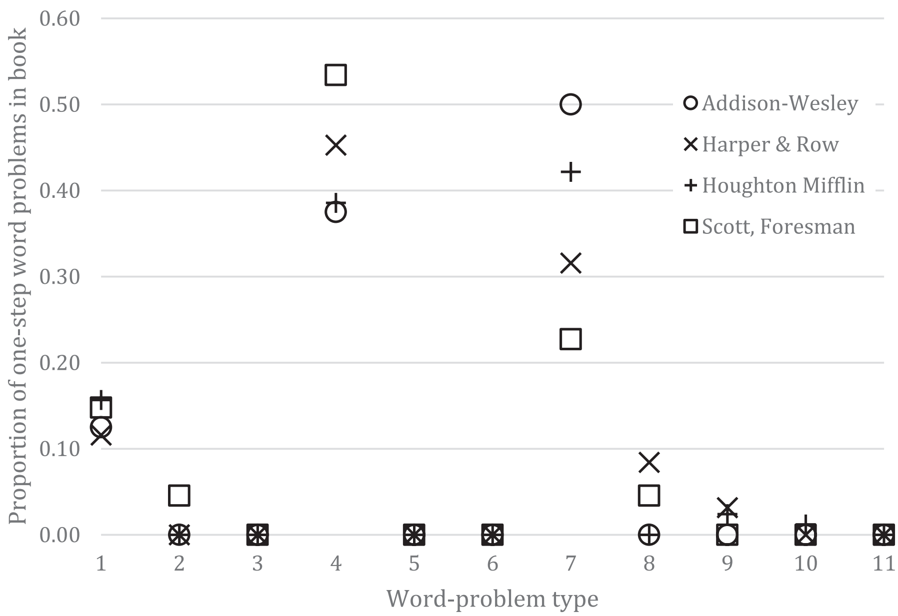

FIGURE 1 Distribution of one-step word problems in each of the four 1980s first-grade student books (secondary analysis, based on findings reported by Stigler et al., 1986)

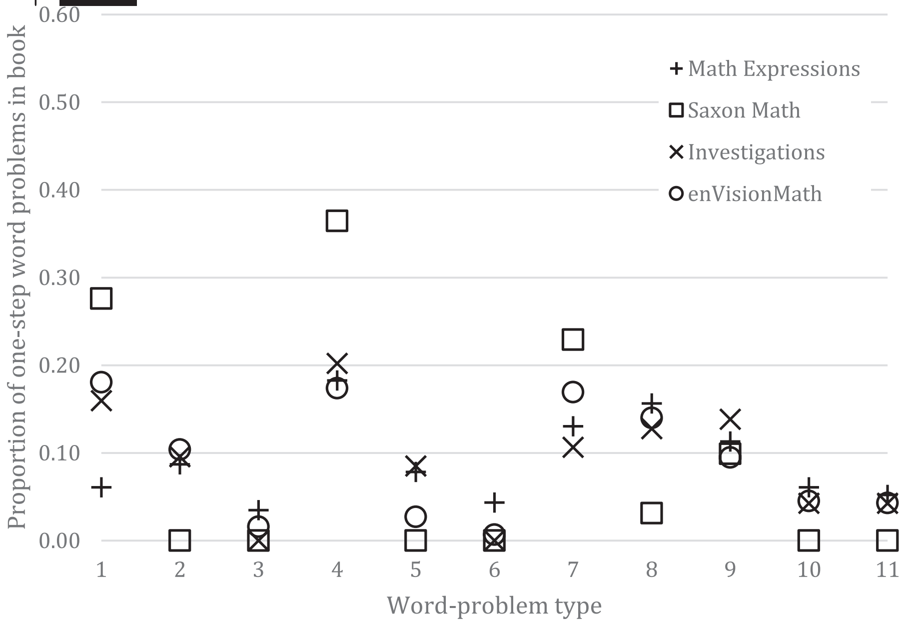

FIGURE 2 Distribution of one-step word problems in each of the four Common Core–era, first-grade student books

The empirical research that resulted in the creation of the taxonomy of problem types investigated how first graders responded to additive word problems of various types and how successful they were in solving different types of problems. Those results must be interpreted with respect to the context in which the studies were conducted. Given what is known about the textbooks of the 1980s, the students in such studies may never have been asked nontraditional problem types, such as Join Start Unknown, before their participation in the research study. The observation of Stigler et al. (1986) that the most common word problem types in textbooks were also those that were easiest for students to solve was, therefore, confounded by the issue of opportunity. The degree to which historical findings regarding children's responses to additive word problems were influenced by children's familiarity with the traditional problem types and lack of familiarity with the nontraditional problem types remains unknown, but it is a fundamentally important question, because the basic research on word problems that occurred in the latter half of the twentieth century has had a major impact on present-day mathematics curriculum and research.

Our results provide evidence that mainstream, Common Core–era, first-grade textbooks are considerably different from their predecessors in at least one respect. The Common Core–era, first-grade textbooks contain a greater number of word problems and greater variation in word-problem types than the mainstream first-grade textbooks published 30 years before them. Interestingly, Stigler et al. (1986, p. 160–161) considered the number of word problems to "vary considerably from series to series" in their analysis of the 1980s books. We conclude that the setting in which present-day curriculum and research is being conducted has changed from that of the 1980s. Although this change has been well established with respect to the policy environment, our findings illustrate that the differences also exist in the substance of the mathematics textbooks.

Stigler et al. (1986) noticed a distinct pattern in the types of problems included (and not included) in the four U.S. textbooks. We also found a remarkably consistent pattern among three of the Common-Core–era textbooks, albeit a pattern different from that observed in the 1980s books. *Saxon Math* is the exception. The distribution of types of problems in the *Saxon Math* book appears to be much more consistent with the pattern observed in the 1980s books, except that the *Saxon Math* book includes more compare-type problems.

6.1 | Limitations

One limitation of our study results from its focus on the student editions of each textbook. The role of the students' book differs in different curricula, and we cannot determine whether the distributions of word problems might have been different if we had included the teacher edition and/or other ancillary materials. We also cannot infer the extent to which these textbooks are implemented in full. We can only examine them at face value, recognizing that the intended and the enacted curricula will differ.

On the basis of the research design we used, we cannot claim that the CCSSM caused the observed differences in textbooks. Another plausible explanation is that textbook authors are generally more knowledgeable about the problem-type taxonomy than they were 30 years ago, and they may have incorporated this knowledge into the design of textbooks throughout that time. In other words, the dissemination of information from research done in the latter half of the 20th century could have resulted in a certain type of innovation creep, wherein various types of word problems may have infiltrated first-grade textbooks over the past 30 years without the influence of the CCSSM.

6.2 | Implications

In their 1986 publication, Stigler et al. acknowledged that the ways in which the distribution of types of problems affects learning was not known. In the past three decades, some cross-cultural comparisons of student achievement and curricula have address this question—with a specific focus on exposure to various types of word problems and relative difficulty of those various problem types. We are certain that the writers of the CCSSM were aware of the literature related to word problem taxonomies and related ideas such as the relative difficulty of the various problem types. We conjecture that the CCSSM writers crafted this policy from a working hypothesis that exposure to a wider variety of word problem types should increase student learning and mathematical ability. Whether this change in exposure affects students' problem-solving ability overall or the relative difficulty of the various problems types remains a very important empirical question. Our findings lay the foundation for such investigations. We hope that our report amplifies the decades-old call by Stigler et al. to investigate the effect of exposure to various problem types on student learning.

At least in terms of additive word problems, our findings provide some evidence to support the claim that the mathematics curriculum in the Common Core–era is different from that found in earlier eras in U.S. mathematics education. With the exception of *Saxon Math*, we find that current first-grade textbooks contrast sharply with those of the 1980s in terms of distribution of additive word-problem types. As a result, research on children's thinking and success rates when they solve various types of additive word problems must be revisited. The degree to which research findings from previous eras remain relevant in the Common Core–era remains unclear. Our findings, therefore, signal exciting new opportunities for research. Given that mathematics textbooks have changed significantly, researchers are now positioned to investigate how those changes might have affected students' mathematical thinking and learning. By focusing on important microcosms of mathematics curricula such as additive word problems, the field can investigate specific hypotheses regarding the possible effects of such changes.

Blazar et al. (2019) reported that the recent changes in curriculum standards and the resulting changes in textbooks have not resulted in overall increases in student learning. Our study took a more focused look at a cornerstone topic in the elementary curriculum (addition/subtraction) and a central component in the teaching of that topic (word problems) to determine whether the substantive changes called for by the CCSSM did, indeed, occur. We found that they did. An important next step in this important line of inquiry will be to find out whether the verified changes in curriculum (both standards and the textbooks) have resulted in changes in student performance.

7 | CONCLUSION

Amid an ongoing inherence–exposure debate, the results of the present study indicate that students today are exposed to a wider variety of additive word problems. This phenomenon may have resulted because textbook authors and publishers incorporated research findings into their books voluntarily, or it may have resulted from changes in policy. Although our study cannot speak to the mechanism or process that caused the change, authors and publishers have clearly incorporated a wider variety of the various types of additive word problems into first-grade textbooks in the U.S. than were included in those used during the mid-1980s. The results, therefore, lay the groundwork for a much-needed, focused investigation to determine whether these changes in curricula—which result in changes in the exposure or U.S. students to different types of word problems—have resulted in changes in student abilities.

ACKNOWLEDGMENTS

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through grant numbers R305A120781 and R305A180429 to Florida State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

REFERENCES

Agodini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness*, *3*(3), 199–253. https://doi.org/10.1080/19345741003770693

Akers, J., Bastable, V., Battista, M. T., Hickey, K., Clements, D., Cochran, K., Earnest, D., Economopoulos, K., Hollister, A., Horowitz, N., Kliman, M., Leidl, E., Mokros, J., Murray, M., Nemirovsky, R., Oh, Y., Perry, B. W., Rubin, A., Russell, S. J., Schifter, D., Sillman, K., &Tierney, C. (2017). *Investigations in number, data, and space 3*. Pearson.

Bergeron, J. C., & Herscovics, N. (1990). Psychological aspects of learning early arithmetic. In P. Nesher, & J. Kilpatrick (Eds.), *Mathematics and cognition*: *A research synthesis by the international group for the psychology of mathematics education* (pp. 31–52). Cambridge University Press.

Blazar, D., Heller, B., Kane, T. J., Polikoff, M., Staiger, D., Carrell, S., Goldhaber, D., Harris, D., Hitch, R., Holden, K. L., & Kurlaender, M. (2019). *Learning by the book: Comparing math achievement growth by textbook in six Common Core states* (Research Report). Center for Education Policy Research, Harvard University.

Cady, J. A., Hodges, T. E., & Collins, R. L. (2015). A comparison of textbooks' presentation of fractions. *School Science and Mathematics*, *115*(3), 105–116. https://doi.org/10.1111/ssm.12108

Caldwell, J. H., & Goldin, G. A. (1979). Variables affecting word problem difficulty in elementary school mathematics. *Journal for Research in Mathematics Education*, *10*(5), 323–336. https://doi.org/10.5951/jresematheduc.10.5.0323

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively guided instruction: A knowledge base for reform in primary mathematics instruction. *The Elementary School Journal*, *97*(1), 3–20. https://doi.org/10.1086/461846

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Heinemann.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, *26*(4), 385–531. https://doi.org/10.3102/00028312026004499

Carpenter, T. P., & Moser, J. M. (1983). The acquisition of addition and subtraction concepts. In R. Lesh, & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 7–44). Academic Press.

Champagne, Z., Schoen, R. C., & Riddell, C. (2014). Variations in Both Addends Unknown problems. *Teaching Children Mathematics*, *21*(2), 114–121. https://doi.org/10.5951/teacchilmath.21.2.0114

Charles, R. I., Bay-Williams, J., Berry, R. Q., Caldwell, J. H., Champagne, Z. C., Copley, J., Crown, W., Fennell, F., Karp, K., Murphy, S. J., Schielack, J. F., & Suh, J. M. (2016). *enVisionmath 2.0*. Pearson.

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. *Journal for Research in Mathematics Education*, *18*(5), 363–381. https://doi.org/10.5951/jresematheduc.18.5.0363

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education*, *27*(4), 403–434. https://doi.org/10.2307/749875

Fuson, K. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. National Council of Teachers of Mathematics.

Fuson, K. (2013). *Math expressions common core*. Houghton Mifflin Harcourt.

Garcia, A. I., Jimenez, J. E., & Hess, S. (2006). Solving arithmetic word problems: An analysis of classification as a function of difficulty in children with and without arithmetic LD. *Journal of Learning Disabilities*, *39*(3), 270–281. https://doi.org/10.1177/00222194060390030601

Gibb, E. G. (1954). Take-away is not enough!. *The Arithmetic Teacher*, *1*(2), 7–10. https://doi.org/10.5951/AT.1.2.0007

Gibb, E. G. (1956). Children's thinking in the process of subtraction. *The Journal of Experimental Education*, *25*(1), 71–80. https://doi.org/10.1080/00220973.1956.11010564

Groves, F. H. (2016). A longitudinal study of middle and secondary level science textbook vocabulary loads. *School Science and Mathematics*, *116*(6), 320–325. https://doi.org/10.1111/ssm.12183

Hiebert, J. (1982). The position of the unknown set and children's solutions of verbal arithmetic problems. *Journal for Research in Mathematics Education*, *13*(5), 341–349. https://doi.org/10.5951/jresematheduc.13.5.0341

Hong, D. S., Choi, K. M., Runnalls, C., & Hwang, J. (2019). How well aligned are common core textbooks to students' development in area measurement? *School Science and Mathematics*, *199*(5), 240–254. https://doi.org/10.1111/ssm.12336

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38*(3), 258–288.

Jerman, M. E., & Mirman, S. (1974). Linguistic and computational variables in problem solving in elementary mathematics. *Educational Studies in Mathematics*, *5*(3), 317–362. https://doi.org/10.1007/BF01424553

Larson, N., & Matthews, L. (2012). *Saxon math 1*. Houghton Mifflin Harcourt.

Lewis, A. B., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology*, *79*(4), 361–371. https://doi.org/10.1037/0022-0663.79.4.363

McNeil, N. M., Grandau, L., Knuth, E. J., Alibali, M. W., Stephens, A. C., Hattikudur, S., & Krill, D. E. (2006). Middle-school students' understanding of the equal sign: The books they read can't help. *Cognition and Instruction*, *24*, 367–385. https://doi.org/10.1207/s1532690xci2403_3

Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage Publishers.

NCTM (National Council of Teachers of Mathematics). (1989). *Curriculum and evaluation standards for school mathematics*. Author.

Nesher, P., Greeno, J. G., & Riley, M. S. (1982). The development of semantic categories for addition and subtraction. *Educational Studies in Mathematics*, *13*(4), 373–394. https://doi.org/10.1007/BF00366618

NGA & CCSSO (National Governors Association Center for Best Practices & Council of Chief State School Officers). (2010). *Common Core State Standards for Mathematics*. Author.

Olkun, S., & Toluk, Z. (2002). Textbooks, word problems, and student success on addition and subtraction. *International Journal for Mathematics Teaching and Learning*, *3*, 1–7. http://www.cimt.org.uk/journal/olkuntoluk.pdf

Polikoff, M. S. (2015). How well aligned are textbooks to the Common Core standards in mathematics? *American Educational Research Journal*, *52*(6), 1185–1211. https://doi.org/10.3102/0002831215584435

Powell, S. (2012). Equations and the equal sign in elementary mathematics textbooks. *The Elementary School Journal*, *112*(4), 627–648. https://doi.org/10.1086/665009

Remillard, J. T., Harris, B., & Agodini, R. (2014). The influence of curriculum material design on opportunities for student learning. *ZDM—International Journal on. Mathematics Education*, *46*(5), 735–749. https://doi.org/10.1007/s11858-014-0585-z

Remillard, J. T., & Heck, D. J. (2014). Conceptualizing the curriculum enactment process in mathematics education. *Zdm—international Journal on Mathematics Education*, *46*(5), 705–718. https://doi.org/10.1007/s11858-014-0600-4

Reys, B., Reys, R., & Rubenstein, R. (2010). *Mathematics curricula: Issues, trends, and future directions*. National Council of Teachers of Mathematics.

Riley, M. S., & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction*, *5*(1), 49–101. https://doi.org/10.1207/s1532690xci0501_2

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–196). Academic Press.

Schoen, R. C., LaVenia, M., & Tazaz, A. M. (2018). *Effects of the first year of a three-year CGI teacher professional development program on grades 3–5 student achievement: A multisite cluster-randomized trial* (Research Report No. 2018–25). Florida State University.

Schoen, R. C., LaVenia, M., Tazaz, A., Farina, K., Dixon, J. K., & Secada, W. G. (2020). *Replicating the CGI experiment in diverse environments: Effects on grade 1 and 2 student mathematics achievement in the first program year* (Research Report No. 2020–02). Florida State University. https://doi.org/10.33009/fsu.1601237075.

Stigler, J. W., Fuson, K. C., Ham, M., & Kim, M. S. (1986). An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. *Cognition and Instruction*, *3*(3), 153–171. https://doi.org/10.1207/s1532690xci0303_1

Tran, D. (2016). Statistical association: Alignment of current U.S. high school textbooks with the common core state standards for mathematics. *School Science and Mathematics*, *116*(5), 286–296.

Verschaffel, L., & De Corte, E. (1993). A decade of research on word-problem solving in Leuven: Theoretical, methodological, and practical outcomes. *Educational Psychology Review*, *5*(3), 1–18. https://doi.org/10.1007/BF01323046

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole numbers concepts and operations. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. National Council of Teachers of Mathematics.

Xin, Y. P. (2007). Word problem solving tasks in textbooks and their relation to student performance. *The Journal of Educational Research*, *100*(6), 347–359. https://doi.org/10.3200/JOER.100.6.347-360

How to cite this article: Schoen RC, Champagne Z, Whitacre I, McCrackin S. Comparing the frequency and variation of additive word problems in United States first-grade textbooks in the 1980s and the Common Core era. *School Science and Mathematics*. 2021;121:110–121. https://doi.org/10.1111/ssm.12447