International Journal of Educational Research 99 (2020) 101493

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

Examining discourse structures in Chinese and U.S. elementary mathematics classes

Shuai Wang[a], Michelle Perry[b,*], Leigh Mingle[c], Marc McConney[b]

[a] SRI International, 1100 Wilson Blvd # 2800, Arlington, VA 22209, USA
[b] University of Illinois at Urbana-Champaign, Department of Educational Psychology, 1310 S. Sixth Street, Champaign, IL 61820-2990, USA
[c] Ensemble Learning, 12 Sierra Avenue, Piedmont, CA 94611, USA

ARTICLE INFO

Keywords:
Classroom discourse
Cross-cultural comparison
Elementary mathematics
Student-teacher interaction

ABSTRACT

Public policy, theory, and empirical research suggest that mathematical learning is supported in U.S. elementary classrooms when students follow “horizontal” discourse patterns. In contrast, in Chinese classrooms, where students have been reported to excel in mathematics, both tradition and expectations necessitate that teachers remain in control and follow “vertical” discourse patterns. The investigation reported here examined these two patterns and an additional, hybrid pattern—teacher-facilitated horizontal discourse—in transcripts from 31 upper elementary U.S. and Chinese fractions lessons. Using a generalized linear mixed model, results indicated that although vertical discourse was the predominant pattern in both samples, horizontal discourse was more common in U.S. classrooms and teacher-facilitated horizontal discourse was more common in Chinese classrooms.

1. Introduction

Evidence suggests that, in general, Chinese students outperform U.S. students in mathematics (see, e.g., Gonzales et al., 2009; Organisation for Economic Co-operation & Development, 2012a, 2012b, 2013; Wang & Lin, 2005). However, the question of why this difference exists has not yet been settled and, although we do not intend to settle this issue in this report, we examine educational practices that likely impact student learning. Among the competing hypotheses, we focus on those that implicate classroom practices because these proximal causes are potentially malleable in ways that more distal causes (e.g., poverty, family composition, cultural press for academic success) are not (e.g. Schleppenbach, Perry, Miller, Sims, & Fang, 2007). And, among the practices that have been underexplored, we have chosen to examine one that has been promoted as productive in policy and theory, but which has only rarely been examined from a cross-cultural perspective in practice: *horizontal discourse*, the practice in which students talk directly with each other to work out mathematical ideas.

In general, reform efforts and research indicate that horizontal discourse supports student learning (e.g., Inagaki, Hatano, & Moritas, 1998; Nathan & Knuth, 2003; National Council of Teachers of Mathematics, 2000, Ni et al., 2017; Zhao, Mok, & Cao, 2016). Intriguingly, Chinese students are well known for their deep respect of their teachers, and we would expect their teachers, following from the Confucius tradition (Leung, 2001; Wong, 2004), to be the main authority on mathematics in the classroom and to control the classroom discourse. In classrooms like this, the discourse is presumed to be between the teacher and students, which has been termed *vertical discourse*. In contrast, U.S. reform efforts and U.S. policy (e.g., Common Core State Standard Initiatives, 2014; National Council of Teachers of Mathematics, 2000), which echoes U.S. democratic principles, push for student-centered learning and for U.S.

* Corresponding author.
E-mail address: mperry@illinois.edu (M. Perry).

https://doi.org/10.1016/j.ijer.2019.10.003
Received 15 July 2019; Received in revised form 27 September 2019; Accepted 8 October 2019

students to have a voice in the lesson and share ideas, which would lead to relatively more student-to-student discourse when compared to their Chinese peers.

A potential contradiction exists between what we expect theoretically—that horizontal discourse supports student learning and that we should find more horizontal discourse in U.S. than in Chinese classrooms—and what we know empirically—that upper-elementary Chinese students generally outperform their U.S. peers in mathematics. One hypothetical explanation for this contradiction is that our expectation does not correspond with what is actually happening in the classrooms. For example, research comparing student talk in U.S. and Chinese classrooms (e.g., Sims, 2008) found that Chinese elementary students talk more and that their talk is more mathematical than U.S. students, indicating that perhaps Chinese teachers encourage students to talk more than we would expect (e.g., Li, 2011). In addition, although research has been conducted to investigate horizontal discourse and vertical discourse within these cultures separately (e.g., Li, 2003, 2005; Nathan & Knuth, 2003; Wang & Cai, 2007), Chinese and U.S. horizontal and vertical discourse have not been directly compared empirically. The current study examines these concomitantly, with an explicit cross-cultural perspective.

A second explanation for this contradiction is that the dichotomous categorization of horizontal and vertical discourse could be missing aspects of discourse not covered by this dichotomy. For example, research has highlighted that Chinese students rarely respond directly to each other—that is, they rarely engage in horizontal discourse—unless sanctioned by the teacher (e.g., Chan, 1999; Li, 2003, 2005; Wang & Cai, 2007) and thus the teacher plays a prominent role in this discursive form. This may be a function of the curriculum reforms implemented in China, which encourage interactive and open classroom discourse where conversation between students and teachers varies more than before the curriculum reforms were implemented (see, e.g., Ni, Zhou, Li, & Li, 2014). In the current study, we propose a relatively nuanced stance toward thinking about how teachers lead their students in mathematical discourse. In particular, we examine both ends of the continuum—vertical, teacher-led discourse in which the students respond directly to the teacher, and horizontal, student-student discourse in which students respond directly to each other—as well as a practice that sits between these, a practice we label *teacher-facilitated horizontal discourse*.[1] In this discursive form, we see the teacher at the helm, directing students to respond not to the teacher but to other students.

A third explanation for the discrepancy would be that different structures of discourse could play different roles in different cultures, which could lead to different levels of effectiveness of each discursive type in different cultures. For example, it has been argued that Chinese teachers who command respect get students to pay attention to the teachers' directives and this obedience leads students to study hard and perform well (e.g., Li, 2001), which makes vertical discourse potentially a more effective practice to promote learning in Chinese lessons than in U.S. lessons. Thus, different discursive moves should be examined and compared within their respective cultures. The current study compares the use of vertical discourse, horizontal discourse, and teacher-facilitated horizontal discourse, within both U.S. and Chinese lessons separately.

A final potential explanation would be that the discourse in Chinese classrooms has little to do with Chinese students' academic performance. It could be that Chinese teachers are not promoting effective classroom discourse, but are doing well in other respects (e.g., classroom management). Although this is certainly possible, this does not discount the importance of discourse, given that numerous previous studies have associated student discourse with student learning (e.g., Nathan & Knuth, 2003). More importantly, significant research has been generated through cross-cultural comparisons (e.g., Perry, Vanderstoep, & Yu, 1993), and new findings may be discovered about classroom discourse in this context, which further necessitates the current study.

In sum, the present study asks what the structure of classroom discourse looks like in both U.S. and Chinese mathematics lessons, and, more particularly, what sort of discursive structures promote student mathematical talk. We do this in light of recent theorizing by Sfard (2017) in her framework for how students comport themselves in classrooms, following the structure that teachers provide, by trying "to participate in the new discourse even if, at a given point in time, they can do it only by imitating the moves of an expert participant" (p. 45). Following Sfard (2017), we take seriously the sociocultural position (e.g., Vygotsky, 1978) in which ideas that develop within individuals have their roots in the socially accepted knowledge expressed in public discourse (also see, e.g., Ni et al., 2017). We take our findings to develop conjectures about how the amount of mathematical talk plays a role in student learning.

1.1. Background: policy and historical views on student-student mathematical talk

In the United States, the Common Core State Standards Initiative (Common Core State Standard Initiatives, 2014), Standards for mathematical practice—and prior to the CCSSI, the National Council of Teachers of Mathematics' (National Council of Teachers of Mathematics, 1991, 2000) Standards—have argued for students' roles in contributing to the mathematical discourse in elementary school classrooms, and for the discourse to be structured as student-to-student rather than student-to-teacher. The CCSSI Standards for Mathematical Practice state that students should: "construct viable arguments and critique the reasoning of others"

[1] Although the term we use here, *teacher-facilitated horizontal discourse,* is new, related discursive moves have been described and examined in previous literature (including re-voicing and redirecting, e.g., Forman & Ansell, 2002; O'Connor & Michaels, 1993; Riley, 1981; Wright & Nuthall, 1970). However, what we describe here introduces a different discursive function. Re-voicing, in which the teacher restates a student's idea, often using more precise language, has the function of reinforcing a student's contribution. We do not include re-voicing in our definition of teacher-facilitated horizontal discourse. Redirecting, in which the teacher turns the same question to another student, functions to give another student a chance to produce the correct response. Teacher-facilitated horizontal discourse includes redirecting, but goes beyond redirecting by also including attempts to get students to evaluate, expand, or contradict another student, not just produce a correct answer when the first student failed to produce a correct answer.

(CCSS.Math.practice.MP3). They go on to say "Mathematically proficient students… justify their conclusions, *communicate them to others, and respond to the arguments of others*…" (emphasis added).

This does not appear markedly different from what the National Council of Teachers of Mathematics (1991) and the National Council of Teachers of Mathematics (2000) urged, prior to the CCSSI. For example, in their *Principles and Standards of School Mathematics*, the National Council of Teachers of Mathematics (2000) stated:

> To support classroom discourse effectively, teachers must build a community in which students will feel free to express their ideas. … Starting in grades 3–5, students should gradually take more responsibility for participating in whole-class discussions and *responding to one another directly*. They should become better at listening, paraphrasing, questioning, and interpreting others' ideas (p. 61, emphasis added).

Although there are other dimensions of discourse practices that promote student success, given the prevalence and consistency of this recommendation over the years and across guidelines, we assume that currently, and at least for the past several decades, U.S. elementary school teachers have been practicing in a national context that has promoted having students voice their understanding directly to each other.

Similarly, Chinese Mathematics Curriculum Standards for Compulsory Education (MCSFCE, Ministry of Education of the People's Republic of China, 2011) also has specifications about classroom discourse and student interactions. For instance, students are encouraged to "learn to collaborate with others," and "form the habit of cooperation and communication." The trial version of MCSFCE was first published in 2001 and only a limited number of Chinese school districts initially participated in the curriculum reform experiment. Furthermore, under the influence of deeply rooted Confucius tradition, teachers still are expected to control the classroom discourse; in Chinese primary schools with an average class size of 39 students and lower secondary schools with an average class size of 52 students (Organisation for Economic Co-operation & Development, 2012a, 2012b), unauthorized student discourse can be easily interpreted as classroom management issues. In this context, teachers often talk to students, but encounter difficulty advancing ideas through dialogic classroom discourse (Stein, Remillard, & Smith, 2007). Thus, we expect less student-student talk in Chinese classrooms compared to U.S. classrooms, and the current study examines empirical instructional practices against our understanding of policy and historical expectations across U.S. and Chinese cultures.

1.2. Theoretical and empirical support for student-to-student discourse

The policy urging students to talk directly to each other is supported by constructivist theory (Piaget, 1926, 1985) and socio-cultural perspective (Vygotsky, 1978). Theoretically, we have every reason to believe that learning is enhanced in an environment where students work to create that knowledge compared to information that is told to students (e.g., Boaler, 2000; Imm & Stylianou, 2012; Inagaki et al., 1998). In the context of learning in classrooms, telling students what they need to know (structured as teacher-to-student, or vertical, discourse) should be less effective than having students work through and develop their own understanding (e.g., Hiebert & Grouws, 2007). Discursive structures designed to enhance learning should have students working directly with each other rather than directly responding only to the teacher.

A long history of empirical research both from Asian perspectives (Hatano & Inagaki, 1991; Inagaki et al., 1998) and from U.S. perspectives (e.g., Ball, 1993; Bauersfeld, 1995; Hufferd-Ackles, Fuson, & Sherin, 2004; Lampert, 1990; Silver, Kilpatrick, & Schlesinger, 1990) provide support for the theory that students who work together through mathematical ideas have superior learning outcomes than when learning from teacher-led instruction. For example, Inagaki et al. (1998) found that learning was enhanced for students who contributed to a class-level discussion of mathematical concepts. Additionally, in a classic study, Lampert (1990) provided evidence for impressive learning opportunities when students were encouraged to work together to build a communal body of knowledge and to support and defend publicly made claims. From these research findings, we would expect that students who work together to create mathematical knowledge would outperform students who learn directly from a teacher. The current study examines these empirical practices concomitantly, with an explicit cross-cultural perspective.

1.3. Measuring student engagement with the mathematics

Measuring student engagement has been difficult. For instance, Ding, Pepin, and Jones, (2015) pointed out that standardized tests are inappropriate for measuring student engagement with mathematics, especially when it comes to Chinese classrooms, because although Chinese students usually perform well on these assessments, they may not feel confident and interested in mathematics. Accordingly, in this investigation, we were interested in the structure of the discourse not only as an end, but also as it may provide supports to student engagement with the mathematics. As level of engagement with mathematical ideas is notoriously problematic to measure (Gettinger & Walter, 2012), we chose an easy-to-recognize behavior as a proxy for engagement with the mathematics: use of mathematical terms. The reason for this focus is that voicing mathematical ideas serves as both an opportunity for building a community's knowledge base and for formative assessment of what else students might need to know (Black & Wiliam, 1998; Clarke & Xu, 2008). In contrast to the use of clear mathematical terms, the use of vague (and non-mathematical) terms may undermine others' access to interact and participate meaningfully (Lack, Swars, & Meyers, 2014).

1.4. Statement of the problem

Based on theory (e.g., Hiebert & Grouws, 2007), policy (e.g., National Council of Teachers of Mathematics, 2000), and empirical results (e.g., Lampert, 1990, cf. Nathan & Knuth, 2003), we might expect certain cultural practices to correlate with certain outcomes in elementary mathematics classrooms. These converging practices might lead us to expect that U.S. elementary teachers would encourage students to take on each other's contributions and talk directly to each other, in *horizontal* discourse structures, and Chinese students to address the teacher directly rather than each other, in *vertical* discourse structures. But does this happen? We have little evidence about the frequency with which these practices are enacted in U.S. and Chinese classrooms, and thus we explore this empirical issue by looking at how frequently students respond directly to students in horizontal discourse versus respond directly to the teacher in vertical discourse structures.

To understand how these forms work, we take a comparative look at U.S. and Chinese classrooms because Chinese culture, with its reverence for elders and authority, places the teacher at the helm of the class, with the responsibility to deliver content (e.g., Pratt, 1992). Because Chinese students have been outperforming U.S. students in mathematics for decades (see, e.g., Stevenson et al., 1990), a potential contradiction exists between what we expect theoretically—that horizontal discourse supports student learning and that we should find more horizontal discourse in U.S. than in Chinese classrooms—and what we know empirically—that upper-elementary Chinese students outperform their U.S. peers in mathematics.

We realize that the apparent contradiction may have been fueled by an overly simplistic dichotomy between horizontal and vertical discursive structures (e.g. Hansson, 2010). Given this realization, we propose here a more nuanced stance toward thinking about how teachers lead their students in mathematical discourse. In particular, we examine both ends of the continuum—vertical, teacher-led discourse in which the students respond directly to the teacher, and horizontal, student-student discourse in which students respond directly to each other—as well as a practice that sits between these, a practice we label *teacher-facilitated horizontal discourse*. In this discursive form, we see the teacher at the helm, directing students to respond not to the teacher but instead to other students.

If we are also questioning what should happen, in the sense of what is most likely to promote student learning, we need to ask more than just how often we observe these discursive forms in elementary mathematics classrooms. Thus, we have chosen to look at how students' responses may differ, in their production of mathematical language, when the prompt for the response comes directly from the teacher, directly from another student, or from another student but facilitated through the teacher. We chose this particular way of examining student input based on research that connects student discussion (Hiebert et al., 2003; Stigler & Hiebert, 1999) and student precise and explicit mathematical language (Clarke & Xu, 2008; Nathan & Knuth, 2003; Sfard & Kieran, 2001; Sims, 2008) to successful mathematical lessons and to student learning.

The present study investigates both the frequency of these three types of discourse (horizontal, vertical, and teacher-facilitated horizontal) in U.S and Chinese classrooms as well as the extent to which these types of discourse structures promote the production of mathematical language, as a proxy for engagement with mathematics. More specifically, we ask two research questions:

1 What are the frequency differences between U.S. and Chinese lessons, and within each site, among the three types of discourse?
2 What are the frequency differences of mathematical terms—between U.S. and Chinese lessons, and within each site, among the three types of discourse?

2. Method

2.1. Data source

2.1.1. Background

The data reported here were collected as part of a larger project, in which we sought to develop effective ways to use video records of classroom processes as a means to help teachers reflect on the processes of teaching and learning (Schleppenbach et al., 2007). In this project, our colleagues videotaped one lesson from each of 31 fourth- and fifth-grade classrooms: 14 from the United States (from 6 schools in and within a 30-mile radius of a university town in the Midwest) and 17 from China (from 8 schools in Beijing). We readily admit that it is not fair to generalize from a relatively small cross-cultural data set to draw inferences beyond these data. Still, we can hope to learn about what these patterns of discourse look like by taking a careful look within this relatively small sample.

2.1.2. Classroom characteristics and length of observations

The average class size for the U.S. classes was 22 students. The U.S. teachers taught a variety of subjects including mathematics. The average class size for the Chinese classes was 55 students. The Chinese teachers were responsible only for teaching mathematics. We did not find any significant difference in the average length of U.S. lessons (40.24 min, *s.d.* = 7.86 min) and Chinese lessons (43.11 min, *s.d.* = 3.26 min), $t(15) = 1.23$, $p = .24$.

2.1.3. Details of video recording

All of the lessons were video recorded in the spring semester, the second semester of the school year in both locales. In each lesson, two videographers were present in the classroom so that the video records captured the class from both the teacher's and the students' perspectives. These lessons were transcribed, and all of the Chinese lessons were translated to English and back-translated for accuracy.

2.1.4. Lesson topics

Before observing each lesson, the researcher arranged with the teacher to observe a lesson about equivalent or adding fractions. Equivalent and adding fractions were chosen because these topics are central to mathematics curricula in both countries and are relatively demanding for students in both countries. We also thought it was crucial to control for the topics because different topics might differentially constrain discourse and we wished to curtail this potential confound as much as possible.

The Chinese teachers introduced these related topics exclusively in the fifth grade, whereas some of the U.S. teachers introduced these in fourth and others in fifth grade. Although we realized that some of the observed differences might be due to differences in age of students in the two sites, we reckoned that this was less important than keeping the topics consistent in our observations. In addition, the NCTM lays out suggestions for communication over a range of grades, expecting all teachers, from third through fifth grade, to approach teaching materials in similar ways (National Council of Teachers of Mathematics, 2000). Thus, we expected students from both grades we observed, at least in the United States, to be similar with respect to their capabilities to communicate mathematically, which was the focus of this investigation.

2.2. Coding

We used the videos and transcripts to code the data. To begin coding, the classroom discourse that involved classroom management and classroom routines was eliminated from analysis, because the focus of this study was to investigate student discourse patterns when talking about mathematics. For example, when the teacher asked the student to close the book, and the student said "okay," this was not a case in our analysis, as it was part of classroom routines. Next, the lessons were divided into whole-class and non-whole-class portions of the lesson. To compare the forms of discourse, only whole-class portions were used, for three reasons: (1) it was often hard to hear conversations during non-whole-class time, which was typically spent with students working alone or in small groups, so we could not be certain we were consistently hearing and coding whether the students were even talking to each other, which could have rendered our results unreliable; (2) by only focusing on whole-class time, it was possible to observe each type of discourse; it would have been impossible for vertical or teacher-facilitated horizontal discourse to occur during these non-whole-class moments, thereby potentially inflating our reporting of horizontal discourse; and (3) if the majority of the class cannot hear a particular discussion, then this type of interaction can influence only those students within earshot, and cannot have the effect of influencing anyone else in the class. As further support for omitting non-whole-class time, we note that we did not disadvantage one site over the other: we found no significant difference between the average length of the non-whole class portions between the Chinese (4.70 min, $s.d. = 4.10$ min) and U.S. lessons (1.92 min, $s.d. = 3.86$ min), $t(28) = 1.88$, $p = .07$, which amounted to small portions of overall class time in both locations. We noted that one U.S. class was conducted exclusively using non-whole class time. Because we could not observe two of the three forms of discourse we were investigating here, we chose to exclude this class from our analyses.

2.2.1. Whole-class interactions

In the portions of each lesson that were identified as whole-class time, we coded each student statement to identify to whom the student was directing that statement. We identified three forms of discourse during whole-class interactions, which depended on to whom the student was responding: vertical (the student statement was directed only towards the teacher), horizontal (the student statement was directed towards another student), or teacher-facilitated horizontal (the student statement was directed by the teacher to respond to another student). For clarity, we describe and exemplify these three forms of discourse.

2.2.2. Vertical discourse

This was coded whenever the teacher or student initiated and the other responded. Here is an example of vertical discourse in a U.S. classroom:

Teacher: Krista, what did you find?
Krista: Two-fourths.
Teacher: So we can write down one-half equals….
Krista: Two-fourths.

Here is an example of vertical discourse in a Chinese classroom:

Jia: I choose four-ninths and two-ninths.
Teacher: Okay. Do you want to make up an addition or a subtraction problem?
Jia: Subtraction.
Teacher: Go ahead.
Jia: Four-ninths minus two-ninths equals two-ninths.

2.2.3. Horizontal discourse

In horizontal discourse, students responded directly to each other, without teacher intervention. Here is an example of horizontal discourse in a U.S. lesson (we include the question from the teacher, which provides context for the horizontal response by Kimberly):

Teacher: 2.7, right?

Jonathan: Don't you have to put a zero there? [as in 2.70]
Kimberly: No.

Here is an example of horizontal discourse in a Chinese lesson (here, again, we provide the comment from the teacher as context):

Teacher: Very good. The third question is a little bit difficult. Who would like to try it? Be more active. Xiu Li.
Xiu Li: Four-fifths minus one-sixth equals twenty-four thirtieths minus six-thirtieths, which equals eighteen-thirtieths.
Zhi: No, it's wrong.

In both of these examples, the second student responded directly to the first student, without any intervention from the teacher.

2.2.4. Teacher-facilitated horizontal discourse

We identified this discursive form when, instead of the student responding directly to the teacher or directly to another student, the teacher intervened and asked for one student to respond to another student. This form of discourse includes the initial student's statement, the teacher's facilitation discourse, as well as another student's response. Here is an example from a U.S. lesson (with the initial student's comment and the responding student's comment noted in bold):

Teacher:Art, how many centimeters is one inch?
***Art*: Three and two-tenths.**
Teacher: He says it's three and two-tenths. Tom's shaking his head "no." Tom what do you say?
***Tom*: Three and a half.**

In this excerpt, the students were not talking directly to each other, but communicating with each other *through* the teacher.
Here is an example from a Chinese lesson:

Teacher: Can you tell me the rationale of consistent quotient?
Jing: The quotient will stay consistent if two numbers are multiplied or divided by the same number.
Teacher: Sit down, please. Does anyone have different opinions? You, please.
***Shi:* The quotient will stay consistent if two numbers are multiplied or divided by the same number at the same time.**
Teacher: Ok. Anything else? You, please.
***Chun:* Except zero.**

In this example, two instances of teacher facilitations were involved: (1) "Does anyone have different opinions? You, please," and (2) "Ok. Anything else? You, please," so we coded two separate instances of teacher-facilitated horizontal discourse. In this excerpt, the second student, Shi, responded to the first student, Jing, with the teacher's facilitation. Also, the third student, Chun, responded to the first two students, and made a complement to the first two students' answers, with the teacher's facilitation. In these instances, the teacher was intervening in the discussion to ask a student to respond to the ideas of another student, which is what we intend to capture when coding teacher-facilitated horizontal discourse.

2.3. Reliability

Two authors coded a random 25 % of the lessons (4 Chinese and 3 U.S. lessons). Disagreements were discussed with a third author until all three of us came to an agreement. Reliability for identifying the student statements as vertical, horizontal, or teacher-facilitated horizontal, using Cohen's (1960) kappa, was 0.86.

2.4. Analytic plan

We conducted analyses in two phases. First, we identified and counted the frequencies of the three forms of discourse. Using a generalized linear mixed model with log transformation, we then looked for significant differences in the frequencies between U.S. and Chinese classrooms. We performed log transformations because the residuals were not normally distributed and the residual variance was not constant. We chose a mixed-effects model because this allowed us to take into account the clustering feature of the data set (different forms of discourse, nested within classes). Our mixed-effects model consisted of fixed-effect terms (country and discursive form) and a random-effect term (classes). By treating the classes as a random effect, we allowed variations among the classes, but without estimating an effect for each class. We report both the between and within degrees of freedom because the counts of discourse differed greatly across lessons.

Second, following Clarke and Xu (2008), we settled on using the number of different mathematical terms provided in the student utterance as our outcome measure. We operationalized *mathematical terms* as the number of words that were numbers, operations, and mathematical vocabulary included in the student's response. Our intent in doing this was to take a more careful look at the how each of the three discursive structures was related to student engagement with the mathematics, using the prevalence of mathematical terms used by students in their responses as a proxy for engagement with the mathematics.

As an example, when a teacher asked "is 2/6 equivalent to 1/3?" one student answered "yes, that is exactly what I got. *2/6* can be considered the *same as 1/3*" (15 words total, 3 mathematical terms, 20 % of words are mathematical terms), while the other student answered, "the *numerator* and *denominator* are *divided by 2* simultaneously, so the *fraction* stays *the same*" (15 words total, 5 mathematical terms, 33.33 % of words are mathematical terms). Both responses were examples of vertical discourse, but the second

Table 1
Descriptive Statistics for Number of Instances of Three Discursive Forms in U.S. and Chinese Mathematics Lessons.

	Vertical		Horizontal*		Teacher-Facilitated Horizontal*	
Country	N (%)	*M* (*SD*)	N (%)	*M* (*SD*)	N (%)	*M* (*SD*)
U.S.	1220 (91.39)	93.85 (45.26)	45 (3.37)	3.46 (7.09)	70 (5.24)	5.38 (5.98)
China	2006 (88.37)	118.00 (33.58)	14 (0.62)	0.82 (1.01)	250 (11.01)	14.71 (8.41)

Note. Reported percentages indicate conditional proportions (i.e., percent within the data from each country).
* $p < .001$, comparing U.S. and Chinese classrooms.

student provided much richer mathematical information than the first student. Thus, by examining the density of mathematical terms in the student responses, we hope to capture some measure of the student's engagement with the mathematics.

3. Results

3.1. Overview

We found a total of 3605 instances of discourse: 1335 in the 13 U.S. lessons and 2270 in the 17 Chinese lessons. The overwhelming majority of these were vertical. We found a total of 3226 instances of vertical discourse: 1220 in U.S. lessons and 2006 in Chinese lessons. We found relatively few instances of horizontal discourse. Only 59 were horizontal: 45 in U.S. lessons and 15 in Chinese lessons. We also discovered 320 instances of teacher-facilitated horizontal discourse: 70 in U.S. lessons and 250 in Chinese lessons. See Table 1 for a summary of frequencies and proportions of these discursive forms.

3.2. Between-country findings

Because we made multiple comparisons, we corrected our analyses using Bonferroni adjustments and all *p*-values we report include these corrections. Accordingly, given that we conducted analyses to compare three types of discourse—horizontal, vertical, and teacher-facilitated-horizontal discourse—the *p*-values resulting from the generalized linear mixed model were multiplied by 3.

We did not find significant differences in vertical discourse between the U.S. lessons and the Chinese lessons, $F(1, 56) = 1.36, p = .74$. We did, however, find significantly more horizontal discourse in the U.S. lessons than in the Chinese lessons, $F(1, 56) = 8.53, p = .015$. We found significantly more teacher-facilitated horizontal discourse in the Chinese lessons than in the U.S. lessons, $F(1, 56) = 17.09, p < .001$, (see Table 1). We present the estimated generalized linear mixed model for these findings in the Appendix.

3.3. Production of mathematical terms found in each discursive pattern

We also sought to characterize the degree to which student responses included mathematical talk. In particular, we were interested in how student responses to each other might be affected by teacher intervention. To examine this issue, we relied on an index capturing the student's adherence to mathematical discourse by counting the number of distinct mathematical terms the student uttered in a response.

Using generalized linear mixed models and lsmeans statements in SAS, we examined differences between U.S. and Chinese lessons in instances of vertical[2], horizontal, and teacher-facilitated-horizontal discourse (see the Appendix). We began by looking at differences between sites. We did this to examine whether production of mathematical terms in student responses was notably different, even when responding within the same discursive structure, between U.S. and Chinese students. We then looked at differences within site, depending on whether the students responded to the teacher, responded directly to each other, or responded to each other via their teacher's intervention. We report Bonferroni-adjusted *p*-values because we made multiple comparisons.

3.4. Production of mathematical talk in U.S. and Chinese classrooms

3.4.1. Vertical discourse

We found significantly more mathematical terms in the Chinese students' responses made directly to their teachers than in the U.S. students' responses made directly to their teachers, $F(1, 38) = 24.80, p < .001$ (see Table 2 and Appendix Table A2).

3.4.2. Horizontal discourse

When considering responses that students provided directly to other students, without teacher intervention, we found no significant differences between U.S. and Chinese students in the number of mathematical terms, $F(1, 38) = 1.80, p = 1.0$ (see Table 2

[2] Given the vast amount of vertical discourse, and the multitude of ways in which it might be used, we chose not to examine all vertical discourse in our sample but rather only vertical discourse that was adjacent to teacher-facilitated horizontal discourse, as we reckoned that this would be a fair comparison between discourse elicited by the teacher that was either directed back to the teacher vs. to another student.

Table 2
Production of Mathematical Terms in Different Discursive Structures.

Country	Vertical* M (SD)	Horizontal M (SD)	Teacher- Facilitated* M (SD)
U.S.	1.43 (1.26)	1.47 (1.95)	1.67 (1.39)
China	3.71 (4.17)	2.29 (4.07)[†]	5.55 (5.44)[†]

* $p < .001$, comparing U.S. and Chinese lessons.
[†] $p = .01$, comparing responses within site.

and Appendix Table A2). This lack of significant distinction in horizontal discourse between the two sites indicates that, although the absolute frequency of this form of discourse differed, the way in which students talked when speaking directly to other students did not differ in their production of mathematical terms.

3.4.3. Teacher-facilitated horizontal discourse

We found significant differences in the number of mathematical terms used in response to another student when the teacher intervened, with U.S. students producing significantly fewer mathematical terms than Chinese students, $F(1, 38) = 52.71, p < .001$ (see Table 2 and Appendix Table A2). This indicated that the U.S. students were less likely than the Chinese students to demonstrate responses with many mathematical terms when the teacher facilitated students talking to each other.

3.5. Production of mathematical talk within sites

3.5.1. Comparing vertical with horizontal discourse within U.S. lessons

U.S. students responded similarly, whether responding directly to the teacher (vertical discourse) or directly to each other (horizontal discourse) in their use of mathematical terms, $F(1, 38) = 0.03, p = 1$ (see Table 2 and Appendix Table A2).

3.5.2. Comparing vertical with teacher-facilitated horizontal discourse within U.S. lessons

We were also interested in seeing whether students responded differently when prompted by the teacher to respond directly to the teacher (vertical discourse) or to another student (teacher-facilitated horizontal discourse). We found no differences in the number of mathematical terms, $F(1, 38) = 0.45, p = 1$ (see Table 2 and Appendix Table A2).

3.5.3. Comparing horizontal with teacher-facilitated horizontal discourse within U.S. lessons

When responding to each other, U.S. students produced comparable amounts of mathematical terms, whether they responded directly to each other or whether the teacher intervened, $F(1, 38) = .18, p = 1$ (see Table 2 and Appendix Table A2).

3.5.4. Comparing vertical with horizontal discourse within Chinese lessons

Chinese students produced comparable amounts of mathematical terms, whether responding directly to the teacher (vertical discourse) or directly to each other (horizontal discourse), $F(1, 38) = 2.43, p = 1$ (see Table 2 and Appendix Table A2).

3.5.5. Comparing vertical with teacher-facilitated horizontal discourse within Chinese lessons

Chinese students responded with more mathematical terms when the teacher directed students to respond to each other (in teacher-facilitated horizontal discourse) compared to when the teacher expected the students to respond directly to the teacher (in vertical discourse), $F(1, 38) = 17.81, p < .001$ (see Table 2 and Appendix Table A2).

3.5.6. Comparing horizontal discourse with teacher-facilitated-horizontal discourse within Chinese lessons

When responding to each other, Chinese students' responses contained more mathematical terms when the teacher intervened, in teacher-facilitated-horizontal discourse, than when speaking directly to each other, in horizontal discourse, $F(1, 38) = 8.64, p < .05$ (see Table 2 and Appendix Table A2).

4. Discussion

In some ways, the results are not surprising. For example, we rarely saw horizontal discourse in Chinese classrooms, and we saw more horizontal discourse in U.S. lessons. This is consistent with our cultural expectations in U.S. and Chinese lessons. In other ways, what we found was striking: teacher-facilitated horizontal discourse was common in Chinese classrooms. This approach appears to capture the intent that horizontal discourse is designed to play in U.S. classrooms. This hybrid form may potentially be significant in helping to understand how Chinese teachers appear to engage their students to learn mathematics successfully, in ways theoretically supported, but at the same time maintaining control of classroom discourse, in concert with cultural expectations.

Before considering the implications of these findings, we return to the contradiction we raised at the beginning of this report: a contradiction exists between what we expect theoretically—that horizontal discourse supports student learning and that we should find more horizontal discourse in U.S. than in Chinese classrooms—and what we know empirically—that upper-elementary Chinese students generally outperform their U.S. peers in mathematics. Our thought about this contradiction, after examining the data and findings from our study, is that purely horizontal discourse is rare and teacher-facilitated horizontal discourse appears to be an age-appropriate medium for having elementary-school students voice their ideas and engage with their peers about mathematical ideas.

One implication of these findings is that teacher-facilitated horizontal discourse potentially provides students with opportunities to interrogate each other's mathematical ideas without taking the responsibility of initiating the challenge. Theoretically, this hybrid form may act like horizontal discourse, in that students must reckon with each other's ideas and, in this way, provoke cognitive discord that provides fertile ground for cognitive change (Piaget, 1985).

This implication—that teacher-facilitated horizontal discourse provides support that allows students to wrestle with each other's ideas without some of the downsides of actual horizontal discourse—has several corollaries. For example, teacher-facilitated horizontal discourse may permit students to concentrate on the mathematics and alleviate them from the potentially socially awkward or problematic position of challenging a peer directly. Alleviating the associated stress could allow students to perform to their potentials rather than taking up valuable space in working memory (Beilock & Carr, 2005; Häggström, 2008). Future research could examine the extent to which students feel more comfortable challenging another student's ideas when the teacher invites this challenge compared to when students must take on this role without the explicit and direct support of the teacher, and whether this impacts successful mathematics performance.

A further potential benefit is that, by having the teacher shape the student's response, the student can borrow the structure and sophistication from the teacher's prompt, thereby having the opportunity to augment their own contribution. This sort of phenomenon has been documented in informal teaching situations (e.g., Wood, Bruner, & Ross, 1976) and may be supporting student learning formally, in elementary mathematics classrooms. We have not examined our data for this, but it certainly is possible and thus deserves further exploration.

Another implication is that the differences we located—especially in which the Chinese students produce more mathematical talk when the teacher entreats students to respond to each other than when students respond directly to each other or when responding directly to the teacher—contributes to students' understanding and use of mathematics. The students' responses in horizontal discourse, whether in U.S. or in Chinese classrooms, seemed remarkably similar. And, when the teacher intervened in the U.S. lessons, there was no real change in student responses. However, when the teacher intervened in the Chinese lessons, we found dramatic differences in student responses: the responses included more than twice as many mathematical terms. Thus, the Chinese teachers' interventions appeared to elevate student responses in a way that the students appeared to become more engaged with the mathematics than if they responded directly to each other or directly to the teacher (see, e.g., Ni et al., 2014). Exactly why this happened in the Chinese sample, but not in the U.S. sample, and how the production of mathematical terms affects student learning is curious and certainly demands further inquiry.

Our final thoughts relate to the issue of whether different structures of discourse play different roles in different cultures. In our careful reading of these truly horizontal discourse moments in U.S. classrooms, students add quick quips, but do not typically appear to engage deeply in serious mathematics. As a counterpoint, Chinese teachers seem to have found a way to follow the spirit of this practice, while, at the same time, adhere to traditional Asian cultural practices. In teacher-facilitated horizontal discourse, teachers invite students to parlay with each other and continue to explore a mathematical point through to its logical conclusion. The teacher often appears to act as a legitimizer of conversation about the issue, while supporting students talking to each other.

In this study we examined differences between theoretically derived practices, providing a basis for understanding the differences between U.S. and Chinese students in their development of mathematical understanding and mathematics classrooms norms. Because we conducted the study using classroom observations, we cannot draw causal conclusions. Future explorations could examine causal relationships between structures of discourse and student achievement by means of carefully designed experimental or quasi-experimental studies. In addition, although we found variations in the use of classroom discourse structures in these two sites, we realize that the contexts and contents of discourse also play non-negligible roles. Thus, future analysis may include student background characteristics, such as age and gender, as well as teacher background characteristics, including education and teaching experience, in statistical models to increase prediction accuracy; detailed qualitative discourse analysis may also provide a benefit to our understanding of classroom discourse.

In sum, looks can be deceiving. The look and feel of a Chinese classroom is of a teacher-centered, not learner-centered, approach (see Schuh, 2004). But, by taking into account the nuances of the approach that Chinese teachers use, we come to a different conclusion: that the Chinese teachers are adept at holding control of the class while, at the same time, inviting students to respond to each other's ideas. Teachers who provide the structure for students to consider each other's ideas may be teaching their students not only how to respond, but also providing the structure for developing more sophisticated mathematical knowledge.

Acknowledgments

This work was supported by the National Science Foundation, Grant No. 0089293. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix A

Table A1
Generalized Linear Mixed Model Fixed Effects for Frequency of Discourse.

Effect	Level	*B*	*SE*	*t*	*p*	*Exp (B)*
Intercept		4.5	.18	24.81	< .01	90.00
Discourse Type	Vertical	0	–	–	–	1.00
	Horizontal	−3.59	.27	−13.64	< .01	.03
	TFH	−3.02	.23	−13.05	< .01	.05
Country	U.S.	0	–	–	–	1.00
	China	0.28	.24	1.17	.25	1.32
Interactions	Vertical*U.S.	0	–	–	–	1.00
	Horizontal*U.S.	0	–	–	–	1.00
	TFH*U.S	0	–	–	–	1.00
	Vertical*China	0	–	–	–	1.00
	Horizontal*China	−1.45	.41	−3.74	< .01	.23
	TFH*China	0.88	.29	3.45	< .01	2.41

TFH = Teacher-facilitated-horizontal discourse.
We performed log transformations during statistical modeling to achieve better model fit. The *B*s represent the coefficients for the model with the log values; *Exp (B)* are the exponentiated values of *B*, which we use to interpret frequencies and predicted counts.
Because we needed to select a reference for the fixed effects, we arbitrarily chose to set the United States and vertical discourse as the reference points for country and for discourse type. Following convention, these were set at 1. To make the table clearer, we will explain how we interpret and use these values, using teacher-facilitated-horizontal discourse responses in China as the example. In the Generalized Linear Mixed Model, we report that the estimate for teacher-facilitated-horizontal discourse in Chinese lessons is 2.41 (see the last line in the table presented above), which is actually the exponentiated estimate of the interaction between Chinese lessons and teacher-facilitated-horizontal discourse responses. (Note that we performed log transformations on the predicted values for better model fit and because of the particularities of the distribution, we exponentiated the log value to provide more interpretable values.) Coefficients that are added in the log-linear model must be multiplied once they have been exponentiated. Thus, to obtain the estimates for teacher-facilitated-horizontal discourse in Chinese lessons, we multiplied the intercept (90.27) with the estimate for teacher-facilitated-horizontal discourse responses (across both countries, 0.05), with the estimate for all responses in Chinese lessons (1.32), and with the estimate for the interaction between teacher-facilitated-horizontal discourse and responses in Chinese lessons (2.41), or 90.27*0.05*1.32*2.41 = 14.36. The result, 14.36, indicates that it is predicted that *on average* in the Chinese lessons (population), there are 14.36 teacher-facilitated-horizontal discourse episodes per lesson. In the sample data, we actually have a total of 250 teacher-facilitated-horizontal discourse episodes in Chinese lessons, and we have 17 Chinese lessons. Thus in our sample, we have *an average* of 250/17 = 14.7 teacher-facilitated-horizontal discourse instances in Chinese lessons. The difference between the predicted value (14.36) and the actual sample value (14.7) is extremely small, which is an indication that our model has good accuracy.

Table A2
Generalized Linear Mixed Model Fixed Effects for Mathematical Terms.

Effect	Level	*B*	*SE*	*t*	*p*	*Exp (B)*
Intercept		0.34	.18	1.92	.07	1.40
Discourse Type	Vertical	0	–	–	–	1.00
	Horizontal	0.04	.25	.16	.87	1.04
	TFH	0.14	.21	.67	.51	1.15
Country	U.S.	0	–	–	–	1.00
	China	0.98	.20	4.98	< .01	2.66
Interactions	Vertical*U.S.	0	–	–	–	1.00
	Horizontal*U.S.	0	–	–	–	1.00
	TFH*U.S	0	–	–	–	1.00
	Vertical*China	0	–	–	–	1.00
	Horizontal*China	−0.51	.39	−1.30	.20	.60
	TFH*China	0.26	.23	1.12	.27	1.30

TFH = Teacher-facilitated-horizontal discourse.
Interpretation for this table is the same as Table A1. In this analysis, we arbitrarily chose to set the United States and vertical discourse as the reference points for country and for discourse type.

Appendix B. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijer.2019.10.003.

References

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal, 93*, 373–397.

Bauersfeld, H. (1995). The structuring of structures: Development and function of mathematizing as a social practice. In L. Steffe, & J. Gale (Eds.). *Constructivism in education* (pp. 137–158). Hillsdale, NJ: Erlbaum.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science, 16*(2), 101–105.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Boaler, J. (2000). Mathematics from another world: Traditional communities and the alienation of learners. *The Journal of Mathematical Behavior, 18*(1), 1–19.

Chan, S. (1999). The Chinese learner-a question of style. *Education and Training, 41*, 294–304.

Clarke, D., & Xu, L. H. (2008). Distinguishing between mathematics classrooms in Australia, China, Japan, Korea and the USA through the lens of the distribution of responsibility for knowledge generation: Public oral interactivity and mathematical orality. *ZDM, 40*(6), 963–972.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurements, 20*, 37–46.

Common Core State Standard Initiatives (2014). *Standards for mathematical practice.* Retrieved fromhttp://www.corestandards.org/Math/Practice.

Ding, L., Pepin, B., & Jones, K. (2015). Students' attitudes towards mathematics across lower secondary schools in Shanghai. In B. Pepin, & B. Roesken-Winter (Eds.). *From beliefs to dynamic affect systems in mathematics education: Advances in mathematics education* (pp. 157–178). Switzerland: Springer International.

Forman, E. A., & Ansell, E. (2002). Orchestrating the multiple voices and inscriptions of a mathematics classroom. *Journal of the Learning Sciences, 11*(2–3), 251–274.

Gettinger, M., & Walter, M. J. (2012). Classroom strategies to enhance academic engaged time. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.). *Handbook of research on student engagement* (pp. 653–673). New York: Springer.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2009). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context.* Washington, DC: National Center for Educational Statistics.

Häggström, J. (2008). *Teaching systems of linear equations in Sweden and China: What is made possible to learn?* Gothenburg: University of Gothenburg.

Hansson, Å. (2010). Instructional responsibility in mathematics education: Modeling classroom teaching using Swedish data. *Educational Studies in Mathematics, 75*(2), 171–189.

Hatano, G., & Inagaki, K. (1991). Sharing cognition through collective comprehension activity. In L. Resnick, J. Levine, & S. Teasley (Eds.). *Perspectives on socially-shared cognition* (pp. 331–348). Washington, DC: American Psychological Association.

Hiebert, J., Gallimore, R., Garnier, H., Givving, K. B., Hollingsworth, H., Jacobs, J., ... Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study, (NCES 2003-013), U.S. Department of Education.* Washington, DC: National Center for Education Statistics.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.). *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.

Hufferd-Ackles, K., Fuson, K. C., & Sherin, M. G. (2004). Describing levels and components of a math talk learning community. *Journal for Research in Mathematics Education, 35*, 81–116.

Imm, K., & Stylianou, D. A. (2012). Talking mathematically: An analysis of discourse communities. *The Journal of Mathematical Behavior, 31*(1), 130–148.

Inagaki, K., Hatano, G., & Moritas, E. (1998). Construction of mathematical knowledge through whole class discussion. *Learning and Instruction, 8*, 503–526.

Lack, B., Swars, S. L., & Meyers, B. (2014). Low- and high-achieving sixth grade students' access to participation during mathematics discourse. *The Elementary School Journal, 115*(1), 97–123.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29–63.

Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics, 47*(1), 35–51.

Li, J. (2001). Chinese conceptualization of learning. *Ethos, 29*(2), 111–137.

Li, J. (2003). The core of Confucian learning. *The American Psychologist, 58*, 146–147.

Li, J. (2005). Mind or virtue: Western and Chinese beliefs about learning. *Current Directions in Psychological Science, 14*(4), 190–194.

Li, Y. (2011). Elementary teachers' thinking about a good mathematics lesson. *International Journal of Science and Mathematics Education, 9*(4), 949–973.

Ministry of Education of the People's Republic of China (2011). *Chinese mathematics curriculum standards for compulsory education.* Beijing: Beijing Normal University Press.

Nathan, M. J., & Knuth, E. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21*(2), 175–207.

National Council of Teachers of Mathematics (1991). *Professional standards for teaching mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Ni, Y., Ho, G., Cai, J., Cheung, A., Chen, G., & Ng, O.-L. (2017). Research protocol: Teacher interventions aimed at engaging students in dialogic mathematics classroom discourse. *International Journal of Educational Research, 86*, 23–35.

Ni, Y., Zhou, D., Li, X., & Li, Q. (2014). Relations of instructional tasks to teacher-student discourse in mathematics classrooms of Chinese primary schools. *Cognition and Instruction, 32*(1), 2–43.

O'Connor, M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology & Education Quarterly, 24*, 318–335.

Organisation for Economic Co-operation and Development (2012a). *Education at a glance 2012: OECD indicators.* Retrieved fromhttps://www.oecd-ilibrary.org/education/education-at-a-glance-2012_eag-2012-en.

Organisation for Economic Co-operation and Development (2012b). *PISA 2012 results.* Retrieved fromhttp://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm.

Organisation for Economic Co-operation and Development (2013). *Asian countries top OECD's latest PISA survey on state of global education.* Retrieved fromhttp://www.oecd.org/newsroom/asian-countries-top-oecd-s-latest-pisa-survey-on-state-of-global-education.htm.

Perry, M., VanderStoep, S. W., & Yu, S. L. (1993). Asking questions in first-grade mathematics classes: Potential influences on mathematical thought. *Journal of Educational Psychology, 85*, 31–40.

Piaget, J. (1926). *The language and thought of the child.* New York, NY: Routledge.

Piaget, J. (1985). *The equilibration of cognitive structures.* Chicago, IL: University of Chicago Press.

Pratt, D. D. (1992). Chinese conceptions of learning and teaching: A Westerner's attempt at understanding. *International Journal of Lifelong Education, 11*(4), 301–319.

Riley, J. P. (1981). The effect of preservice teachers' cognitive questioning level and redirecting on student science achievement. *Journal of Research in Science Teaching, 11*(4), 81–94.

Schleppenbach, M., Perry, M., Miller, K. F., Sims, L., & Fang, G. (2007). The answer is only the beginning: Extended discourse in Chinese and U. S. mathematics classrooms. *Journal of Educational Psychology, 99*, 380–396.

Schuh, K. L. (2004). Learner-centered principles in teacher-centered practices? *Teaching and Teacher Education, 20*(8), 833–846.

Sfard, A. (2017). Ritual for ritual, exploration for exploration or what the learners get is what you get from them in return. In J. Adler, & A. Sfard (Eds.). *Research for educational change: Transforming researchers' insights into improvement in mathematics teaching and learning* (pp. 41–63). London: Routledge.

Sfard, A., & Kieran, C. (2001). Cognition as communication, rethinking learning-by-talking through multi-faceted analysis of students' mathematical interactions. *Mind, Culture and Activity, 8*(1), 42–76.

Silver, E. A., Kilpatrick, J., & Schlesinger, B. (1990). *Thinking through mathematics: Fostering inquiry and communication in mathematics classrooms.* New York, NY: College Entrance Examination Board.

Sims, L. (2008). Look who's talking: Reflecting on differences in math talk in U.S. and Chinese classrooms. *Teaching Children Mathematics, 15*(2), 120–124.

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

Stevenson, H. W., Lee, S.-Y., Chen, C., Lummis, M., Stigler, J. W., Fan, L., & Ge, F. (1990). Mathematics achievement in China and the United States. *Child Development, 61*(4), 1053–1066.

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

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wang, T., & Cai, J. (2007). Chinese (Mainland) teachers' views of effective mathematics teaching and learning. *ZDM Mathematics Education, 39*(4), 287–300.

Wang, J., & Lin, E. (2005). Comparative studies on US and Chinese mathematics learning and the implications for standards-based mathematics teaching reform. *Educational Researcher, 34*(5), 3–13.

Wong, N.-Y. (2004). The CHC learner's phenomenon: Its implications for mathematics education. In L. Fan, N.-Y. Wong, J. Cai, & S. Li (Eds.). *How Chinese learn mathematics: Perspectives from insiders* (pp. 503–534). River Edge, NJ: World Scientific.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology, 17*(2), 89–100.

Wright, C. J., & Nuthall, G. (1970). Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons. *American Educational Research Journal, 7*, 477–491.

Zhao, W., Mok, I. A. C., & Cao, Y. (2016). Curriculum reform in China: Student participation in classrooms using a reformed instructional model. *International Journal of Educational Research, 75*, 88–101.