Journal of Educational Psychology
2007, Vol. 99, No. 2, 380–396

0022-0663/07/$12.00 DOI: 10.1037/0022-0663.99.2.380

The Answer Is Only the Beginning: Extended Discourse in Chinese and U.S. Mathematics Classrooms

Meg Schleppenbach and Michelle Perry
University of Illinois at Urbana–Champaign

Kevin F. Miller
University of Michigan

Linda Sims
University of Illinois at Urbana–Champaign

Ge Fang
Chinese Academy of Sciences

The authors investigated the use of a particular discourse practice—continued questioning and discussion *after* a correct answer was provided, which they called *extended discourse*—and examined the frequency and content of this practice in 17 Chinese and 14 U.S. elementary mathematics classes. They found that the Chinese classrooms had more, and spent more time in, extended discourse than did the U.S. classrooms. The content of these episodes differed: The Chinese classrooms focused more on rules and procedures than did the U.S. classrooms, whereas the U.S. classrooms focused more on computation than did the Chinese classrooms. These findings shed light on interesting practices of discourse in both countries and also have implications for current U.S. reforms in mathematics pedagogy.

Keywords: classroom discourse, elementary mathematics learning, Chinese and U.S. elementary classrooms

For more than a decade, mathematics educators have been concerned with how language affects student learning in mathematics and how discourse mediates what counts as mathematical knowledge for students and teachers in classrooms. Influential documents such as the National Council of Teachers of Mathematics's (1991, 2000) *Professional Standards for Teaching Mathematics* and *Principles and Standards for School Mathematics* have called for teachers to emphasize communication that allows students to develop conceptual, or so-called higher level, understanding of mathematics. According to these documents and other research on classroom discourse (e.g., Ball, 1993; Hiebert & Wearne, 1993; Kazemi, 1998; Kazemi & Stipek, 2001; Lampert, 1990, 1992; O'Conner, 1998; Whitenack & Yackel, 2002), such high-level communication consists, in large part, of encouraging students to present mathematical conjectures, pushing students to both explain and justify their conjectures to their colleagues, and otherwise promoting debate and discussion of mathematical ideas. At the very least, this body of work has indicated that extended conversations about mathematical ideas (as opposed to the simple statement and acceptance of "correct" answers) provide a necessary, but not sufficient, foundation for such high-level talk (Ball, 1991; Kazemi & Stipek, 2001).

At the same time that interest in classroom discourse in mathematics has risen, so has an alarm over the performance of U.S. students on cross-national comparisons of education. In the past, cross-national studies (e.g., Crosswhite, Dossey, Swafford, McKnight, & Cooney, 1985; Stevenson, Chen, & Lee, 1993) typically took the form of achievement tests, and indeed a robust group of cross-national studies has indicated that a number of countries outperform the United States on standardized achievement tests in mathematics. More recently, large-scale video studies such as the Third International Mathematics and Science Study have indicated not only achievement differences between countries but also fairly substantial cultural differences in teaching practice (Hiebert & Stigler, 2000; Stigler & Hiebert, 1997, 1999, 2004). Of course, cultural differences in beliefs and attitudes toward education make it difficult to assess whether the teaching practices in other countries are actually responsible for student performance.

Assuming that teaching is, at least partially, culturally embedded (Santagata & Stigler, 2000), it would be difficult or perhaps even impossible to directly transfer practices from one country to another. Despite these difficulties, mathematics educators have become increasingly interested in what the examination of teaching practices allows in terms of the discovery of new and poten-

Meg Schleppenbach, Michelle Perry, and Linda Sims, Department of Educational Psychology, University of Illinois at Urbana–Champaign; Kevin F. Miller, Combined Program in Education and Psychology, University of Michigan; Ge Fang, Institute of Psychology, Chinese Academy of Sciences.

This article is based upon work supported by National Science Foundation Grant 0089293. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We thank the teachers and students who participated in this research and Michele Crockett for her comments on this article. In addition, we thank Sujai Kumar for his invaluable assistance in applying dynamic time-warping analysis to this data set.

We reported portions of this study in a symposium at the American Educational Researchers Association annual meeting in April 2005 in Montreal, Quebec, Canada.

Correspondence concerning this article should be addressed to Meg Schleppenbach, Department of Educational Psychology, University of Illinois at Urbana–Champaign, Champaign, IL 61820. E-mail: megschleppenbach@gmail.com

tially useful practices, as well as the illumination of currently used practices that would likely go unnoticed without a comparative lens to bring them into focus.

The research described in this article combines these current interests in classroom discourse and cross-cultural comparisons in mathematics education. Looking at video from U.S. and Chinese lessons, we examined the frequency and character of extended conversations around mathematical ideas in these lessons. We focused on these extended conversations, which we term *extended discourse episodes* because they represent moments in the classroom dialogue in which a student has already provided a correct answer and is still required to discuss, explain, or justify that answer. In other words, extended discourse is the forum for the kinds of "higher level" talk prescribed by mathematics educators (Kazemi & Stipek, 2001). In this article, we compare how often a sample of teachers in the United States and China engaged their students in extended discourse and what the talk during extended discourse looked like. Using this cross-cultural lens, we hope to reveal ways that teachers can use extended discourse to promote higher level thinking in mathematics.

What is Extended Discourse?

In extended discourse, a student's answer to a question serves as the beginning to a larger discussion about the mathematical algorithms, rules, and reasoning needed to find that answer. An example will help explain the practice.

Teacher:	What is the answer?
Student:	5/8 equals 40/64.
Teacher:	How did you do it?
Student:	5 multiplied by 8 equals 40, so 8 should be multiplied the same number of times, and 8 times 8 equals 64.
Teacher:	How can you put it more simply? The numerator and the denominator both. . . . ?
Students (together):	Multiplied 8 times.

To give context to this practice, we first present two sections on how looking at extended discourse fits into the larger literature on the form and content of higher level classroom discourse in mathematics. Then, we explain why China makes a good partner for research into the issue of classroom discourse in mathematics.

The Form of Extended Discourse

Researchers of sociolinguistics in education have long examined how the form of classroom conversation limits or expands the kinds of talk that can occur in a lesson (Cazden, 2001; Gee, 1999; Mehan, 1979). In his influential work, *Learning Lessons*, Mehan (1979) argued that conversations in lessons take the form of three-part instructional sequences. These sequences include an initiation statement, a response to that statement, and an evaluation of the response (I-R-E). Usually, the teacher makes the final, evaluative statement in the form of a positive remark that indicates approval of the successful completion of the I-R-E sequence. However, the teacher is occasionally displeased with the student's response and must continue (or extend) the conversation to attain the successful completion of the I-R-E sequence.

Many researchers have argued that the I-R-E sequence potentially limits the verbal opportunities (and, in some cases, the thinking opportunities) for students. For instance, Cazden (2001) defined I-R-E as the traditional form of classroom discourse and noted that the discourse-intensive classrooms advocated by the National Council of Teachers of Mathematics and other mathematics educators are based on nontraditional forms of classroom discourse. In particular, she noted that nontraditional discourse includes such features as continuing a conversation even after a correct answer has been given, developing the classroom norm that providing explanations is as important as providing answers, and encouraging students to reference and critique each other's solution methods.

Continuing this line of thinking, Nassaji and Wells (2000) further discussed the limits of the I-R-E sequence and how it could be expanded to move classroom talk beyond the simple provision and acceptance of correct responses. To begin, Nassaji and Wells explained how the I-R-E sequence, in Mehan's (1979) view, was continued only when the response element of the sequence was incorrect. In such cases, the teacher withheld an evaluation of the answer and instead made verbal moves to elicit the correct answer from the student, thereby extending the conversation.

Noting these extended conversations, Nassaji and Wells (2000) posited that the move following the student response, which is the *evaluation* move in the I-R-E sequence, is the turning point of an instructional sequence. When evaluation of a response is withheld or is replaced with a different type of follow-up statement, the sequence carries on. For instance, even when a student response is correct, the teacher might extend the conversation by making a follow-up comment or asking for clarification rather than evaluating the response.

Nassaji and Wells (2000) went on to examine the types of follow-up moves made by teachers after student responses, and found that the type of initial question was closely related to the type of follow-up move made after the student response. However, Nassaji and Wells argued that the type of question did not control which follow-up move the teacher chose and, in fact, posited that choosing a follow-up move that extends, rather than concludes, a sequence may be the teacher's most important role in the discourse.

Moving toward a focus on discourse in mathematics teaching, Kazemi and Stipek (2001) connected discourse forms to conceptual learning in mathematics. After investigating a number of fourth- and fifth-grade classrooms, Kazemi and Stipek catalogued several social and sociomathematical norms that were present in classrooms with a "high press" toward conceptual learning. One of the sociomathematical norms they linked to conceptual learning was teachers not only having students present problem-solving strategies but also pushing students to explain and justify those strategies. To create opportunities for these sorts of discussions of problem-solving strategies, Kazemi and Stipek argued that teachers must engage students in extended exchanges about mathematics. In short, they argued that sustained instructional sequences are needed if discourse that promotes conceptual understanding is to arise, writing, "Sustained exchanges allow for but do not ensure conceptual thinking" (p. 68).

In this study, we focus exclusively on such sustained exchanges, which we term *extended discourse*. Extended discourse is essentially an I-R-E sequence that the teacher has extended typically by withholding evaluation of a correct answer and instead asking the student follow-up questions. It is important to study extended discourse because its form makes higher level conversation, or conversation that supports conceptual understanding, possible. Although these higher level exchanges do not always occur during extended discourse, this form of discourse is a most likely place for higher level conversations about mathematical ideas to occur. In this article, we examine first how frequently our sample of U.S. and Chinese teachers created this important form of classroom discourse. In the next section, we explain our basis for looking more deeply at this discourse for how the content of the extended discourse episodes may contribute to conceptual understanding.

The Content of Extended Discourse

In addition to looking at the various traditional and nontraditional forms of discourse found in classrooms, research on discourse in classrooms has also looked at the content of individual utterances and statements made by teachers and students. Traditionally these discourse studies (e.g., Perry, VanderStoep, & Yu, 1993) have taken a number of forms, from counting the number of words in teacher and student statements (e.g., Miller, Correa, Sims, Noronha, & Fang, 2005), to classifying the types of questions asked by teachers (e.g., Hiebert & Wearne, 1993), to looking at the kinds of knowledge discussed by students (e.g., Lampert, 1992).

In mathematics education in particular, researchers have looked at what students' language says about the level of understanding they have and about the level of understanding they are comfortable expressing. As we have already alluded to, discourse specialists in mathematics education have frequently examined how discourse is used to create and express conceptual, or higher level, understanding of mathematics. Although only a few studies (e.g., Hiebert & Wearne, 1993; Kazemi, 1998) have directly linked the discourse found in classrooms to increased student achievement (or, for that matter, increased conceptual understanding), researchers have provided several reasons why certain discourses might promote higher level understanding.

One position is that discourse in which students question each other and are pushed to explain their thinking parallels the discourse of expert mathematicians and thus engages students in doing authentic mathematics (Ball, 1993; Lampert, 1990, 1992). Another argument is that when students present and discuss their conjectures, their underlying thinking is revealed, which helps the students to clarify their ideas and perhaps see mathematical connections they might otherwise have missed (Brown, Stein, & Forman, 1996; Sfard, 2001a, 2001b; Whitenack & Yackel, 2002).

Student discussion of mathematical ideas, coupled with the use of situation-specific imagery, is also linked to "mathematizing," or a movement from context-specific problems to abstract mathematical terms and ideas (McClain & Cobb, 1998). In addition, an emphasis on discourse helps create a community in which students learn to value learning from other students (Hiebert et al., 1996) and develop a shared language about what constitutes legitimate mathematical arguments (O'Conner, 1998). Finally, some researchers interested in equity issues suggest that a higher level discourse more befits the cultural communication styles of certain minority groups (Berry, 2003; Ladson-Billings, 1997).

Although respecting the diverse types of research that could be performed on the data presented here, we chose to locate our study of the content of extended discourse on the level of questions asked by teachers during extended discourse and the level of answers given by students. By *level*, we refer to a taxonomy of mathematical knowledge that verges from computational to procedural to conceptual (e.g., Hiebert & Lefevre, 1986; Rittle-Johnson & Siegler, 1998; Smith, Desimone, & Ueno, 2005; Star, 2005). Many mathematics educators (e.g., Hatano, 1988; Kilpatrick, Swafford, & Findell, 2001) agree that mathematics students need to learn both procedural and conceptual knowledge in order to effectively, efficiently, and flexibly solve mathematical problems. However, these researchers also argue that conceptual knowledge is the higher level knowledge in the sense that strong conceptual knowledge leads to the ability to use procedures correctly and appropriately, whereas procedural knowledge alone does not ensure correct use of those procedures (nor does it ensure understanding of concepts on which the procedures are based).

In spite of, or perhaps because of, the inevitable integration of conceptual and procedural knowledge in students with advanced understanding of mathematics, a large amount of debate persists in the mathematics community over what exactly constitutes each of the various levels of mathematical knowledge. Hiebert and Lefevre (1986) made a distinction of complexity between conceptual and procedural knowledge. They defined conceptual knowledge as knowledge rich in relationships, whereas procedural knowledge was seen as knowledge of rules and procedures for solving mathematical problems.

However, as Star (2005) pointed out, these definitions essentially separate conceptual and procedural knowledge on the basis of the quality of knowledge rather than the type of knowledge. In Hiebert and Lefevre's (1986) definition, conceptual knowledge is inherently deep, whereas procedural knowledge is inherently superficial. Star argued that procedural knowledge can be deep, as indicated by flexible use of procedures, whereas conceptual knowledge can be superficial. However, Star did not acknowledge the role that a student's conceptual knowledge may indeed play in making so-called procedural knowledge deep and flexible.

Other definitions have proven useful for this current study. In particular, Rittle-Johnson and Siegler (1998) defined conceptual knowledge as "understanding of the principles that govern the domain and of the interrelations between pieces of knowledge in a domain" (p. 77). In contrast, procedural knowledge was simply defined as "action sequences for solving problems" (p. 77). In addition, in a study of mathematics teachers' goals for learning, Smith et al. (2005) placed reasoning, estimation, and conjecture under the heading of "conceptual learning goals" and placed memorization and computation under the heading of "procedural learning goals" (p. 77).

In this study, we drew on these definitions to create our codes for marking the content of student and teacher statements, but we also acknowledge the limitations of these definitions and our own codes. We will explicate the various codes we used to examine the content of the extended discourse episodes later in this article. However, it is worth noting here that we examined the content of extended discourse by distinguishing between several levels of mathematical knowledge expressed by students and teachers in their statements, which we labeled as *computation*, *procedures*, *reasoning*, and *rule recall*.

Because it is impossible to see the quality or depth of students' understanding directly through their discourse, we used discourse instead to look at the level of abstraction expressed by the students. We distinguished between procedures and what we call *reasoning* by examining the abstraction of the knowledge presented by students. Our definition of procedural statements is similar to Rittle-Johnson and Siegler's (1998) definition in that procedural statements refer to applying action sequences for solving particular problems. In contrast, reasoning is defined as conversation about mathematical ideas on a relatively abstract level. This includes discussion of why a general principle or procedure is appropriate for a particular problem, what patterns or relationships exist between numbers, or why a mathematical rule contains certain elements. Again, this definition of reasoning is similar to Rittle-Johnson and Siegler's definition of conceptual knowledge. However, we acknowledge that this definition is not similar to many definitions of conceptual knowledge that emphasize complex webs of knowledge and defer to these differences by calling our abstraction-based definition *reasoning*.

We should note that, in our more qualitative analysis of extended discourse episodes, we connected student and teacher statements with similarly coded content to some of the broader literature already cited on using discourse to construct different types of mathematical knowledge. In particular, we analyzed how rule statements could be expressed informally or formally and how that contributed to what was viewed as a shared mathematical term in the classroom (e.g., O'Conner, 1998). We also looked at how teachers used discussion of procedures and concepts to move students from thinking about specific problems to thinking about general mathematical rules and how such moves relate to mathematizing (McClain & Cobb, 1998). Finally, we also examined how students' reasoning statements could be used to construct links between procedural and more truly conceptual knowledge (e.g., Hiebert & Lefevre, 1986; Star, 2005).

In sum, with this article we plan to look at both the frequency of extended discourse in lessons from two countries and the content, or level of discussion, found within that extended discourse. By looking at the frequency of extended discourse, we were able to see how often teachers in our U.S. and Chinese samples sought to push students to express their thinking beyond providing correct answers. By looking at the content of that extended discourse, we could see whether the use of this form of discourse could be linked to conversation that promotes conceptual, or higher level, mathematical thinking.

Why Study China?

There are several major structural differences between Chinese and U.S. schools that impact mathematics education, such as the influence of a national curriculum on Chinese teaching and the fact that Chinese elementary mathematics teachers, including the ones in this study, teach only one subject (Stevenson & Stigler, 1992; Wong, Han, & Lee, 2004; Yao, 1992). Despite these differences, we believe there are three major reasons for using Chinese classrooms as a venue for understanding the practice of extended discourse.

First, previous research suggests that urban Chinese mathematics teachers have a deeper and more conceptually connected understanding of mathematics than do U.S. teachers (Ma, 1999), which may lead to higher level teaching practices (Ball, 1996). This knowledge of mathematical tasks, along with knowledge of student thinking, is critical for effective teaching (An, Kulm, & Wu, 2004; Fennema et al., 1996; Fennema, Franke, Carpenter, & Carey, 1993; Hiebert et al., 1996). Thus, looking at extended discourse in China should give us insight about a discourse practice used by teachers to uncover and extend student understanding.

Second, Chinese educators promote high student achievement despite some of the same structural factors that are often blamed for failure in the United States, such as funding issues and diversity of student ability. Chinese education is given less funding (in proportion to the country's gross national product) than it is in the United States, and this funding is perhaps even more inequitably divided than in the United States (Hannum & Park, 2002; Stevenson & Stigler, 1992). Also, despite differing student abilities, Chinese teachers do not track or group students in the elementary schools. Sometimes astonishing to U.S. observers is the fact that Chinese elementary schoolteachers ignore ability differences and do not segregate students or demand less from them in terms of achievement (Stevenson & Stigler, 1992). A look at the discourse in Chinese classrooms may reveal something about how the Chinese teachers manage to reach out to all these students.

The third and final reason for looking at Chinese education is the existing international comparisons of achievement in mathematics that have long shown that Asian schools are top international performers in general and outperform U.S. schools, in particular, on many mathematical tasks (Fan & Zhu, 2004; Stevenson, Chen, & Lee, 1993; Stigler & Hiebert, 1997; Stigler & Perry, 1988). Looking directly at the United States and China, Stevenson et al. (1990) illustrated that, in a comparison of classes of students in Chicago and Beijing, the lowest scoring Chinese classes received better scores than did the highest scoring U.S. classes. These differences were pervasive on tests of various mathematical concepts (although it should be noted that other studies, e.g., Cai, 2001; Cai & Silver, 1995, have found that Chinese and U.S. students perform similarly on process-open or complex word problems). Although there may be no definitive link between discourse and this achievement, looking at how high-achieving students engage in extended discourse allowed us to see the possibilities for such discourse.

Chinese Mathematics Education and Discourse

An observer of Chinese elementary mathematics lessons would perhaps first be struck by the emphasis on whole-class instruction. As Stevenson and Stigler (1992) reported, Chinese teachers are the leaders of their classes 90% of the time (as compared with 47% of the time in the United States). However, despite the pervasive belief that Chinese education is traditional and teacher centered, this whole-class time is not spent in lecture. Chinese teachers engage their students in discussion of a small number of difficult mathematical problems (Stevenson & Stigler, 1992). In doing so, Chinese teachers frequently move between having students work out problems in heterogeneous ability groups and having students discuss their solution methods during whole-class instruction. As Huang and Leung (2004) pointed out, Chinese mathematics teaching may be teacher led or traditional in the sense that it is dominated by whole-class teaching, but it could also be characterized as student-centered in terms of the mathematical conversations held between teacher and students.

This idea that classroom discourse may be a more accurate indicator of the true nature of teaching in a culture than are other factors (such as the amount of time spent in whole-class instruction) was the impetus for this research and other research on discourse in Chinese mathematics classrooms. Comparisons of the discourse of Chinese and U.S. mathematics teachers have been the focus of several recent studies (Miller et al., 2005; Perry, 2000; Perry et al., 1993). For example, Perry et al. (1993) used observational analysis to examine different kinds of question-asking activities by teachers in first- and fifth-grade classrooms in Japan, Taiwan, and the United States. They found that U.S. teachers asked for computing in context, problem-solving strategies, and conceptual knowledge less frequently than did their Asian counterparts. In addition, qualitative analyses of each question type supported the notion that U.S. teachers asked these kinds of questions in less successful ways. For instance, U.S. teachers provided students with computation problems in which the contexts were often irrelevant and confusing to the students. Perry et al. argued that the use of higher level questioning (defined as questioning that probes conceptual knowledge), as seen in both Asian countries, could lead students to develop a deeper and more conceptual understanding of mathematics.

In a more qualitative study that focused on just one type of discourse—teacher explanations—Perry (2000) once again found significant differences across countries. In terms of amount of explanations, Perry discovered that Japanese and Chinese children in Taiwan were exposed to more explanations per lesson than were U.S. children. The content and quality of these explanations also varied. The explanations in Asian countries were more substantive and less varied in topic than were the U.S. explanations. Perry proposed that these differences in explanation frequency and quality are important because students not only learn more about mathematics through explanations but also come to understand the importance of hearing and giving explanations in general.

Finally, Miller et al. (2005) reported that there was more mathematics talk in Chinese than in American elementary mathematics classes and that the division of labor between teachers and students was quite different. U.S. teachers produced a much greater proportion of mathematical explanations and statements than did their students, whereas the opposite was the case for Chinese classrooms, with students producing most of the mathematical statements and explanations. These findings relate quite directly to this current study. In particular, although Miller et al. reported that Chinese students talk more than do their U.S. counterparts, they did not look at the content and development of the mathematical discourse. Therefore, we felt a more in-depth analysis of the content of extended discourse in this study was worthwhile because it would control for the form of classroom discourse and would decipher what this mathematical talk entailed.

Method

Data Source

The participating schools for this study included eight schools from Beijing, China, and six schools from the area around a midsize university town in the Midwest of the United States. To study the extended discourse in mathematics lessons in the schools, we videotaped single lessons. When completed, the data set included 17 lessons in Chinese mathematics classrooms and 14 lessons in U.S. mathematics classrooms. By design, each teacher presented a lesson about equivalent or adding fractions. We asked teachers to inform us when they would be teaching an equivalent or adding fractions lesson, and we then made arrangements to capture these lessons on video. We chose equivalent fractions because of the relatively demanding nature of this topic and its centrality to the curriculum of late elementary school in both countries.

In the U.S. schools, 12 different teachers (11 female, 1 male) participated in the study. These teachers were responsible for teaching a variety of subjects, including mathematics. Most (11) of the lessons were in fourth-grade classrooms, and 3 were in fifth-grade classrooms. All of the lessons were videotaped in the spring semester. The average class size for the U.S. classes was 22 students.

In the Chinese schools, we observed 15 different teachers (13 female, 2 male). These teachers were responsible for teaching only mathematics. The classrooms in China were all fifth grade. The lessons were videotaped in the spring semester of the students' fifth-grade year. The average class size for the Chinese classes was 55 students.

As indicated in the previous paragraphs, the Chinese students studied this material a year later than did most of the U.S. students. Although this age difference might initially appear problematic, we believe that our use of different grade levels is justifiable for several reasons. First, our major emphasis was to uncover and illustrate discourse practices, not to judge one practice as better than another. For this to be accomplished, it was essential that all the examined discourse be centered on the same topic to ensure that a similar depth of discussion was possible in all the lessons we studied. To match the content being taught, we found it necessary to videotape fifth graders in China and fourth graders in the United States. In other words, the matching of discussion topics was more important than the matching of discussant ages (Perry, 2000).

In a study that also discussed discourse in fourth- and fifth-grade classrooms, Kazemi and Stipek (2001) made a similar argument regarding examining different grade levels. Although they acknowledged their initial concern with making comparisons in student and teacher talk across grade levels, they said that their emphasis on analyzing mathematical conversations, as opposed to evaluating student mathematical knowledge, alleviated such concerns. They wrote, "Although it may be reasonable to expect fifth-grade students to know more than fourth-grade students, it is also reasonable to expect that both fourth and fifth graders could engage in conceptual conversations about mathematics" (p. 62). We believe this argument is applicable to our own work, for although the older Chinese students in our study may arguably have had more mathematical knowledge than did their younger U.S. counterparts, this knowledge should not have precluded any of the students and teachers from engaging in conversations about all kinds of mathematical knowledge, from computational to conceptual. In fact, the current movement toward discourse-oriented teaching in mathematics education relies on the fact that no matter what their age, all elementary school students can and should engage in conversations that probe conceptual knowledge, and we believe this study contributes to that movement by looking at conversations, not achievement.

Second, to further scrutinize our decision to use different grade levels, we made use of the fact that three of our U.S. classrooms were fifth-grade classes. To ensure that the differences we discovered were not a product of the age of the students, we looked for differences between the fourth- and fifth-grade U.S. classes in the use and content of extended discourse. There were no apparent differences between the fourth- and fifth-grade U.S. classes on any of our measures, so we were confident that differences between U.S. and Chinese classrooms were not likely a result of the difference in the average age of the students.

In the sampling of schools, we procured as diverse a sample as we could obtain among the area schools that agreed to be videotaped for our study. In the U.S. data set, there were significant demographic differences among the schools. The schools in the university town itself tended to have a great deal of diversity and a low overall socioeconomic status. Some rural, low-socioeconomic-status schools outside of the university town were used for the study, as were some high-socioeconomic-status schools in cities surrounding the university town. Although the schools studied in China were all located in the urban Beijing area, most of the students in the Chinese classrooms were from either middle-class or working-class backgrounds.

In addition, we made an effort to account for the prestige of the schools in both countries. There was some variability in the perceived prestige of the Chinese schools. Two of the schools were viewed as top-level schools, whereas the others were viewed as middle-to-high-level schools. The differences in prestige were virtually the same in the U.S. schools, with most being viewed as middle to high range, whereas one was viewed as top range. Despite our best efforts to find a range of prestige levels and student demographics in the schools used for this study, it is important to note that these schools are at best representative of the regions from which they were procured and cannot be said to be representative of China and the United States as a whole. As we discuss later, the point of this article is thus not to make cross-national comparisons but to use data from schools in both the United States and China to explore and consider a potentially valuable educational practice.

Finally, in all the classrooms, the teachers and students did not know what aspects of the lessons were going to be analyzed. The teachers, students, and parents in both countries were told that the classes were being videotaped for researchers to examine how students develop an understanding of mathematics through their class experience and how videos of classroom instruction can be used to train teachers. This research thus represents a portion of a larger project aimed at understanding differences between U.S. and Chinese mathematics teaching practices, with the purpose of enabling teachers to improve their practices.

Analytic Plan

Our method for examining the data was fourfold. First, we used two quantitative measures, one to examine how frequently extended discourse occurred in the lessons in our sample and a second to find differences in student and teacher statements within these episodes of extended discourse. Next, we used a process called dynamic time warping to look at the content of the extended discourse, not in terms of the quantity of each type of statement but in terms of how the statements were sequenced. Finally, we looked more qualitatively at the data by examining examples of extended discourse episodes and discussing subtle differences within them.

The justification for this fourfold method is as follows. The analytic approach that we followed here was a mixed-methods approach. As Greene and Caracelli (1997) have suggested, we followed the dialectical position, which is “shaped by both interpretivist and postpositivist paradigms” (p. 10). We followed this approach because we felt it was necessary that both generality and particularity were purposefully addressed (Rocco et al., 2003). The necessity for addressing issues of generality came from the responsibility to report, for example, the number of occurrences of extended discourse episodes and any significant differences in prevalence across the lessons from the two locales. The necessity for addressing issues of particularity came from the responsibility to report what these occurrences looked like and how they transpired in distinct classrooms within the two countries we sampled for our investigation. Furthermore, we note that, especially when addressing issues cross-culturally, it is critical to provide direct examples so as to expose our own cultural perspectives (Moghaddam, Walker, & Harré, 2003).

To deal with examining issues of generality, we first devised a coding system to categorize different types of extended discourse. By doing so, we could determine the prevalence of this discourse practice and also compare the prevalence across countries. Here, we used inferential statistics, including analysis of variance. We also computed effect sizes so that the importance of the findings could be judged.

We relied on analyses of variance because they are robust with respect to the condition of normality, and some of our data did not follow neat, normal distributions. In particular, we found that some of the U.S. classrooms relied on some types of statements considerably more than did other classrooms, whereas we were less likely to observe uneven distributions in the Chinese classrooms. Also, because the total lesson length varied, especially within the U.S. lessons, we conducted analyses on the proportion of time devoted to extended discourse. Analyses of variance on proportional data require transformations. Thus, analyses of the amount of time devoted to extended discourse in both countries were conducted on arcsine-transformed proportions of time spent on extended discourse out of the total time in each lesson. The remaining analyses were performed on original data (as opposed to transformed data).

In our analyses, we began by examining general characteristics of extended discourse episodes in both locations. We examined the frequency, length, and other general quantifiable features of extended discourse episodes. We next took a look at the content of extended discourse episodes, and we did so separately for both teachers and students. Finally, we examined the structure of these episodes—what came first, second, and so forth—to document how these episodes proceeded. We relied on a statistical technique, dynamic time warping (Kruskal & Liberman, 1983), to help us identify patterns in the data across episodes of different lengths. We explain this technique in more detail when we discuss the results from this analysis.

To deal with examining issues of particularity, we took a close look at examples of extended discourse. We discuss some of the distinctive features that we noticed in these episodes and offer examples of these in Chinese and U.S. lessons. We focused on three behaviors that transpired differently in the Chinese and U.S. lessons we observed. In particular, we present examples and discuss generalized versus context-specific knowledge, components of reasoning, and rule formality. In these analyses, we seek to look beyond the codes, which we used to quantify the data, and examine these episodes more qualitatively, attempting to make sense of how different types of extended discourse episodes might differentially influence the development of students' mathematical knowledge.

In the next two sections, we continue to explicate the coding systems we used for measuring the amount of extended discourse in the data and the content of the statements made by students and teachers during this extended discourse. The other methods used are discussed further in the *Results* section.

Identifying Episodes of Extended Discourse

The first step in analyzing the data was finding and marking episodes of extended discourse, or periods in which one mathematical question was discussed in depth by the students and teacher. For ease of reading, please note that we term any exchanges involving extended discourse *episodes*, which should not be confused with the *sequences* we discuss later in our dynamic time-warping analysis. Extended discourse episodes were defined by several criteria. First, an episode must have begun with a teacher-initiated question or with a student-initiated question that was accepted by the class for continued discussion. Then, the episode must have contained at least two substantive (more than one-word) student responses to the initial question.

The ends of the episodes were marked differently, depending on the type of initial question asked. When the question was about manipulating numbers, the episode ended when the numbers involved in the question changed or basically when the numerical question was altered. When the question was about mathematical ideas, such as a question about mathematical terms or rules, the episode ended when the teacher moved to a numerical question or to an unrelated question about mathematical ideas.

In general, extended discourse episodes began only after a student gave a complete answer. If the teacher had to prompt a student to get him or her to report an answer, this prompting was not included as part of the episode. The criteria for extended discourse episodes were designed to capture extended discussion about the answer to one question, not all interactions between students and teachers. More specific examples of extended discourse episodes are presented in the *Results* section.

As an aside, it is important to note that the only kind of discourse examined in this study is verbal, spoken discourse. Although we respect the contributions of nonverbal communication (e.g. gestures, drawings, written assignments) to discourse and the creation of mathematical communities in classrooms, the study of nonverbal language is beyond the scope of this study (see Flevares & Perry, 2001, for an example of how nonspoken communication is used in elementary mathematics classes). Although we do not examine this here, we acknowledge that it is likely that the use of nonverbal forms of communication influences the kinds of verbal discourse, such as extended discourse, that gets expressed in classrooms (Ball, 1993).

Coding Content of the Statements in Extended Discourse Episodes

After locating the extended discourse episodes within each lesson, we looked closely at what kinds of discourse occurred in these episodes. We then examined to what extent this discourse differed between China and the United States. To do so, we coded individual student and teacher statements within each episode for their intended functions in the mathematical dialogue. The statement, rather than the utterance, was the unit of analysis for this study (Bakhtin, Holquist, & Emerson, 1986; Stieglitz & Oehlkers, 1989). We initially considered coding each teacher or student turn, or what Bakhtin et al. called an "utterance," as a whole. However, we decided instead to assign a code to each individual statement within an utterance. We did so because we wanted to capture (albeit somewhat crudely) what the teacher or student emphasized within their utterance. For instance, if a teacher made three statements explaining a mathematical idea and one statement praising a student, we wanted to capture the relative emphasis on explanation in that utterance by coding each statement, instead of simply labeling the entire utterance as explanation and praise.

Table 1 shows the nine codes for teacher statements and six codes for student statements. Each of the six student codes is matched by a corresponding teacher request code. The three additional teacher codes had no corresponding student codes. All coding was done on a statement-by-statement basis, although both the original coder and the reliability coder were aware of whether each statement was spoken by the teacher or by the student. For a more in-depth explanation of each of the codes, please see the Appendix.

Reliability

We computed reliability at the following two stages: identifying extended discourse episodes and identifying the content of these episodes. To begin, one primary coder identified all the episodes of

Table 1
Content Codes for Function of Teacher and Student Statements in Extended Discourse

Teacher discourse codes	Student discourse codes
1. Request for computation	1. Computation
2. Request for procedure or method	2. Procedure or method
3. Request for reasoning	3. Reasoning
4. Request for rule or term recall	4. Rule or term recall
5. Check for student understanding and/or agreement	5. Indication of understanding and/or agreement
6. Request for short answer	6. Short answer
7. Teacher explanation	
8. Restating student answer	
9. Praise	

extended discourse. We achieved reliability by having another independent coder examine 25% of these lessons. In this case, the independent coder examined 4 of the 17 Chinese lessons and 3 of the 14 U.S. lessons. Simple agreement between coders for locating episode beginnings and endings was .86, with a Cohen's Kappa of .71, which is considered substantial reliability (Cohen, 1960; Landis & Koch, 1977).

To determine reliability for the content of the statements in the extended discourse episodes, one coder coded all 31 lessons. As with identifying episodes, we achieved reliability in content coding by having another independent coder code 25% of the lessons. Simple agreement was .81, with a Cohen's Kappa of .80, which is considered substantial reliability (Cohen, 1960; Landis & Koch, 1977).

Results

We report the features of extended discourse at several levels. First, we report basic features of extended discourse, including the frequency and length of these episodes. Next, we report analyses of the content of the statements made in these episodes based on the codes developed explicitly for this investigation. We report the analyses of the content separately for the teachers and students. Next, we examine the structure of these episodes by looking at the sequences of statements from each episode in the Chinese and in the U.S. lessons. Finally, we present examples of extended discourse to give a clear sense of what these episodes looked like and to discuss the different ways extended discourse could be used to promote mathematical knowledge.

As a caveat, we remind the reader that the statistics presented in this section are exploratory in the sense that they are presented as the basis for discussion of a particular discourse structure. These statistics confirm and enhance our initial observations of this discourse structure and thus are used to describe extended discourse in this article. But by no means can these statistical results be generalized to Chinese and U.S. lessons or teachers in general. The teachers and lessons in this sample were not randomly selected and thus cannot be said to be representative of the nations as a whole. However, we present statistical analysis here to spur discussion of the implications of the differences we found within our small cross-cultural sample.

Frequency, Length, and General Description of Extended Discourse Episodes

In all, we found 400 extended discourse episodes, with 279 occurring in the Chinese lessons and 121 in the U.S. lessons. To obtain a general picture of the amount of extended discourse used in the lessons from both countries, we measured and compared several features of classroom practice of extended discourse. In particular, we examined the amount of time spent in extended discourse per lesson, the number of episodes of extended discourse per lesson, the length of each episode of extended discourse, and the number of different students who spoke in each episode of extended discourse.

Please note that the degrees of freedom used for the statistical analysis varied in this section. The reason for this was simple but perhaps not self-evident. The first two sets of results looked at time spent in extended discourse and number of extended discourse episodes *per lesson*. Because these results were averaged for each of the 31 lessons in our sample, the degrees of freedom were 1 (country) and 29 (lessons). The other two measures were the length of and number of students participating in each extended discourse *episode*. As we found 400 such episodes in this data set, the degrees of freedom when we were analyzing episodes were 1 and 398.

Time spent in extended discourse. Extended discourse appeared in a significantly higher percentage of the Chinese than the U.S. lessons. The Chinese classrooms we observed spent about 37% of the lesson time in extended discourse, whereas the U.S. classes spent only 21% of time in extended discourse, $F(1, 29) = 14.47$, $p < .001$, $\eta^2 = .335$.

Number of extended discourse episodes. We found a significant difference in the average number of extended discourse episodes per lesson in each country, $F(1, 29) = 17.29$, $p < .001$, $\eta^2 = .373$. On average, each Chinese lesson had 16.4 extended discourse episodes, and each U.S. lesson had 8.6.

Length of extended discourse episodes. We found no significant differences between countries in the length of extended discourse episodes, $F(1, 398) = 3.19$, $p > .05$. Both countries' mean episode lengths were around 1 min: the Chinese lessons averaged 58.86 s per episode, and the United States lessons averaged 67.37 s per episode.

Number of different students contributing to an episode of extended discourse. We found no significant differences between countries in the average number of different students participating in each episode, $F(1, 398) = 2.39$, $p > .05$. The number of different students participating in each episode of extended discourse was 2.37 in the Chinese lessons and 2.62 in the U.S. lessons.

In sum, we observed more extended discourse in the Chinese than in the U.S. lessons, but the length of the episodes and the number of students participating in the episodes did not differ.

Content of Extended Discourse Episodes

Analysis of the content of extended discourse episodes was done separately for teacher statements and student statements. For analysis of the content, we began by analyzing the proportion of each type of statement. We argue that analyses of the proportional data were necessary because the Chinese lessons had more overall extended discourse; thus, an analysis of the simple number of each type of statement would be biased in favor of the Chinese exhibiting more of each type of statement. Given the issues with using analysis of variance on a restricted range, which results from proportional data, all analyses were conducted on arcsine-transformed data. Please note that because all the data in this section were averaged for each of the 31 lessons, the degrees of freedom were 1 and 29.

Teacher statement codes. We displayed the mean proportion of each type of teacher statement by country in Table 2, as well as F values and effect sizes (η^2). We include effect sizes here, measuring the proportion of variance accounted for by being observed in our China sample versus in the U.S. sample. In general, effect sizes greater than .3 suggest that the difference between the countries was large (Cohen, 1969; Glass, McGaw, & Smith, 1981).

Table 2
Mean Proportion of Teachers' Use of Different Types of Statements in Chinese and U. S. Fractions Lessons

Statement type	China *M* (n = 17)	U. S. *M* (n = 14)	$F(1, 29)$	Effect size (η^2)
Request for computation	.11	.29	28.86**	.499
Request for procedure	.13	.08	7.76*	.211
Request for reasoning	.18	.09	15.27**	.345
Request for rule/term	.08	.01	53.75**	.650
Check for understanding	.11	.05	15.70**	.351
Request for short answer	.03	.03	1.76	.057
Praise	.08	.06	3.29	.102
Restatement/revoicing	.13	.16	0.79	.027
Explanation	.14	.24	17.62**	.378

* $p < .01$. ** $p < .001$.

As can be seen in Table 2, the Chinese teacher statements were significantly ($p < .05$) more frequently coded as requests for procedures, requests for reasoning, requests for rules or terms, and checks for understanding or agreement than were the U.S. teacher statements. The U.S. teacher statements were significantly more often requests for computation or explanations than were the Chinese teacher statements. There were no significant differences by country for request for simple answer, praise, or restatement/revoicing.

Student statement codes. Not surprisingly, categories of student discourse that corresponded to teacher discourse followed similar patterns. In Table 3 we display the proportion of each type of student statement by country. The Chinese students uttered a significantly greater proportion of procedure, reasoning, rule/term recall, and indication of understanding or agreement statements than did the U.S. students. In turn, the U.S. students made computation statements significantly more often than did the Chinese students. We found no significant difference between the lessons in each country in the proportion of short answers.

Examining the Structure of Extended Discourse Episodes by Looking at Statement Sequences

Although the analysis of student and teacher statements in each episode of extended discourse illustrates the kinds of questions and answers most frequently uttered during extended discourse, this analysis gave us only some idea of what actually went on within each episode. We knew how many kinds of each statement occurred in extended discourse in general, but we did not know much about the sequences of these statements. For instance, in two episodes of extended discourse, we might have counted eight overall statements, including five computation statements and three procedural statements. These counts were already included in our data, but they obscure the sequence these statements formed. One episode could contain the five computation statements in a row, whereas the other episode contained the three procedural statements. Or perhaps one episode could contain two computation statements followed by one procedural statement, whereas the other episode contained two procedural statements followed by three computation statements. A clearer picture was needed of the sequences of the statements in each extended discourse episode and when and how these different sequences were qualitatively different.

To get this general picture, we used a statistical technique, only recently adapted for exploring classroom discourse, called dynamic time warping (Kruskal & Liberman, 1983; Kumar, 2004). This approach provided us with a way of quantifying similarity between extended discourse episodes by looking at the "distance" between the sequences of statements in these episodes on a two-dimensional scale.

Dynamic time warping was developed for speech recognition systems as a way of dealing with the problem that different

Table 3
Mean Proportion of Students' Use of Different Types of Statements in Chinese and U. S. Fractions Lessons

Statement type	China *M* (n = 17)	U. S. *M* (n = 14)	$F(1, 29)$	Effect size (η^2)
Computation	.21	.55	36.97**	.560
Procedure or method	.29	.21	4.65*	.138
Reasoning	.20	.12	7.47*	.205
Rule/term recall	.16	.03	42.61**	.595
Indication of understanding	.09	.05	6.85*	.191
Short answer	.05	.05	1.10	.037

* $p < .05$. ** $p < .001$.

speakers speak at different rates (Myers & Rabiner, 1981). It provides a measure of how dissimilar or distant two sequences of events are. To sensibly examine the data, we introduced some limits on how sequences could be viewed. Following Kumar, we treated Sequence AABBCC as identical to Sequence ABCCCC. This algorithm compresses sequences while maintaining alignment between the sequences (in this example, both sequences were identical in the sense that they followed an ABC pattern). Although treating AABBCC as identical to ABCCCC was, at some level, arbitrary, we adopted this decision rule for two reasons. First, we needed some principled way to combine sequences, to reduce the number of different sequences if we were to efficiently describe the world of sequences we encountered. When we considered alternatives, treating, for example, repeating computations but counting every time the discourse switched to giving reasons seemed a reasonable compromise that generally represented the data while leaving the flow of the discourse intact. Following this decision rule, the distance between two identical sequences (including sequences made to be identical, by our decision rule) would be 0, and the more dissimilar two sequences were, the greater the distance between them. Essentially, it ignored the duration of events of the same time, while looking at switches among different things (be they phonemes or classroom activities of a particular kind).

In the present investigation, we looked at how the student statements progressed in each episode of extended discourse and how similar or dissimilar these sequences of statements were to each other. Using only the most prevalent student codes (computation, procedure, rule recall and reasoning), we characterized each extended discourse episode by the sequence of student responses. To make the pattern analysis a bit simpler and thus easier to interpret, we combined the reasoning and rule recall codes because both types of statements represented a move to more abstract, generalized thinking about mathematical ideas. As we noted in the introduction, our codes generally measured the level of abstraction expressed by the student and not the depth or complexity of the student's knowledge. Because of that distinction, we collapsed the two abstract codes, reasoning and rule recall, in this analysis. Thus, for instance, if students said, in order, a computation statement, two process statements, and either a reasoning or a rule recall statement, the episode exhibited a computation–process–process–reasoning sequence.

After marking each episode for its distinct sequence of statements, we used the dynamic time-warping technique to find the distance from every episode to every other episode, resulting in a 400×400 distance matrix. As the dynamic time-warping distance is a measure of dissimilarity between two sequences, all the diagonal elements of this distance matrix were 0 (each episode was identical to itself). The matrix was also symmetric about the diagonal (because, e.g., the distance between Episode D and Episode E was the same as the distance between Episode E and Episode D).

To visualize the distance matrix, we used multidimensional scaling to reduce the distance data to a two-dimensional plot (see Figure 1). Note that the x- and y- axes of this plot are not informative because multidimensional scaling preserves only the distances in the input distance matrix. Thus, two episodes that contain identical or very similar sequences are placed closer together on the plot, whereas two dissimilar episodes have a greater dynamic time-warping distance between them and are further apart on the plot. For example, episodes that began with computation

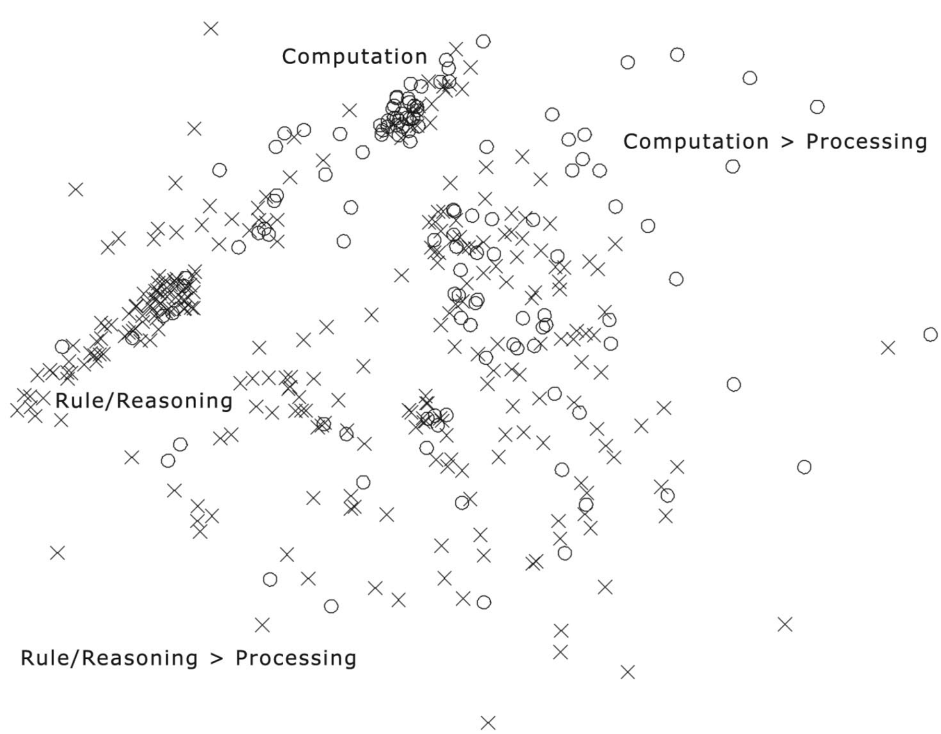

Figure 1. Dynamic time-warping plot of the extended discourse episodes in each country.

statements were fairly similar and thus are close together. Likewise, episodes that contained only computation statements are closer to each other than episodes that contained both computation and reasoning statements.

In Figure 1, each Chinese lesson is represented by an X and each U.S. episode is represented by an O on the plot. Although we have not drawn lines in this plot to distinguish clusters or patterns, we note that we might have easily included two diagonals, one proceeding from the southwest corner to the northeast corner and one proceeding from the northwest corner to the southeast corner. These diagonals separate the data into four clusters: rule/reasoning statements (west), computation statements (north), computation statements followed by processing statements (east), and rule statements followed by processing statements (south). As illustrated in Figure 1, the Chinese classrooms clustered most heavily in the quadrant that included rule and reasoning statements, and the U.S. classrooms were essentially absent from this quadrant. The U.S. classrooms clustered mainly in the quadrant that focused on computation statements without process statements. Process statements seemed to be secondary in the lessons in both countries; they appeared to be used in conjunction with rules or computation and rarely alone.

What the data in Figure 1 suggest is that the Chinese and U.S. teachers in this study requested different types of knowledge from their students: The Chinese students were prompted to answer with rule or reasoning statements, whereas the U.S. students were prompted to answer with computation statements.

Examples of Extended Discourse

Although the reliance on rule and reasoning statements was found predominantly in the Chinese lessons and the reliance on computation statements was found predominantly in the U.S. lessons, classrooms in both countries used computation, procedure, rule, and reasoning statements. However, when one looks at examples of these statements in action, it becomes clear that even when the Chinese and U.S. teachers were using and eliciting the same types of statements, the discourse in each country still had a unique character. In fact, it seems that the teachers in this study used similar statements in pedagogically different ways. In this section, we provide illustrations of extended discourse in practice and discuss how statements made in extended discourse were used to construct different types of mathematical knowledge.

Generalized versus context-specific knowledge. Close examination of the use of similarly coded statements in both countries revealed that the Chinese teachers used discourse around specific problems to help students develop general mathematical knowledge. This is an interesting practice because the use of language to move context-specific problems to more abstract rules and mathematical ideas is an oft-reported goal for reform discourse (McClain & Cobb, 1998). To frame this use of discourse as it appears in Chinese lessons, first we present an example of an episode in which a U.S. teacher asked for computation statements only. (See the computation cluster on the upper right-hand part of Figure 1 to see where this episode falls in the dynamic time warping analysis.)

Teacher: Well, what about something else that equals 1/4? Student 1? (Request for computation)

Student 1: Um. . .3/16ths? (Computation)

Teacher: 3/16ths. (Restatement) 3 times 4. . .? (Request for computation)

Student 1: That equals 12. (Computation)

Teacher: 3 times 4 would be 12. (Restatement)

Student 1: 3/12ths. (Computation)

Teacher: Can we say 3/12ths, Student 1? (Check for understanding)

Student 1: Yeah. (Indication of understanding)

Teacher: Yeah, that'll work. (Praise) That follows our pattern. (Explanation)

In this example, the student used computation statements only to provide answers and a correction to an answer. The teacher requested computation by prodding the student to the right answer by saying "3 times 4. . .?" Neither the teacher nor the student pointed out the more general conceptual reason for the student's error of saying 3/16ths, that the numerator and denominator were multiplied by different numbers (and thus multiplied by a quantity unequal to 1) or even the procedural idea that because 1 was multiplied by 3, 4 must also be multiplied by 3. Instead, the teacher simply helped the student solve this problem without assisting this student or the class with understanding how to solve similar problems. In this way, the discourse around computation in this U.S. lesson was quite context specific.

Now here is an example of a Chinese extended discourse episode in which the teacher asked for computational statements as well. (This episode, like the other in this section, also falls in the computation cluster of Figure 1.)

Teacher: Fill in numbers. Start from that side. Student 1, the first one. (Request for computation)

Student 1: 12/36 equals 9 over. . . ? (Computation)

Teacher: What? Speak louder. (Request for computation)

Student 1: Equals 3/9. (Computation)

Teacher: From 36 to 9, the denominator is divided by what? (Request for computation)

Students (together): 4. (Computation)

Teacher: To keep the fraction unchanged, the numerator should also be divided by what? (Request for computation)

Students (together): 4. (Computation)

Teacher: So that is 3/9. (Explanation)

In this example, the Chinese student, like the U.S. student, used computation statements to give an answer and to give affirmations of an answer. Similarly, after the answer was given, the teacher requested computation in a prodding manner, by asking what the denominator was divided by to get from 36 to 9. However, the most striking difference between the practices of the U.S. teacher and the Chinese teacher is illustrated in the last request for computation from the Chinese teacher: "To keep the fraction unchanged, the numerator should also be divided by what?" Although this was a computation request and would result only in a reiteration of the number used in division, the Chinese teacher reinforced procedural ideas with her question. She reminded the students that the numerator and denominator must be divided by the same number, and the answer was provided by the student in the context of this request for computation. Thus, in this example, requests for computational statements used by the Chinese teacher included reference to more generalized procedural thinking than did the U.S. teacher's requests and illustrated more of an interest in eventual mathematizing (McClain & Cobb, 1998).

Components of mathematical reasoning. We have characterized the Chinese teachers in this study as valuing rules as a building block of mathematical knowledge. Yet this belief in rules extends beyond the common requests for rule recall in the Chinese episodes. Another interesting aspect of the extended discourse in the Chinese lessons is that much of the discussion of mathematical reasoning in the Chinese lessons sought to provoke thinking about underlying mathematical rules. For instance, consider this episode from a Chinese classroom in which mathematical reasoning was elicited from the students. (This episode is located on the left-hand side of Figure 1, in the rule/reasoning cluster.)

Teacher:	This one, is it true? (Request for simple answer)
Students (together):	False. (Simple answer)
Teacher:	Why false, Student 1? (Request for reasoning)
Student 1:	Because the numerator is multiplied by 2, while the denominator is not. (Reasoning)
Teacher:	Yeah, so the fraction? (Request for reasoning)
Student 1:	Changes. (Reasoning)
Teacher:	It changes. (Restatement)

It appears that, even when requesting reasoning, this Chinese teacher was still seeking student knowledge of underlying rules. In particular, the student's response made use of a rule that we often observed in the Chinese classrooms, the *rationale of consistent quotient*, when he said the reason that the answer was false was because the numerator and denominator were not multiplied by the same number. With the teacher's prompting, the student explained that the fraction was not equivalent because the numerator and denominator were multiplied by different numbers. An important point here is that the Chinese teacher used reasoning to get the students to talk about the underlying mathematical rule (here, the rationale of consistent quotient), as did many other Chinese teachers in our study. The Chinese teachers we observed seemed to place a heavy importance on mathematical rules.

This emphasis on rules found in the Chinese lessons is quite different from the discourse around reasoning found in the U.S. classrooms. Interestingly, we found that the teachers in the U.S. lessons used reasoning to promote more of an integration of different types of computational and procedural mathematical knowledge, as illustrated in the following example. (This episode, like the previous episode in this section, also resides in the rule/reasoning cluster on the left-hand side of Figure 1).

Teacher:	You can either have 1/2 of my candy bar or you can have 2/4 of my candy bar. Which do you want? (Request for reasoning)
Student 1:	It doesn't matter. (Reasoning)
Teacher:	Why not? (Request for reasoning)
Student 1:	Because they are both the same. (Reasoning)
Teacher:	They are both the same. (Restatement) Because 1/2 is, everybody. . . ? (Request for reasoning)
Students (together):	Equivalent to 2/4. (Reasoning)
Teacher:	That's the whole idea that we're going to work with today. (Explanation)

In this example, the students were reasoning verbally about the amounts 1/2 and 2/4. This reasoning is not easily derived from a rule but is related to computation, computational procedures, and reasoning. To reason about this question correctly, the students used a procedure to compute whether 1/2 and 2/4 were equal or unequal. They then connected this to a general theory of equivalence. Thus, the U.S. teachers in this study seemed to emphasize using procedures and computations in complex ways.

The differences between these two examples may illustrate the distinction Star (2005) made between superficial and deep procedural knowledge. Although the Chinese teacher in this example sought for the child to essentially repeat and apply knowledge of a standard rule, the U.S. teacher in this example sought an integration of procedural knowledge regarding the relationship of 1/2 to 2/4 and more conceptual knowledge about equivalence. This second example from the U.S. classroom could illustrate more of an interest in using reasoning to support "deep" procedural knowledge, as opposed to the somewhat more superficial recitation of a standard rule in conjunction with a mathematical problem.

Formal versus informal rules. As discussed already, the Chinese teachers in this study placed an emphasis on mathematical rules. In addition, the Chinese teachers placed a great importance on formal language when discussing these rules. Here is an example of this formal use of rules in a discussion of the rationale of consistent quotient. (In the dynamic time-warping analysis, this episode is located on the diagonal between the computation cluster and the rule/reasoning cluster, in the northwest quadrant of Figure 1.)

Teacher:	Please fill in the equation on the blackboard. (Request for computation)
Student 1:	Eight divided by 10 equals 4 divided by 5 equals 12 divided by 15. (Computation)
Teacher:	Good. (Praise) What's your rationale? (Request for rule)
Student 1:	It's based on the rationale of consistent quotient. (Rule)
Teacher:	Can you say that in detail? (Request for rule)
Student 1:	The quotient will stay consistent if two numbers are multiplied or divided by the same number. (Rule)
Teacher:	Sit down, please. Anything else? You, please. (Request for rule)

Student 2: The quotient will stay consistent if two numbers are multiplied or divided by the same number at the same time. (Rule)

Teacher: Okay, anything else? You, please. (Request for rule)

Student 3: Except zero. (Rule)

Teacher: Good. (Praise) This is a very important condition. (Explanation) We must pay attention to it. (Explanation)

In contrast, an interesting feature of the discourse in U.S. lessons is how teachers encouraged students to discuss rules informally, or in their own words. Here is an example of a U.S. lesson in which the rationale of consistent quotient was discussed in a more informal manner. (This episode, like the previous episode in this section, is located in the northwest quadrant of Figure 1, between the computation and rule/reasoning clusters.)

Teacher: Now, guys, if we use division to get lowest terms, what could we have used to get a bigger equivalent fraction? (Request for procedure) Multiplication. (Explanation) Six times what gives you 12? (Request for computation)

Students (together): Two. (Computation)

Teacher: Two. (Restatement) Whatever I do to the denominator, I have to do . . . (Request for reasoning)

Students (together): To the numerator. (Reasoning)

Teacher: To the numerator. (Restatement) So 5 times 2 is 10. (Explanation) Your equivalent fraction was 10/12. (Explanation)

In this example, the U.S. teacher seemed interested in promoting the understanding of a rule as part of solving a problem, not as a separate, formal entity. Thus, it appears that both the Chinese and U.S. teacher were interested in knowledge of mathematical rules in the context of these computation-reasoning episodes, but the Chinese teacher asked for more formalized rules. In this sense, the teachers in these examples created very different classroom norms about the kind of language that constitutes an acceptable mathematical argument (O'Conner, 1998). Although the memorization and recitation of formal rules was required for making an argument in this example from a Chinese classroom, putting rules "in your own words" was seen as necessary for legitimate mathematical argument in the U.S. example.

It should be noted that in this discussion of generalized versus context-specific knowledge, components of reasoning, and rule formality, we are not arguing that these findings were universal to every teacher in a particular country and are certainly not drawing any conclusions about the superiority of one pedagogical approach over another. We simply sought to look beyond the codes to examine these episodes for their unique features and in doing so found that even similar extended discourse episodes can be used to construct different ideas about the kinds of mathematical knowledge that are valued in a particular classroom. In the discussion, we examine how these unique features may benefit mathematical learning.

Discussion

Our goals for this discussion are to summarize our findings about the extended discourse in the U.S. and Chinese classrooms and provide four implications of these findings. First, we summarize what we learned from the Chinese classrooms. In comparison with the U.S. classrooms, the extended discourse in Chinese classrooms was more frequent and included more discussion of procedures, rules, and reasoning. In addition, we found evidence that many of the extended discourse discussions in Chinese classrooms connected specific problems to more general mathematical rules, involved students in expressing their understanding of rules, and pushed students to express these rules formally.

In contrast, the extended discourse in the U.S. classrooms was less frequent and included a relatively heavy emphasis on computation. In addition, some U.S. teachers also appeared to engage students in computation that was specific to individual problems and did not push students to connect these problems to more abstract mathematical ideas. Other examples illustrated that some U.S. teachers asked students to reason by integrating concepts and procedures and allowed students to express mathematical rules informally.

We believe there are at least four potential implications for the major findings reported in this investigation. First, although the sample size for this study was too small to support generalizations about the nature of teaching in China and the United States, we do believe these findings complement and echo the results of other cross-national comparisons of mathematics teaching in Asia and the United States. (e.g., Miller et al., 2005; Perry, 2000; Perry et al., 1993; Stigler & Hiebert, 1997, 1999, 2004). In all of these studies, subtle differences in discourse were seen as representative of larger cultural pedagogical patterns in terms of what types of mathematical knowledge are emphasized and valued. In our study, as in the others, the Asian teachers, by their actions, seemed more interested than were their U.S. counterparts in conversations about mathematics in general and in discussion about mathematical rules and concepts more specifically. As the evidence of such differences continues to accumulate, we are more likely to consider that classroom discourse plays a significant role in larger cross-cultural achievement patterns.

Second, we believe that extended discourse, as illustrated and explicated in this article, can provide a foundation for higher level exchanges between students, as Kazemi and Stipek (2001) predicted with their discussion of sustained exchanges as necessary but not sufficient for promoting conceptual learning. The Chinese teachers in our study provided an example of using extended discussions to ask questions about rules and reasoning, while some of the teachers from both countries in our sample also indicated how extended discourse could be used to help students develop deeper procedural knowledge that is integrated with conceptual knowledge (e.g., Hatano, 1988; Star, 2005) and to move students from more context-specific to generalized thinking (e.g. McClain & Cobb, 1998). Although our data do not allow us to interpret whether extended discourse is truly necessary for conceptual learning (and we argue that in some cases it is not), we do believe that engaging students in extended discourse is an important first step toward developing a discourse-oriented pedagogy that emphasizes conceptual knowledge. Echoing the emphasis that Nassaji and Wells (2000) placed on the follow-up move in the I-R-E pattern, we believe that simply asking students a follow-up question after a correct answer may open the door to deeper investigation of the mathematical content behind a particular problem.

Third, we argue that although extended discourse may allow higher level mathematical conversations, that is certainly not its only possible function. As Cazden (2001) posited, using a particular form of discourse (such as I-R-E) does not necessarily result in a particular kind of instruction. For instance, we saw in the data from the U.S. lessons that the teachers primarily used extended discourse as a means for asking students to engage in computation. Allowing students to practice computation is also an essential aspect of mathematics teaching and could usefully be incorporated into extended discourse, but, we argue, should not be its only function. We also saw some, but not all, teachers use extended discourse to push students from context-specific thinking to abstract thinking. In this sense, we must be careful not to assume that the presence of extended discourse guarantees better mathematics instruction but should instead look to how extended discourse is used in excellent mathematics instruction and how we can encourage teachers to use extended discourse in similar, better ways.

Finally, although the extended discourse in lessons from both countries has revealed some intriguing practices, the discourse in neither country suggests that the pedagogy in these classrooms has reached the extent of the reforms proposed by the National Council of Teachers of Mathematics (1991, 2000) or the supporters of discourse-intensive teaching (e.g., Ball, 1991, 1993; Berry, 2003; Lampert, 1990, 1992; Whitenack & Yackel, 2002). For instance, the council has heralded discourse practices as a cornerstone of mathematics education reform, calling for student participation at the level of "presenting and defending conjectures" and debating mathematical ideas. Nowhere in our study did we see clear evidence of this occurring in classrooms in the United States or China. Rather, the content of the extended discourse suggests that, for the majority of questions, students were being asked to provide specific, predetermined answers and to justify those answers using standard algorithms; memorized rules; and, in some cases, mathematical reasoning. Although we believe such discussions can contribute to greater student understanding of mathematics, the evidence we saw regarding the use of extended discourse does not convince us that this talk is different enough from traditional forms of discourse (Cazden, 2001) to promote such grander aims as engaging students in the authentic discussions of mathematicians (Lampert, 1992) or as satisfying the discourse needs of students from diverse backgrounds (Berry, 2003). No matter how interesting the practices embedded in extended discourse, we still have a long way to go toward helping teachers use discourse forms such as extended discourse to implement the kinds of pedagogy proposed by the National Council of Teachers of Mathematics.

References

An, S., Kulm, G., & Wu, Z. (2004). The pedagogical content knowledge of middle school mathematics teachers in China and the U.S. *Journal of Mathematics Teacher Education, 7,* 145–172.

Bakhtin, M. M., Holquist, M., & Emerson, C. (1986). *Speech genres and other late essays* (1st ed.). Austin: University of Texas Press.

Ball, D. L. (1991). What's all this talk about "discourse"? *Arithmetic Teacher, 39*(3), 44–48.

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

Ball, D. L. (1996). Teacher learning and the mathematics reforms: What we think we know and what we need to learn. *Phi Delta Kappan, 77,* 500–508.

Berry, R. Q. (2003). Mathematics standards, cultural styles, and learning preferences: The plight and promise of African American students. *Mathematics Standards, 76,* 244–249.

Brown, C. A., Stein, M. K., & Forman, E. A. (1996). Assisting teachers and students to reform the mathematics classroom. *Educational Studies in Mathematics, 31,* 63–93.

Cai, J. (2001). Improving mathematics learning: Lessons from cross-national studies of Chinese and U.S. students. *Phi Delta Kappan, 82,* 400–404.

Cai, J., & Silver, E. A. (1995). Solution processes and interpretations of solutions in solving a division-with-remainder story problem: Do Chinese and U.S. students have similar difficulties? *Journal for Research in Mathematics Education, 26,* 491–497.

Cazden, C. B. (2001). *Classroom discourse* (2nd ed.). Portsmouth, NH: Heinemann.

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

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Crosswhite, F. J., Dossey, J. A., Swafford, J. O., McKnight, C. C., & Cooney, T. J. (1985). *Second International Mathematics Study summary report for the United States.* Champaign, IL: Stipes.

Fan, L., & Zhu, Y. (2004). How have Chinese students performed in mathematics? A perspective from large-scale international mathematics comparisons. In L. Fan, N.-Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 3–26). Singapore: World Scientific Press.

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

Fennema, E., Franke, M. L., Carpenter, T. P., & Carey, D. A. (1993). Using children's mathematical knowledge in instruction. *American Educational Research Journal, 30,* 555–583.

Flevares, L. M., & Perry, M. (2001). How many do you see? The use of nonspoken representations in first-grade mathematics lessons. *Journal of Educational Psychology, 93,* 330–345.

Gee, J. P. (1999). *An introduction to discourse analysis: Theory and method.* London: Routledge.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* London: Sage.

Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. In J. C. Greene & V. J. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 5–17). San Francisco: Jossey-Bass.

Hannum, E., & Park, A. (2002). Educating China's rural children in the 21st century. *Harvard China Review, 3*(2), 8–14.

Hatano, G. (1988). Social and motivational bases for mathematical understanding. In G. B. Saxe & M. Gearhart (Eds.), *Children's mathematics:* New directions for child development (no. 41, pp. 55–70). San Francisco: Jossey-Bass.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., et al. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher, 25*(4), 12–21.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Erlbaum.

Hiebert, J., & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS video study. *Elementary School Journal, 101,* 3–20.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30,* 393–425.

Huang, R., & Leung, K. S. F. (2004). Cracking the paradox of Chinese learners: Looking into the mathematics classrooms in Hong Kong and Shanghai. In L. Fan, N.-Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 348–381). Singapore: World Scientific Press.

Kazemi, E. (1998). Discourse that promotes conceptual understanding. *Teaching Children Mathematics, 4,* 410–414.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *Elementary School Journal, 102,* 59–80.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: National Academy Press.

Kruskal, J. B., & Liberman, M. (1983). The symmetric time warping algorithm: From continuous to discrete. In J. B. Kruskal & D. Sankoff (Eds.), *Time warps, string edits and macromolecules* (pp. 125–162). Reading, MA: Addison-Wesley.

Kumar, S. (2004). *Finding patterns in sequences: Comparison of motif extraction, dynamic time warping, and hidden Markov model approaches with applications to the TIMSS 1999 video study.* Unpublished master's thesis, University of Illinois at Urbana–Champaign.

Ladson-Billings, G. (1997). It doesn't add up: African American students' mathematical achievement. *Journal for Research in Mathematics Education, 28,* 697–708.

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

Lampert, M. (1992). Practices and problems in teaching authentic mathematics in school. In F. Oser, A. Dick, & J.-L. Patry (Eds.), *Effective and responsible teaching: The new synthesis* (pp. 295–314). New York: Jossey-Bass.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Erlbaum.

McClain, K., & Cobb, P. (1998). The role of imagery and discourse in supporting students' mathematical development. In M. Lampert & M. L. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 56–81). New York: Cambridge University Press.

Mehan, H. (1979). *Learning lessons.* Cambridge, MA: Harvard University Press.

Miller, K. F., Correa, C. A., Sims, L., Noronha, N. C., & Fang, G. (2005, April). Deconstructing classroom dialog: U.S. and Chinese comparisons. In M. Perry (chair), *Discourse in elementary U.S. and Chinese mathematics classrooms: Building understanding of practices and possibilities.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Moghaddam, F. M., Walker, B. R., & Harré, R. (2003). Cultural distance, levels of abstraction, and the advantages of mixed methods. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods* (pp. 111–134). Thousand Oaks, CA: Sage.

Myers, C. S., & Rabiner, L. R. (1981). A comparative study of several dynamic time-warping algorithms for connected word recognition. *The Bell System Technical Journal, 60,* 1389–1409.

Nassaji, H., & Wells, G. (2000). What's the use of "triadic dialogue"?: An investigation of teacher-students interaction. *Applied Linguistics, 21,* 376–406.

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.

O'Conner, M. C. (1998). Language socialization in the mathematics classroom: Discourse practices and mathematical thinking. In M. Lampert & M. L. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 17–55). New York: Cambridge University Press.

Perry, M. (2000). Explanations of mathematical concepts in Japanese, Chinese and U.S. first- and fifth-grade classrooms. *Cognition and Instruction, 18,* 181–207.

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.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 75–110). East Sussex, England: Psychology Press.

Rocco, T. S., Bliss, L. A., Gallagher, S., Perez-Prado, A., Alacaci, C., Dwyer, E. S., et al. (2003). The pragmatic and dialectical lenses: Two views of mixed methods use in education. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods* (pp. 595–615). Thousand Oaks, CA: Sage.

Santagata, R., & Stigler, J. W. (2000). Teaching mathematics: Italian lessons from a cross-cultural perspective. *Mathematical Thinking and Learning, 2,* 191–208.

Sfard, A. (2001a). Learning mathematics as developing a discourse. In R. Speiser, C. Maher, & C. Walter (Eds.), *Proceedings of the twenty-first annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 23–44). Columbus, Ohio: Clearing House for Science, Mathematics, and Environmental Education.

Sfard, A. (2001b). There is more to discourse than meets the ears: Looking at thinking as communication to learn more about mathematical learning. *Educational Studies in Mathematics, 46,* 13–57.

Smith, T. M., Desimone, L. M., & Ueno, K. (2005). "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational Evaluation and Policy Analysis, 27,* 75–109.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36,* 404–411.

Stevenson, H. W., Chen, C., & Lee, S.-Y. (1993, January 1). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science, 259,* 53–58.

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,* 1053–1066.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap.* New York: Summit.

Stieglitz, E. L., & Oehlkers, W. J. (1989). Improving teacher discourse in a reading lesson. *The Reading Teacher, 42,* 374–379.

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction: An overview of the TIMSS video study. *Phi Delta Kappan, 79,* 14–21.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap.* New York: Free Press.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61,* 12–17.

Stigler, J. W., & Perry, M. (1988). Mathematics learning in Japanese, Chinese, and American classrooms. In G. B. Saxe & M. Gearhart (Eds.), *Children's mathematics. New directions for child development* (No. 41, pp. 27–54). San Francisco: Jossey-Bass.

Whitenack, J., & Yackel, E. (2002). Making mathematical arguments in the

primary grades: The importance of explaining and justifying ideas. *Teaching Children Mathematics, 8,* 524–527.
Wong, N-Y., Han, J., & Lee, P. Y. (2004). The mathematics curriculum: Toward globalization or Westernization? In L. Fan, N.-Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 27–70).
Yao, Y. (1992). American and Chinese schools: A comparative analysis. *Education, 113,* 232–235.

Appendix

Explanation of Content Codes for Function of Teacher and Student Statements in Extended Discourse

Request for Computation (Teacher) or Computation (Student)

A teacher exhibited a request for computation when he or she asked for the result of a calculation, such as "What is 9 minus 3?" or when he or she asked the students to fill in a blank in an equation or write an answer on the board. A student statement was coded as computation if it was a brief numerical answer, such as "3/4" or "1/2 equals 2/4."

Request for Procedure (Teacher) or Procedure (Student)

A teacher requested procedures or methods by asking students what they did to produce a certain answer or how they could illustrate the answer. A student statement was coded as a procedure statement when the student outlined the mathematical procedure for producing a certain computation or short answer. For instance, a student might explain the answer of 1/3 by saying, "I divided both the numerator and denominator of 3/9 by 3." A student statement was also coded as a procedure statement if it was a method derived using mathematical manipulatives. An example of this could be a student explaining 1/3 by illustrating one shaded rectangle out of three rectangles. It is important to note that the procedures being coded here were not student-invented approaches to problems but were student applications of school-taught algorithms for solving certain types of problems.

Request for Reasoning (Teacher) or Reasoning (Student)

Requesting reasoning could be characterized in three different ways. First, teachers could ask students why a certain procedure or computation was appropriate (e.g., "Why did you multiply by 5?") or why a certain aspect of a rule was necessary. Second, teachers could ask for a pattern or final conclusion about certain numerical problems (e.g., "What did you notice about those three numbers?"). Last, teachers could question what happens when students apply certain processes (e.g., "What happens when you multiply the numerator and denominator by the same number?").

Student reasoning was characterized correspondingly in three ways. First, a student could provide rule-based reasoning about mathematical processes. For instance, a student might explain that he multiplied the numerator and the denominator of a fraction by 3 because "the fraction doesn't change if the numerator and denominator are multiplied at the same time by the same number." Second, reasoning included noticing or explaining a pattern or final conclusion about a set of numbers. For example, students sometimes explored the equivalency of certain fractions. A student might have found that 1/3, 2/6, and 3/9 were equivalent and reasoned that "if the multiple is equivalent, the fraction is equivalent." Finally, reasoning could involve explaining aspects of rules. The rationale of consistent quotient, for instance, states that one can multiply the numerator and denominator by any number, except zero, to keep the fraction the same. It was common for the students to reason that the phrase "except zero" is in this rationale because the denominator cannot legally be zero. Thus, the students explained the necessity of an aspect of a rule.

The type of reasoning being coded here does not correspond with the idea of conceptual reasoning in mathematics. The reasoning code is not intended to indicate inductive, deductive, or intuitive reasoning. The code is simply meant to capture statements in which students discuss certain problems or questions in generalized, abstract ways. These abstract responses do indicate ability to reason about specific problems but do not suggest the presence or absence of specific conceptual reasoning capabilities.

Request for Rule Recall (Teacher) or Rule Recall (Student)

Requesting rule recall occurred when the teacher asked for the names of rules or terms or when the teacher asked for a rule or term to be described in detail. Rule recall was the code when students stated the name of a rule or term (such as "Those are real fractions"), recited the full definition of a memorized rule, or recited part of the full definition of a memorized rule (such as "The denominator can be multiplied by any number, except zero").

Check for Understanding or Agreement (Teacher) or Indication of Understanding or Agreement (Student)

A teacher's statement was coded as a check for understanding or agreement if it was used to assess student understanding of a discussion or if it was used to assess student agreement with an answer or explanation. A check for understanding or agreement

(*Appendix continues*)

statement frequently took the simple form of “Do you agree?” or “Understand?” Students exhibited this understanding or agreement in indication of understanding or agreement statements such as “yes” or “I agree.” Students sometimes failed to respond to teacher checks for understanding, but they rarely said “no” in response to a check for understanding. If they did respond negatively to a check for understanding, it was still coded as an indication of understanding, as it alerted the teacher to a lack of complete comprehension.

Request for Short Answer (Teacher) or Short Answer (Student)

Request for short answer was used to code statements in which teachers posed multiple-choice questions such as true–false and yes–no. Short answer coded student responses to such questions.

Teacher Explanation (Teacher)

Teacher explanation was coded when teachers summarized the mathematical ideas that were discussed within an episode of discourse or explained new material within an episode.

Praise (Teacher)

Praise indicated teacher approval of student answers and statements.

Restatement/Revoicing (Teacher)

Restatement or revoicing captured statements in which a teacher repeated or rephrased a student answer or statement.

Received February 14, 2006
Revision received November 14, 2006
Accepted December 31, 2006 ■