Journal of Language, Identity & Education

ISSN: 1534-8458 (Print) 1532-7701 (Online) Journal homepage: https://www.tandfonline.com/loi/hlie20

Mathematical Classroom Discussion as an Equitable Practice: Effects on Elementary English Learners' Performance

Leslie C. Banes, Rebecca C. Ambrose, Robert Bayley, Rachel M. Restani & Heather A. Martin

To cite this article: Leslie C. Banes, Rebecca C. Ambrose, Robert Bayley, Rachel M. Restani & Heather A. Martin (2018) Mathematical Classroom Discussion as an Equitable Practice: Effects on Elementary English Learners' Performance, Journal of Language, Identity & Education, 17:6, 416-433, DOI: 10.1080/15348458.2018.1500464

To link to this article: https://doi.org/10.1080/15348458.2018.1500464

JOURNAL OF LANGUAGE, IDENTITY & EDUCATION
2018, VOL. 17, NO. 6, 416–433
https://doi.org/10.1080/15348458.2018.1500464

Routledge
Taylor & Francis Group

Check for updates

Mathematical Classroom Discussion as an Equitable Practice: Effects on Elementary English Learners' Performance

Leslie C. Banes, Rebecca C. Ambrose, Robert Bayley, Rachel M. Restani, and Heather A. Martin

University of California, Davis

ABSTRACT

This mixed-method study examines the relationship between classroom discussion and student performance in twenty 3rd and 4th grade classrooms in northern California with 50% English language learners (ELLs). Discussions were scored on features including use of multiple approaches for solving problems, students' opportunities to speak, equitable participation, explanations, and connections between ideas. A linguistically-modified math assessment measured student performance. Quantitative analysis using hierarchical linear modeling (HLM) showed that the discussion features *variety of approaches* and *equitable participation* significantly contributed to the explanation of between-class variation in assessment scores, above and beyond that explained by prior mathematics performance and English proficiency. Importantly, mathematical discussion was equally beneficial for students classified as ELLs and those not classified as ELLs. Two classroom vignettes illustrate the different features of discussion and offer insight into ways ELLs contributed to the discussion in different contexts.

KEYWORDS

Classroom discourse; English learners; math; mathematical classroom discussion; mixed method; student performance

A growing body of research has captured the ways in which English language learners (ELLs) can engage productively in mathematical discussions in classrooms with English as the language of instruction (e.g., Hansen-Thomas, 2009; Khisty & Chval, 2002; Moschkovich, 2007; Takeuchi, 2015; Turner, Dominguez, Maldonado, & Empson, 2013). Although demonstrating ELLs are capable of engaging in rich mathematical discussions while learning English, this research often assumes communicating about mathematics is an end in itself. Many in the broader community do not share this assumption and only value classroom discussion if it promotes success on written measures of performance. Few studies, however, have analyzed the effect of discussion on written assessments, and even fewer have looked specifically at ELLs. Many stakeholders will continue to balk at reform-based instruction, including mathematical discussion, if educators cannot demonstrate that discussion positively affects students' performance on assessments.

Achievement is one of four dimensions addressing equity in mathematics classrooms, along with access, identity, and power (Gutiérrez, 2007). From this view, achievement involves more than performance on assessments, though because high-stakes decisions impacting course selection, retention, and graduation are often based on students' test scores, they remain crucial in considerations of equity. The lack of attention to how discussion-based math instruction may impact ELLs' performance on assessments raises questions about the supposed benefits of discussion for these students. The growing population of language learners in the United States and abroad, and the

CONTACT Leslie C. Banes lcbanes@ucdavis.edu School of Education, University of California, Davis, 1 Shields Ave, Davis, CA 95616.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hlie.

Supplemental data for this article can be accessed here.

increased emphasis on mathematical communication within the Common Core State Standards makes attention to this issue particularly critical.

Important educational reforms tend to bypass classrooms with high percentages of ELLs. It may be assumed that ELLs do not possess the linguistic skills necessary to participate in reform-based classrooms (Spillane, 2001). Indeed, several studies suggest reform approaches to instruction may affect students differentially. Baxter, Woodward, and Olson (2001) argue low-performing students in elementary classrooms tend to avoid participation in math discussion and tune out if the verbal or cognitive demand is too high. Others point to the extraordinary discursive demands math discussion places on ELLs (Barwell, 2012), suggesting they can get lost in the spontaneity of interactions between multiple speakers. Furthermore, a recent study of first grade math classrooms found the clarity offered by direct instruction can be more effective than hearing strategies from peers. However, this study controlled for students' ELL status, and thus, obscured the impact of particular strategies on ELLs (Morgan, Farkas, & Maczuga, 2015).

Several researchers exploring the participation of students not yet proficient in the language of instruction, including ELLs and learners of other languages, have found that many students struggle to produce meaningful explanations in their second language and their participation in mathematics discussions can be quite limited (Adler, 1997; Gorgorió & Planas, 2001; Secada & De La Cruz, 1996). As well, researchers have found that students who are left out of the discussion do not acquire the same positive disposition toward mathematics as those who participate (e.g., Ball, 1993; Boaler, 2000).

On the other hand, some studies of reading and math instruction in elementary classrooms suggest students at the low end of the preparedness/achievement spectrum can benefit from class discussion as much as those at the high end (Connor, Morrison, & Katch, 2004; Griffin, 2004). Webb (1991), for example, found the act of producing "elaborate explanations" was correlated with high performance on assessments. Further, Chapin, O'Connor, and Anderson (2003) found benefits of mathematical discussion for elementary and middle school students, some of whom were ELLs, involve building a deeper understanding of concepts and improving motivation. Importantly, many math educators argue the ability to communicate mathematically is an essential part of what it means to do mathematics, regardless of its impact on assessments (e.g., NCTM 2000; Turner et al., 2013). For instance, Niemi (1996) contends, "Students who cannot effectively explain the meaning of, and justify the use of, mathematical symbols, concepts, and operations are not yet fully fledged members of the community of discourse" (p. 361).

The evidence of the possible benefits of discussion for ELLs is unclear. In an attempt to clarify the issue, we analyzed the impact of discussion on the math performance of ELLs in 20 third and fourth grade classrooms using hierarchical linear modeling (HLM). The following questions guided our analysis: (1) Do mathematical classroom discussions affect all students in a similar manner, regardless of their status as English language learners?, and (2) How do discussions affect English language learners?

Theoretical framework

This article focuses on benefits of discussion that can be associated with improved performance on achievement measures for ELLs. We share the view of many within mathematics education that students construct their understanding of mathematics by going through an iterative process of working on problems and then discussing their attempts under the guidance of a teacher who orchestrates discussion by strategically eliciting student contributions and encouraging students to make sense of one anothers' ideas (e.g., Hufferd-Ackles, Fuson, & Sherin, 2004; Kazemi & Stipek, 2001).

Drawing on the work of Forman, McCormick, and Donato (1998), we define classroom discussion in mathematics as an academic activity in which students actively participate by listening, speaking, and engaging in thinking about mathematical ideas. Classroom research has shown that *quantity* of student talk is not enough to produce an effective math discussion. The *quality* of student talk is equally important (Smith & Stein, 2011; Truxaw & DeFranco, 2008). Based on the literature, we identified five key features of effective math discussion: (1) variety of approaches, (2) opportunities to speak,

(3) equitable participation, (4) explanations, and (5) connections between ideas. In the following sections, we introduce these features which are illustrated in detail in the methods section.

Literature review

Studies of class discussion with ELLs

Sociolinguistic studies of classroom discourse have found whole-class interactions often emerge as a three-part exchange, the *Initiate-Respond-Evaluate* (IRE) sequence (Mehan, 1979), still prevalent across all grade levels and content areas (Cazden, 2001). In the first stage of this sequence, the teacher initiates an interaction by asking a question, to which there is usually one acceptable answer. In the second stage, the student responds to the question, often producing a numerical answer. The final stage is the teacher's evaluation of the response as right or wrong. In this sequence, students speak only in response to the teacher's questions and seldom have opportunities to voice ideas that fall outside the range of acceptable answers. Class discussion can provide an alternative to the IRE sequence when teachers are genuinely interested in the way their students understand the mathematics, students listen and respond to one another's ideas, and misconceptions are allowed to surface (Cirillo, 2013).

Few studies of the effect of discussion on written measures of mathematics achievement have included ELLs. In this section, we primarily review the work of Chapin et al. (2003) and Lane, Silver, and Wang (1995), who specifically look at how discussion impacts ELLs' mathematics performance. Chapin and colleagues report on findings from Project Challenge, which included 400 students in fourth through seventh grade, of whom 75% spoke a language other than English and 85% qualified for free or reduced-price lunch. Participating students were placed in project math classrooms where they received daily instruction from teachers who worked closely with the researchers. Teachers employed a special curriculum emphasizing conceptual understanding and discussion. After three years, researchers found 90% of the students in the original cohort surpassed the performance of students from affluent schools on the Massachusetts Comprehensive Assessment System (MCAS). Although promising, these findings combine the effects of the curriculum with class discussion, making it difficult to evaluate the impact of discussion alone.

Lane et al. (1995) reported findings from the Quantitative Understanding: Amplifying Students' Achievement and Reasoning (QUASAR) study, an intervention including emphasis on conceptual understanding and discussion. In one school, 20% of students were Latino, most of whom were enrolled in bilingual (Spanish-English) classes. Researcher-developed assessments were administered in both languages. Findings suggest that while students in bilingual classes performed lower, on average, than students in English-only classes in 6th grade, by 8th grade they had mostly caught up. These results indicate that mathematics instruction emphasizing discussion can be as effective for bilingual students in bilingual classrooms as it is for monolingual students in English-only classrooms. However, because Spanish-speaking students had access to instruction, discussion, and assessments in their native language, this study does not help us understand how discussion conducted in English impacts students who have not yet attained English proficiency. A few other studies (e.g., Kosko & Miyazaki, 2012) have investigated the impact of discussion on students' performance on mathematics assessments; however, they do not discuss the extent to which findings hold true for language learners.

Variety of approaches

Discussions in mathematics classroom can vary in several ways, and for the purposes of this study we focused on five features. First, we considered that multiple ways of solving a problem was the primary content of a quality discussion (Akkus & Hand, 2010). The literature suggests instruction emphasizing multiple ways to solve a problem offers students access to ideas they may not have thought of on their own. When allowed to choose their own strategies, students tend to use concrete tools and manipulatives when they are first beginning to understand, and advance to more abstract strategies as their understanding increases (Carpenter, Fennema, & Franke, 1996). Presenting ideas using multiple resources, such as concrete and visual representations, can increase comprehensibility for ELLs who may struggle to understand verbal explanations (Echevarria, Vogt, & Short, 2007). Knowledge of a variety of strategies provides students with back-up strategies should their first attempt at a problem fail, thus supporting students in successful performance on assessments.

Opportunities to speak

Given the potential benefits of discussion for students' understanding and self-concept in mathematics, students would ideally engage in talk during every lesson. However, in traditional math classrooms, teachers still do most of the talking (Stigler & Hiebert, 1998). Thus, following Truxaw and DeFranco (2008), we set out to capture the extent to which student voices were prevalent in the lessons. Our notion of opportunities to speak includes tracking the frequency of opportunities for student talk in all configurations (partner, small group, and whole-class), as well as the length of time allotted for each interaction.

Equitable participation

ELLs are often left on the periphery of mathematics discussion, while native English speaking students and high performers tend to receive the majority of talk time (Ball, 1993). Our view of equitable participation accounts for the total number of students who participate in whole-class discussion during a lesson, with special attention to inclusion of ELLs and whether or not a few students are allowed to dominate (Planas & Gorgorió, 2004). In response to Moschkovich's (2013) call to broaden what counts as participation, we include both verbal and non-verbal displays of thinking, such as using hand signals to agree or disagree (discussed further in the classroom vignettes that follow). ELLs may have an easier time understanding explanations from peers than from the teacher because peers' sentence structure and vocabulary tend to more closely align with their own levels of language use (Ellis, 1999; Varonis & Gass, 1985). Repetition also facilitates comprehension for language learners, so hearing the same idea presented by a variety of speakers offers greater access than hearing it only from the teacher (e.g., Chapin et al., 2003).

Explanation

Cognitive scientists have found that when students explain ideas to themselves, they tend to learn more than when they work in silence, and when they explain ideas to someone else, they learn even more (Rittle-Johnson, Saylor, & Swygert, 2008). Further, when students develop conceptual explanations that explain how and why a strategy works, they learn more than when they develop purely procedural explanations (Matthews & Rittle-Johnson, 2009). To realize the benefits of class discussion, students' contributions need to be substantive, consisting of more than one or two word answers so they have the opportunity to articulate and revise their mental models (Chi, 2000). Thus, our notion of explanation is an attempt to capture a meaningful and thorough articulation of concepts, with greater value given to explanations produced by students than by the teacher.

Connection between ideas

Finally, our interpretation of connections between ideas prioritizes the building on and connecting of ideas, rather than a show and tell of unrelated ideas (Hufferd-Ackles et al., 2004). While hearing multiple problem solving approaches can provide students with a specific approach that matches their developmental level, considering the connections between various approaches can deepen their mathematical understanding, allowing students to solve a range of problems (Hiebert & Carpenter, 1992). The robust understanding that grows out of making connections among strategies and across problem types should enable students to transfer their knowledge to the unfamiliar problems that might appear on assessments.

Taken together, the literature suggests these features, further illustrated in the upcoming classroom vignettes, may affect ELL students' performance on written mathematics tests. However, the potential benefits will only be apparent if tests do not contain the undue linguistic complexity that can confound knowledge of mathematics with language proficiency (Abedi & Lord, 2001). Walshaw and Anthony (2008) point out the "enormous complexity" (p. 543) of facilitating discussion with all these features. Complexity is increased in classrooms with linguistically diverse learners at various stages of English development. Given the difficulty of orchestrating productive discussions and the amount of class time required, it is crucial to gain a better understanding of how they impact students.

Methods and data sources

Participants

Data for this study were collected in a small K-6 urban school district with five schools. At the time of this study, 37% of students were Hispanic, 16% white, and 15% African American. Forty-five percent of students in the district were designated English learners and 92% received lunch subsidies, an indicator of high poverty levels. On average, teachers had been in the district for 15 years and the annual teacher mobility was less than 5%. Fourteen out of 20 teachers in the study participated in voluntary professional development provided by the authors, with participation ranging from 22 to 131 hours over three years. Professional development focused on discussion, anticipating students' responses, and extending teachers' own mathematical understandings. Math discussion and problem solving was emphasized by the school district and the newly adopted curriculum, EnVision Math.

Our first research question explores differences in the relationship between discussion and math performance for students who were classified as ELLs and those who were not, and was based on 410 students from 20 classrooms representing nearly all the third and fourth classrooms in the district. The second research question guides us in taking a deeper look at the classroom interactions that include the 217 students classified as ELLs, totaling 53% of the student population across these classrooms. In California, the site of this study, to be classified as an ELL means a student's parents reported at least one family member spoke a language other English in the child's home at the time of enrollment and the student's scores on the California English Language Development Test (CELDT) indicate they are "limited English proficient" (California Department of Education, 2017). In fact, many students classified as ELLs using this system are actually raised bilingually or come from homes in which English is spoken a majority of the time. We use the current ELL classification system, despite its imperfections, as a useful way to investigate the impact of math discussion on students who are in the process of acquiring the language of instruction. Students not currently identified as ELLs (non-ELLs) include native English speakers and those who were initially identified or re-designated fluent English proficient.

Seventeen different languages other than English are represented across the 20 classrooms, including Spanish (62% of ELLs), Hmong (17%), Russian (4%), and Hindi (4%). The average class size was 22 students in third grade classrooms and 31 students in 4th grade classrooms. The percentage of ELLs in each class varied slightly, as did the average English proficiency level of ELLs. Table 1 shows the average percentage of ELLs and average English proficiency scores for each grade, according to the CELDT. The percentage of ELLs is slightly lower in fourth grade than in

Table 1. Descriptive statistics of ELLs by grade.

Grade	Number of Classes	Mean % of Students who are ELLs	Mean CELDT Level of ELLs (1–5)
3	11	55%	2.94
4	8	47.5%	3.12
3/4	1	48%	4.08

third because each year a few students are re-designated fluent English proficient, and thus, are no longer ELLs. The 3rd/4th grade combination class is for Gifted and Talented Education (GATE) students and the ELLs in this class have higher average CELDT scores than those in other classes.

Measures of mathematics performance

To measure mathematics performance, we created our Linguistically Modified Math Assessment (LMMA) by adapting problems from past Trends in Math and Science Study (TIMSS) and National Assessment of Educational Progress (NAEP) assessments. Experts argue for the need to distinguish between ELLs' content knowledge and their English language proficiency, though it is recognized that separating the two completely may not be possible (Abedi, 2004; Martiniello, 2008). Strategic modification of word problems is one way of increasing comprehensibility and reducing linguistic bias for ELLs (Abedi & Lord, 2001).

Our research team, which included a linguist, modified the language of items according to guidelines for assessing ELLs, while keeping the mathematics and content-related vocabulary the same. We performed differential item functioning on the test, which demonstrated that no single item was significantly more difficult for ELLs than for their peers. This suggests our linguistic modification may have reduced any substantial linguistic bias.

The LMMA includes a mix of multiple choice, open response, and explanation items. It also contains multi-step word problems and items with more than one possible solution. The third grade test includes 11 items, while the fourth grade test includes these plus an additional 6 items. Both tests include items covering a variety of mathematical concepts, with about one third of the items on geometry, one half on number concepts, and the remaining items on fractions/decimals and data representation. We believe the LMMA is more closely aligned to the Common Core State Standards and more sensitive to the kinds of mathematical thinking promoted by discussion than standard multiple-choice tests. The LMMA was administered in spring 2013. Students' California State Test (CST) math scores from the previous year were used as a control variable to account for students' baseline mathematics performance. Readers interested in more information on the Linguistically Modified Math Assessment, including item modification guidelines, scoring, and the complete assessments, may refer to the journal's online supplement.

Measure of mathematics discussion

Two raters, unfamiliar with the teachers, attended one lesson for each classroom and, using a rubric developed by our research team (Appendix A), rated discussions according to the five features described above. The rubric has a 4-point scale for each feature, developed using existing literature. Though other researchers studying math discussion have developed rubrics, none fit the particular needs of this study. We required a rubric that would allow us to code live, without the use of video, and because teachers all used the district-adopted curriculum, we wanted to maintain a focus on more general features of math discussion, and not the nature of the math tasks determined by the curriculum.

During the classroom visits, raters moved quietly throughout the room to see students' written work and use of manipulatives, and to listen to partner, group, and whole class discussions. They took notes focusing on what was said, who was called on, and how much time was allotted for talk.

They also captured work shown on the board or projector and noted the nature of the learning environment and teacher-student rapport. The raters each had several years of teaching experience contributing to their understanding of classroom interactions. Teachers provided seating charts indicating where ELLs sat so raters could pay special attention to how they participated.

After the lesson, raters independently used the rubric to generate scores for the discussion, then met to debrief and resolve any differences in scores. Inter-rater agreement was computed for overall ratings. Scores within 10% of each other were considered to be consistent. Inter-rater agreement was 75%. When raters' initial scores were divergent, they returned to the classroom for another observation. This occurred two times.

Vignettes of math class discussion

The following two vignettes, derived from observations, illustrate how we used the rubric to score each of the five features enacted in two different classroom discussions. These vignettes are representative examples of a high-scoring and a mid-scoring discussion.

High-scoring discussion

Students in a third-grade class were working on the following problem:

> Lou is painting a shelf. She paints 2/8 purple. Then she paints 4/8 more of the shelf gray. How much of the shelf has she painted in all?

After reading the problem aloud clause-by-clause several times, the teacher gave students five minutes to "get a start, but don't solve it yet." He then selected students to present their "starts" to the class; it was apparent this was a regular classroom routine. The following excerpt shows Aria, a particularly outgoing ELL with intermediate English proficiency, presenting her "start" at the front of the room with her paper projected. The teacher acted as facilitator, reminding presenters to "check your audience," and asking them to revoice key ideas. The apparent assumption in this classroom was that all students have something valuable to contribute and deserve to be listened to.

Aria: [presenting] I colored it 2 purples because she said it was 2/8 purple and 4/8 gray. [pause]
Teacher: Can you show us that in your drawing? [Aria points to 2 sections shaded with pencil in her drawing of a fraction bar divided into eight sections.]
Teacher: What made you decide to color that in?
Aria: I colored it in because it's going to show how much she colored in all.
[several students signal agreement using hand signals, unprompted by teacher]
Teacher: Any questions for Aria?
Joey: Why did she add the 2/8 and the 4/8? [looking at teacher]
Teacher: Are you asking Aria?
Joey: How come you added the 4/8 and 2/8? [now looking at student presenter]
Aria: Because it's going to tell how much she colored altogether. [pause] I didn't color it with pencil because it would be mixed up together.

Like other discussions with a score of 4 in *variety of approaches*, this teacher presented an open-ended problem and let students decide how they would solve it. By selecting several students who solved the problem in different ways to share their strategies with the class, the teacher ensured all students access to multiple ways of solving.

More than half the students, including many ELLs, contributed to the discussion at least once, earning this discussion a 4 in *equitable participation*. In addition to contributing verbally, this class used hand signals to communicate agreement with the student presenter, a practice that was prevalent in several classrooms. However, in this classroom, the teacher did not have to remind students to use the signal. Instead, the students used hand signals spontaneously, without prompting to display their thinking. This lesson received a 3 in the category *opportunities to speak* because students were given short opportunities to speak with partners, and several students were given time to produce extended responses without interruption in whole class discussion. If students had received more frequent opportunities to speak at length, this lesson would have received a 4.

Several times during the discussion, the teacher pressed students to *explain* how they decided on a particular strategy. Like Aria, several students in this discussion produced meaningful, *partial* explanations, thus earning a 3 for explanations. We considered Aria's explanation partial because her explanation did not include that the denominator refers to 8 equal parts, or the whole shelf.

Students showed a genuine interest in making sense of their classmates' thinking. When Joey asked Aria *why* she added the fractions, he pressed for a justification that led her to connect back to the situation described in the word problem. Later in the lesson, students had opportunities to compare and connect multiple ways of solving, moving the discussion beyond a show and tell of strategies, and helping the discussion reach a 4 for *connections between ideas*.

This vignette illustrates a small part of the rich interaction that took place during high scoring discussions. The teacher supported student-student interaction by directing questions back to students, giving them agency and positioning them as mathematics doers and thinkers. He alerted students ahead of time they would be presenting, enabling ELLs to produce better, more thorough articulations of their thinking. Encouraging students to refer to visuals while they presented, as this teacher did, may also support the communication of mathematical ideas. Few classrooms we visited achieved this level of math community.

Mid-scoring discussion

This vignette contains an excerpt from a 4th grade discussion that scored a 14 out of 20 on our rubric. The midrange score illustrates a teacher grappling with the complexity of engaging students in discussion, but without the sophistication of teacher moves we saw in the first vignette. In this lesson, students compared and ordered three decimal numbers with the teacher prompting them to explain how they knew one number was greater than another. The following excerpt includes two ELLs, Laura, a Hmong-speaking student with early-advanced English proficiency and David, a Spanish-speaking student with intermediate proficiency. Several students were visibly disengaged and the teacher had to interrupt the lesson several times to remind students of her expectations. The class was discussing how to order 0.67, 0.66, and 0.7 from least to greatest.

Laura: 0.7, 0.66, 0.67
Teacher: How did you do it?
Laura: I looked at the hundreds.
Teacher: Does anybody see it differently? Emily? [short pause. No response.]
Teacher: What should go first? I heard you tell your partner. Just say it. I know you know it. This is your chance to shine. (pause) Bummer. David?
David: 66
Teacher: 66 what?
David: Hundredths
Teacher: What did he say? [asking for a volunteer to revoice David's response]
Sarah: 70 hundredths
Teacher: How many agree with David? [pause, No response.]
Teacher: Remember we said you could put a 0 to make .70. 7 tenths is the same as 70 hundredths. Laura, does that change your mind?
Laura: No.
Teacher: David you have to convince Laura. [short pause. No response.]
Teacher: Laura, why do you believe you're right? [short pause. No response. Teacher moves on to another problem.]

This vignette illuminates the complexity of orchestrating productive mathematical discussions. This teacher demonstrated skill in offering access to multiple ways of solving a problem and getting most students, including ELLs, to contribute verbally at least once during the lesson. This lead us to score the discussion a 4 in *variety of approaches* and *equitable participation*. Additionally, this teacher let a student share a wrong answer without immediately correcting her, offering students the opportunity to share their thinking.

Students in this lesson were provided many *opportunities to speak*, but only in very short bursts. In fact, the one opportunity for pair sharing during this lesson was only 30 seconds, leading to a score of 2 for opportunities to speak. Students responded only to teacher questions and they gave short responses, usually no more than a few words. The teacher asked for a volunteer to revoice David's answer, positioning him as having mathematical ideas worth listening to, but she seemed not to notice when his idea was revoiced incorrectly.

Later in the lesson, the teacher explained why tenths are bigger than hundredths using the example of cutting a pizza: "if you cut the pizza into ten slices each piece would be larger than if you cut it into one hundred slices." We considered this a partial *explanation* and scored it as a 2. Ideally, we would have seen students produce this explanation, which would have increased the score to a 3. Although the teacher encouraged students to describe their strategies, neither students nor the teacher offered substantial connections *between* the strategies or other mathematical ideas, earning this discussion a 2 for *connections between ideas*. In contrast to the first vignette, this teacher had to prompt students to use their hand signals to communicate agreement, and did so even when it seemed most students did not understand or were tuned out.

This class did not achieve the same level of student-student discussion, as in the first example, and it is clear the teacher still has work to do to engage all students in meaningful discussion. All students, but perhaps especially ELLs, need to feel safe in order to publicly share their mathematical thinking, and the classroom atmosphere in this lesson did not afford that level of safety. Nonetheless, there were opportunities for students to articulate their ideas and gain access to others' thinking. This vignette shows in some lessons different features of discussion were assigned various scores. In this case, variety of approaches and equitable participation received scores of 4, while explanation received a score of 3 and opportunities to speak and connections received scores of 2. Having provided readers with a sense of the discussions we observed along with two examples of how observers applied the discussion rubric, we turn now to a discussion of how we related discussion scores to students' performance on our written measure.

Quantitative methods

To analyze our data, we first explored the patterns of discussion features across classrooms by calculating a mean score and standard deviation for each feature. Next, we used HLM to compare the relationship between discussion and performance on the LMMA for students who were classified as ELLs and those who were not. Our goal was to ascertain whether discussion is as beneficial for ELLs as it is for non-ELLs. We also used HLM to determine how much of the between-class variation in performance for ELLs can be explained by math discussion overall and by each feature of discussion. HLM offers a way to uncover the relationship between class-level (discussion scores) and student-level (LMMA scores) variables while ensuring more credible statistical results than traditional regression modeling (Raudenbush & Bryk, 2002).

Table 2 displays the composite models specified for our analysis, beginning with limited models and adding parameters sequentially, keeping only those that proved significant (Hox, 2010). Model 1 is the null model, used to determine the percent of variation in LMMA scores between classes. Model 2 includes student-level variables, 2012 CST scores and ELL status, as fixed effects, meaning their effects were not allowed to vary across classes. Model 3 includes the class level variable, discussion scores, as fixed. In model 4, the effect of CST scores is allowed to vary randomly across classes, creating a better fitting model. Model 5 includes a two-way interaction between discussion and ELL

Table 2. HLM model specification.

Model	Composite Equation
1. Null Model	$LMMA_{ij} = \gamma_{00} + \mu_{0j} + e_{ij}$
2. Model with 2012CST and ELL Status fixed	$LMMA_{ij} = \left[\gamma_{00} + \gamma_{10}\left(2012CST_{ij)} + \gamma_{20}\left(ELL_{ij)}\right)\right] + \left[\ \mu_{0j} + e_{ij}\right]$
3. Model with 2012CST, Discussion, and ELL Status fixed	$LMMA_{ij} = \left[\gamma_{00} + \gamma_{10}\left(2012CST_{ij)} + \gamma_{01}\left(Discussion_j\right) + \gamma_{20}\left(ELL_{ij)}\right)\right] + \left[\ \mu_{0j} + e_{ij}\right]$
4. Model with random slopes for 2012CST	$LMMA_{ij} = \left[\gamma_{00} + \gamma_{10}\left(2012CST_{ij)} + \gamma_{01}\left(Discussion_j\right) + \gamma_{20}\left(ELL_{ij)}\right)\right] + \left[\ \mu_{0j} + \mu_j\left(2012CST_{ij}\right) + e_{ij}\right]$
5. Model allowing Discussion to interact with ELL Status	$LMMA_{ij} = \left[\gamma_{00} + \gamma_{10}\left(2012CST_{ij)} + \gamma_{01}\left(Discussion_j\right) + \gamma_{20}\left(ELL_{ij)}\right) + \gamma_{21}\left(EL_{ij}\right) + \gamma_{21}\left(EL_{ij} * Discussion_j\right)\right] + \left[\ \mu_{0j} + \mu_j\left(2012CST_{ij}\right) + e_{ij}\right]$
6–10. Models with individual features of discussion	$LMMA_{ij} = \left[\gamma_{00} + \gamma_{10}\left(2012CST_{ij)} + \gamma_{01}\left(Feature_j\right) + \gamma_{20}\left(ELL_{ij)}\right) + \gamma_{21}\left(ELL_{ij}\right)\right] + \left[\ \mu_{0j} + \mu_j\left(2012CST_{ij}\right) + e_{ij}\right]$

status, allowing us to determine if the relationship between discussion and LMMA performance is different for ELLs versus non-ELLs. Finally, models 6 through 10 each include one of the five features of discussion, allowing us to investigate the effects of individual features. We ran additional models with CELDT scores exploring effects for ELLs with varying English proficiency. However, because this analysis only included ELLs, the smaller sample size created insufficient power. We did not control for socioeconomic status because variation between schools was very low and the district did not offer access to these data at the student level.

Results

In this section, we first describe LMMA performance and the nature of discussion observed in the 20 classrooms. Next, we address the first research question reporting on findings of the HLM analysis.

LMMA performance

Third graders scored an average of 47.2% (SD 22.85) on the LMMA, and fourth graders scored an average of 51.78% (SD 21.85). This is lower than we expected, with students successfully solving only about half of the items. ELLs scored 48% on average, while non-ELLs scored slightly higher, with a mean of 51%.

Features of discussion in observed lessons

Table 3 illustrates the means for each discussion feature. Variety of approaches and equitable participation were the two highest scoring categories. Data indicate that for at least part of the observed lesson most students had access to more than one strategy or representation for a problem and most students demonstrated independent thinking at least once during the lesson in most classrooms. We found this promising given that in the United States, historically, most mathematics instruction consisted primarily of direct instruction with very little opportunity for discussion. Students in the majority of the 20 classrooms in this study were talking, publicly sharing their thinking, and solving problems more than one way. However, connections between ideas and conceptual explanations were observed much less frequently than the other features. These features correspond to the quality and content of talk during discussions and required teachers to respond in real time to students' contributions and connect them to other mathematics, which may be more difficult to achieve than the other features.

Effects of mathematical discussions on LMMA scores

We begin with the intraclass correlation (ICC) indicating that 12% of the variation in LMMA scores lies between classes (versus between individuals). It is this portion of the variation we sought to explain. To answer the first research question, we examined two variables, ELL Status and Discussion Total, as well as the presence of an interaction effect between these two variables. Table 4 displays the HLM models used to determine whether or not ELL status mediates the

Table 3. Discussion means across classrooms.

Discussion Feature	Mean Score	Standard Deviation	Range
Variety of approaches	3.60	.64	2–4
Equitable participation	3.55	.56	2–4
Opportunities to speak	3.05	.77	1–4
Connection between ideas	2.55	.84	1–4
Explanations	2.30	.86	1–4
Total Score	**15.10**	**2.69**	**10–20**

Table 4. HLM predictions of LMMA performance (Z Scores) by ELL status.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Fixed Effects										
2012CST		0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ELL Status		0.01	0.01	0.01	−0.49	0.01	0.01	0.01	0.02	0.01
		(0.07)	(0.07)	(0.07)	(0.40)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Total Discussion			0.03	0.04*	0.03					
			(0.02)	(0.02)	(0.02)					
ELL*Discussion					0.03					
					(0.03)					
Variety of Approaches						0.22**				
						(0.08)				
Opportunity to Speak							0.10			
							(0.08)			
Equitable Participation								0.21*		
								(0.10)		
Explanations									0.05	
									(0.08)	
Connection between Ideas										0.08
										(0.07)
Constant	−0.00	−3.12***	−3.44***	−3.80***	−3.59***	−3.89***	−3.41***	−3.89***	−3.23***	−3.32***
	(0.09)	(0.19)	(0.39)	(0.39)	(0.43)	(0.36)	(0.32)	(0.43)	(0.28)	(0.29)
Random Effects										
var(Residual)	0.90	0.51	0.50	0.48	0.48	0.48	0.48	0.48	0.48	0.48
var(Cons)	0.12	0.061	0.057	0.25	.25	0.21	0.27	0.30	0.26	0.29
var(CST)				0.00	0.00	0.00	0.00	0.00	0.00	0.00
cov(CST, cons)				0.00	0.00	0.00	0.00	0.00	0.00	0.00
Model Fit										
ICC	0.12									
AIC	1414.5	983.3	916.1	909.9	910.4	899.1	912.2	909.8	913.5	912.6
BIC	1427.2	1003.8	940.2	942.0	946.5	915.2	944.3	941.9	945.6	944.8
Deviance	1408.5	973.3	904.1	893.9	892.4	891.1	896.2	893.8	897.5	896.6

effect of discussion on students' performance on the LMMA. First, the statistically significant coefficient on Discussion Total in model 4 indicates the discussion score at the classroom level is significantly related to students' individual performance on the LMMA. Thus, discussion scores contribute to the explanation of between-class variation in LMMA performance above and beyond that explained by students' prior achievement. The percentage of variance explained was calculated by taking the difference in variance between models with and without the variable in question and dividing by the variance of the model without the variable. Thus, we determined discussion scores account for 6% of the between-class variance in LMMA scores, a small, yet promising effect.

Surprisingly, the interaction between ELL Status and Discussion Total in model 5 is not statistically significant ($p > .05$). The lack of a significant interaction effect suggests the relationship between discussion scores and performance on the LMMA is the same for ELLs and non-ELLs alike. This evidence supports the conclusion that ELLs in these classrooms benefit from class discussion as much as those who are native English speakers or designated fluent English proficient.

Two of the individual features of quality discussion explored in models 6 through 10 indicate a statistically significant relationship with the LMMA: variety of approaches and equitable participation. Opportunities to speak, explanations, and connections between ideas did not prove to be significantly related to achievement on the LMMA. Individually, variety of approaches explains 2.6% of the between-class variation in LMMA scores, while equitable participation explains 1.9%. It should also be noted that a non-significant coefficient when all five features are simultaneously included in the model suggests there may be some overlap in their effects.

To illustrate the relationship between discussion and performance we graphed each student on a scatter plot by their ELL status and their LMMA score. Next, we plotted the line of best fit for the group of students in classrooms that received discussion scores less than 15 and another line for the group of students in classrooms receiving discussion scores of 15 or greater. The slope of the lines in the graph in Figure 1 demonstrates the predicted correlation between students' ELL status and their LMMA score. Moreover, the parallel lines indicate discussions affected ELLs and non-ELLs similarly.

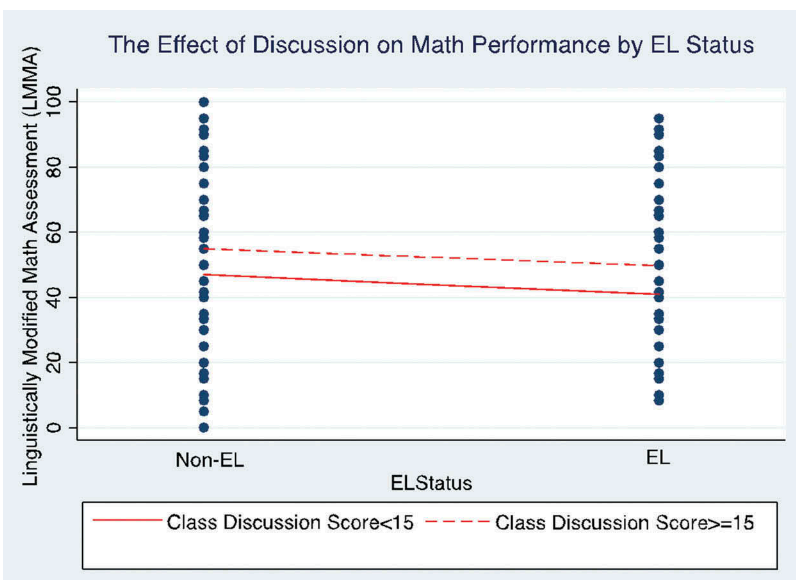

Figure 1. Lines of best fit relating class discussion score to LMMA score by ELL status.

Higher discussion scores appear to benefit both groups. The dashed line shows that the students in the 10 classrooms receiving discussion scores above 15 had higher scores on the LMMA than students in classes with lower scoring discussions.

Discussion

Our findings provide evidence to counter the idea that mathematics classroom discussions only benefit proficient English speakers. Our results show that students' status as ELLs did not mediate the relationship between class discussion and student performance on the LMMA. Therefore, in these twenty classrooms, mathematical discussion was equally beneficial for ELLs as it was for non-ELLs. Given that reform approaches to instruction are often underutilized in classes with high percentages of ELLs, this study adds to the evidence that English learners are capable of participating in high-level mathematics discussions and that even students with low English proficiency may benefit. With the current push for mathematical communication and reasoning in the Common Core State Standards, these findings come at a crucial time. Orchestrating mathematical discussions in classrooms with linguistically diverse learners is a demanding task and teachers will need substantial support from administrators and professional development to develop these skills.

Our study had two limitations. First, the inter-rater reliability of our discussion rubric was less than ideal (75% overall), especially for the features explanation and connections between ideas. Among the features with higher reliability are those found to be associated with performance, variety of approaches and equitable participation (80% and 85% respectively). We believe inter-rater reliability was easier to achieve for these features because, unlike explanations and connections, they include specific counts (number of ways of solving and number of students participating). Reliability is a common issue among studies attempting to measure complex classroom practices (Hill, Charalambous, & Kraft, 2012; Schoenfeld, 2013). However, our process of debriefing to reach agreement after scoring individually resulted in final consensus scores that mitigate issues with inter-rater reliability (Perry & Henry, 2004). Second, our measure of quality discussion may be associated with other aspects of quality instruction, including classroom management, relationship of the teacher with students, pedagogical content knowledge, that could also contribute to the positive association we found. Further, although teachers reported that observed lessons were representative of their daily teaching, it is possible that observed discussions differed from typical discussions in these classrooms. However, we believe teachers would find it very difficult, if not impossible, to orchestrate a high-quality discussion if they did not normally do so. In light of these limitations, findings and generalizations should be interpreted with some caution. Nonetheless, we believe our findings represent a step forward in understanding the impact of discussion on diverse students, a step on which future research can build.

Because our HLM analysis showed the discussion features variety of approaches and equitable participation significantly affected all students' performance, other educators may find it useful to develop these aspects of practice. We suggest the reason math discussions including these features may be especially beneficial to English learners is that they both involve visual displays and repetition of mathematical ideas. Mathematics instruction that includes a variety of different ways to solve problems opens up the discussion to learners who think about concepts or see problems in different ways. Equitable participation emphasizes hearing from many students across the classroom, including ELLs, and considers their non-verbal displays of thinking as important contributions. With each additional idea or approach presented, students hear the same math problem or concept discussed in different ways and they begin to make more sense out of it each time they hear it. Additionally, when students publicly share their ways of solving or representing a problem, they are positioned as important contributors to knowledge building (Takeuchi, 2015), and are learning to participate in multi-semiotic mathematical discourse (O'Halloran, 2005).

Moreover, the mid-scoring vignette demonstrated high scores in the two features found to correlate with students' performance, although it did not achieve a high overall score. We therefore argue that teachers may not need to have all the features mastered in order to see *some* benefits of discussion on measures of performance. In accord with theories of Hufferd-Ackles et al. (2004) and Kazemi and Stipek (2001), we believe working in other features of discussion, such as supporting students to produce conceptual explanations and connections between ideas, may be more difficult and take longer for teachers to develop. Because opportunities to speak, explanation, and connections between ideas did not demonstrate a significant relationship to performance, more research is needed to examine classroom practices related to these features. Taken as a whole, our analysis indicates mathematics classroom discussion can be an equitable practice. Discussion that includes student engagement and multiple solution strategies benefits all learners.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the 2013 NCLB Improving Teacher Quality Grant. California Department of Education. ITQ 10-704.

ORCID

Leslie C. Banes http://orcid.org/0000-0003-2352-3923

References

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, *33*, 4–14. doi:10.3102/0013189X033001004

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*, 219–234. doi:10.1207/S15324818AME1403_2

Adler, J. (1997). A participatory-inquiry approach and the mediation of mathematical knowledge in a multilingual classroom. *Educational Studies in Mathematics*, *33*, 235–258. doi:10.1023/a:1002976114883

Akkus, R., & Hand, B. (2010). Examining teacher's struggles as they attempt to implement dialogical interaction as part of promoting mathematical reasoning within their classrooms. *International Journal of Science and Mathematics Education*, *9*, 975–998. doi:10.1007/s10763-010-9266-8

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, *93*, 373–397. doi:10.1086/461730

Barwell, R. (2012). Discursive demands and equity in second language mathematics classrooms. In B. Herbel-Eisenmann, J. Choppin, D. Wagner, & D. Pimm (Eds.), *Equity in discourse for mathematics education: Theories, practices, and policies* (pp. 147–163). Dordrecht, The Netherlands: Springer.

Baxter, J. A., Woodward, J., & Olson, D. (2001). Effects of reform-based mathematics instruction on low achievers in five third-grade classrooms. *The Elementary School Journal*, *101*, 529–547. doi:10.1086/499686

Boaler, J. (2000). Mathematics from another world: Traditional communities and the alienation of learners. *The Journal of Mathematical Behavior*, *18*, 379–397. doi:10.1016/S0732-3123(00)00026-2

California Department of Education (2017, April 14). California English Language Development Test (CELDT). Retrieved from http://www.cde.ca.gov/ta/tg/el/

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

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann. doi:10.1017/S0047404500014676

Chapin, S., O'Connor, C., & Anderson, N. (2003). *Classroom discussions. Using math talk to help students learn, Grades 1-6*. Sausalito, CA: Math Solutions.

Chi, M. T. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in Instructional Psychology*, *5*, 161–238. Retrieved from http://chilab.asu.edu/papers/advances.pdf

Cirillo, M. (2013). *NCTM research brief: What does research say the benefits of discussion in mathematics class are? (703)* (1–6). Reston, VA: National Council of Teachers of Mathematics. doi:10.1111/jpm.12012

Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading*, *8*, 305–336. doi:10.1.1.474.1163.
Echevarria, J., Vogt, M. E., & Short, D. J. (2007). *Making content comprehensible for English language learners: The SIOP model* (3rd ed.). Boston, MA: Allyn & Bacon.
Ellis, R. (1999). *Learning a second language through interaction*. Amsterdam, The Netherlands: John Benjamins. doi:10.1075/sibil.1
Forman, E. A., McCormick, D., & Donato, R. (1998). Learning what counts as a mathematical explanation. *Linguistics and Education*, *9*(4), 313–339. doi:10.1016/S0898-5898(97)90004-8
Gorgorió, N., & Planas, N. (2001). Teaching mathematics in multilingual classrooms. *Educational Studies in Mathematics*, *47*, 7–33. doi:10.1023/A:1017980828943
Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly*, *19*(1), 173–180. doi:10.1016/j.ecresq.2004.01.012
Gutiérrez, R. (2007). Context matters: Equity, success, and the future of mathematics education. In *Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1–18). Stateline: University of Nevada, Reno. Retrieved from http://www.pmena.org/pmenaproceedings/PMENA%2029%202007%20Proceedings.pdf
Hansen-Thomas, H. (2009). Reform-oriented mathematics in three 6th grade classes: How teachers draw in ELLs to academic discourse. *Journal of Language, Identity, and Education*, *8*, 88–106. doi:10.1080/1534845090284841
Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In A. Grouws & Douglas (Ed), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 65–97). New York, NY: Macmillan.
Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, *41*, 56–64. doi:10.3102/0013189x12437203
Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.
Hufferd-Ackles, K., Fuson, K. C., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education*, *35*, 81–116. doi:10.2307/30034933
Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper elementary mathematics classrooms. *The Elementary School Journal*, *102*, 59–80. doi:10.1086/499693
Khisty, L. L., & Chval, K. B. (2002). Pedagogic discourse and equity in mathematics: When teachers' talk matters. *Mathematics Education Research Journal*, *14*, 154–168. doi:10.1007/BF03217360
Kosko, K. W., & Miyazaki, Y. (2012). The effect of student discussion frequency on fifth-grade students' mathematics achievement in U.S. schools. *The Journal of Experimental Education*, *80*, 173–195. doi:10.1177/0895904815595723
Lane, S., Silver, E. A., & Wang, N. (1995, April). An examination of performance gains of culturally and linguistically diverse students on a mathematics performance assessment within the QUASAR project. Paper presented at the AERA annual meeting, San Francisco, CA. Retrieved from https://eric.ed.gov/?id=ED390927
Martiniello, M. (2008). Language and the performance of English-Language Learners in math word problems. *Harvard Educational Review*, *78*, 333–369. doi:10.17763/haer.78.2.70783570r1111t32
Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, *104*, 1–21. doi:10.1016/j.jecp.2008.08.004
Mehan, H. (1979). "What time is it, Denise?": Asking known information questions in classroom discourse. *Theory into Practice*, *18*, 285–294. doi:10.1080/00405847909542846
Morgan, P. L., Farkas, G., & Maczuga, S. (2015). Which instructional practices most help first-grade students with and without mathematics difficulties? *Educational Evaluation and Policy Analysis*, *37*, 184–205. doi:10.3102/0162373714536608
Moschkovich, J. (2007). Using two languages when learning mathematics. *Educational Studies in Mathematics*, *64*, 121–144. doi:10.1007/s10649-005-9005-1
Moschkovich, J. (2013). Principles and guidelines for equitable mathematics teaching practices and materials for English language learners. *Journal of Urban Mathematics Education*, *6*, 45–57. Retrieved from http://education.gsu.edu/JUME
National Council for Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
Niemi, D. (1996). Assessing conceptual understanding in mathematics: Representations, problem solutions, justifications, and explanations. *Journal of Educational Research*, *89*, 351–363. doi:10.1080/00220671.1996.9941339
O'Halloran, K. 2005. *Mathematical discourse: Language, symbolism and visual images*. London: Continuum.
Perry, J. C., & Henry, M. (2004). Studying defense mechanisms in psychotherapy using the Defense Mechanism Rating Scales. In U. Hentschel, G. Smith, J. G. Draguns, & W. Ehlers (Eds.), *Defense mechanisms: Theoretical, research and clinical perspectives* (Vol. 136). Amsterdam, The Netherlands: Elsevier.
Planas, N., & Gorgorió, N. (2004). Are different students expected to learn norms differently in the mathematics classroom? *Mathematics Education Research Journal*, 16(1), 19–40.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed. ed.). Thousand Oaks, CA: Sage.

Rittle-Johnson, B., Saylor, M., & Swygert, K. E. (2008). Learning from explaining: Does it matter if mom is listening? *Journal of Experimental Child Psychology*, *100*, 215–224. doi:10.1016/j.jecp.2007.10.002

Schoenfeld, A. H. (2013). Classroom observations in theory and practice. *ZDM Mathematics Education*, *45*, 607–621. doi:10.1007/s11858-012-0483-1

Secada, W. G., & De La Cruz, Y. (1996). Teaching mathematics for understanding to bilingual students. In J. L. Flores (Ed.), *Children of la Frontera: Binational efforts to serve Mexican migrant and immigrant students* (pp. 286–308). Charleston, WV: Appalachia Educational Laboratory.

Smith, M., & Stein, M. K. (2011). *5 practices for orchestrating productive mathematics discussions*. Reston, VA: NCTM.

Spillane, J. (2001). Challenging instruction for all students: Policy, practitioners, and practice. In S. Fuhrman (Ed.), *From the capital to the classroom: Standards based reform in the states. 100th yearbook of the National Society for the Study of Education* (pp. 217–241). Chicago, IL: University of Chicago Press.

Stigler, J. W., & Hiebert, J. (1998). The TIMSS videotape study. *American Educator*, *22*(4), 7, 43–45.

Takeuchi, M. (2015). The situated multiliteracies approach to classroom participation: English language learners' participation in classroom mathematics practices. *Journal of Language, Identity and Education*, *14*, 159–178. doi:10.1080/15348458.2015.1041341

Truxaw, M. P., & DeFranco, T. (2008). Mapping mathematics classroom discourse and its implications for models of teaching. *Journal for Research in Mathematics Education*, *39*, 489–525. doi:10.2307/748406

Turner, E., Dominguez, H., Maldonado, L., & Empson, S. (2013). English learners' participation in mathematical discussion: Shifting positionings and dynamic identities. *Journal for Research in Mathematics Education*, *44*, 199–234. doi:10.5951/jresematheduc.44.1.0199

Varonis, E. M., & Gass, S. M. (1985). Miscommunication in native/nonnative conversations. *Language in Society*, *14*, 327–343. doi:10.1017/S0047404500011295

Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, *78*, 516–551. doi:10.3102/0034654308320292

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*, 366–389. doi:10.2307/749186

Appendix A

Classroom Discussion Observation Instrument

	0	1	2	3	4
Variety of approaches and resources	Emphasis on symbol mani-pulation	Emphasis on symbol mani-pulation	Multiple resources used for one approach throughout lesson	Multiple approaches for part of lesson	More than 3 approaches on part of lesson and/or 2–3 throughout
Explanations	No explanation at all	Procedural narration or non-interactive video explanation	Partial explanations OR interactive video OR Meaningful and thorough explanations available only to some students	One meaningful and thorough explanation (excluding video) available to all students	Two or more meaningful and thorough explanations available to all students (excluding video) in whole group discussion
Opportunities to speak	Teacher is the only speaker	Students speak infrequently and only briefly in response to teacher prompts	Students speak infrequently but at length OR Students speak frequently but briefly	Students speak frequently and occasionally at length	Students speak frequently and consistently at length
Equitable participation in whole group interaction	No whole group discussion takes place	1 –3 students participate in discussion	4 –6 students dominate the discussion; most students do not speak or visibly engage	More than 6 students participate (speak or visibly engage) in discussion	Most students, including ELLs, participate (speak or visibly engage) in discussion. Independent thinking is publicly shared.
Connections between ideas	No whole group discussion takes place	Focus is on sharing of procedures	*Teacher* probes strategies OR Little connection made between ideas	*Teacher* makes substantial connections OR *Students* comment on conceptual aspects of others' ideas	*Students* make substantial connections between central ideas

Resources – tools used to complete task: manipulatives, drawings, symbols, number line, graphs, etc.

Approaches – strategies: procedures/steps to complete the problem

Interactive Visual Learning Bridge –teacher stops digital animation (that accompanies curriculum) to have children respond to questions posed by the narrator

Thorough explanation – needs to include justification. Relate to concept, not just procedure. Include appropriate mathematics terminology, emphasis on meaning, and should not be missing an essential component

Partial explanation – must have some conceptual aspect, not just a procedural narration. Might be missing some terminology or an essential component

Connections – might involve comparing and contrasting two or more ideas/strategies/concepts/representations