This is a post-print version of the article:

Chan, M. (2020). A multilevel SEM study of classroom talk on cooperative learning and academic achievement: Does cooperative scaffolding matter? *International Journal of Educational Research, 101*, 101564. https://doi.org/10.1016/j.ijer.2020.101564

Publisher's version available at:
https://www.sciencedirect.com/science/article/pii/S088303551932275X?via%3Dihub

A multilevel SEM study of classroom talk on cooperative learning and academic achievement: Does cooperative scaffolding matter?

1. Introduction

It is now increasingly apparent that constructive forms of classroom talk exert a beneficial influence on students' disciplinary understanding and socio-cognitive advancements (Gilles, 2015; Mercer & Littleton, 2007; Walshaw & Anthony, 2008). However, not all kinds of talk are equivalent. Scholars of contemporary classroom discourse continue to debate over a critical conundrum surfaced by Barnes (1976) and revisited by Simpson, Mercer and Majors (2010): "If learning, particularly that which takes place in the classroom, floats on a sea of talk, what kind of talk? And what kind of learning?" (p. 1).

Despite considerable research on classroom talk (e.g., Gillies, 2014; 2015; Mercer & Dawes, 2014), and in spite of numerous efforts by leading educational scholars to theorize the value and importance of exploratory talk (Alexander, 2008; Barnes, 2008; Mercer, 1992; Mercer & Dawes, 2008), classrooms continue to be led by teacher-dominated discussions using closed lines of questioning (Lefstein & Snell, 2010; Mercer & Dawes, 2014; Nystrand et al., 2003) and revolve around teacher IRF sequences (initiation–response–feedback; Howe & Abedin, 2013). While educators are broadly aware of the specific pedagogical scripts that make talk productive for learning, in practice, they are less common than it should be (e.g., Alexander, 2008; Mercer & Howe, 2012). Rather, one kind of talk dominates contemporary classroom discourse: "the so-called 'recitation script' of closed teacher questions, brief recall answers and minimal feedback that requires children to report someone else's thinking rather than to think for themselves, and to be judged on their accuracy or compliance in doing so" (Alexander, 2008, p. 93). Similarly, at the turn of the 21st Century, Smith, Hardman, Wall and Mroz (2004, p. 410) also observed that "In the whole class sections of literacy and numeracy lessons most of the questions asked were of a low cognitive level designed to funnel students' responses towards a required answer”.

Although research on classroom discourse is well-researched, much of existing studies have relied on qualitative classroom observations. When quantitative evidence are reported, these are often limited to descriptive enumeration of specific classroom activities (see Howe & Abedin, 2013 for a review). For instance, these studies typically entail investigations of background and contextual characteristics on variations in frequencies of qualitative dialogue categories. After decades of classroom research, it remains unclear whether some modes of dialogic organization are more beneficial and for which outcomes (i.e., the consequences of dialogue), what are the relational functions of different dialogue practices (i.e., the processes of dialogue) and how large these effects are (i.e., the impact of dialogue) (Howe & Abedin, 2013; Mercer & Howe, 2012; Wolf, Crosson, & Resnick, 2006). To date, only one recent quantitative research attended to this empirical dilemma in which Howe et al. (2019) used two-level multilevel modeling to analyze survey, classroom recordings and assessment data from 72 Primary school classrooms in England. Their comprehensive study examined different classroom talk patterns and their influence on student academic and attitudinal outcomes, while also controlling for potential confounds (e.g., prior attainment, socio-economic status and classroom climate). They found that active student participation coupled with teacher elaboration and querying prior learning positively predicted student outcomes.

To contribute to this emerging field, the current study draws on quantiative survey and assessment data from Singapore Secondary schools to examine the structure of classroom talk in mathematics and English classrooms. Although multilevel modeling was similarly employed in this study, our conceptual and analytical foci were different. Unlike Howe et al. (2019) whose conceptions of talk categories were largely informed by the authors’ previous work and related classroom discourse research, key facets of classroom talk in this study were motivated by two dominant conceptions that have received broad reception in the literature: *presentational* or *accountable* talk (i.e., talk that is typically closed ended and used for performative or evaluative purposes) and *exploratory* talk (i.e., talk that is typically open-ended and promotes deep learning and involves co-construction of meaning) (Barnes, 2008; Mercer & Littleton, 2007; Resnick, Michaels, & O'Connor, 2011). Next, to accommodate the nested nature of students within classrooms and to concurrently investigate a priori structural relationships, multilevel mediation modeling was employed to examine: 1) the relations of classroom talk on academic achievement and students' orientations towards cooperative work, 2) the extent to which teacher scaffolds of cooperative learning mediated the relationship between classroom talk and student outcomes (since productive talk depends on teachers establishing the right climate; Mercer & Howe, 2012) and 3) the influence of individual and compositional (classroom) effects of student prior ability (Wilkinson & Fung, 2002; Willms, 2010). While these findings may be unique to the educational setting of Singapore, in the context of 21st Century schooling, however, there is every reason to believe that productive dialogue is an important pedagogical endeavour that permeates all educational systems.

2. Literature review

2.1. What is classroom talk?

Drawing on a constructivist view of learning, Barnes (1976, 2008) described exploratory talk as productive academic conversations that entail a set of carefully crafted teacher questions used purposively to engage students to work on understanding. But it may well be Mercer (2000) who provided one of the clearest accounts:

> Exploratory talk is that in which partners engage critically but constructively with each other's ideas. Relevant information is offered for joint consideration. Proposals

> may be challenged and counterchallenged, but if so reasons are given and alternatives are offered. Agreement is sought as a basis for joint progress. Knowledge is made publicly accountable and reasoning is visible in the talk. (p. 98)

For Barnes, Mercer and others (e.g., Cazden, 2001; Mercer & Dawes, 2008; Mercer & Littleton, 2007), exploratory talk generates the most positive learning outcomes. Exploratory talk is not simply a conduit for the transmission of knowledge, of the teacher talking to (and imparting knowledge to) students. Rather, it is an intentional classroom activity that, on one hand, develops students' social, linguistic and conceptual skills through critical but constructive assessment conversations between teachers and students and among students (Alexander, 2008; Mercer & Hodgkinson, 2008), and, on another, provides teachers with relevant assessment information to adjust instruction to optimize students' learning. Duschl and Gitomer (1997) describe these forms of purposive dialogue as assessment conversations which they explain are "a specially formatted instructional dialogue that embeds assessment into the activity structure of the classroom" (p. 39). When teachers use exploratory talk, learners are not only afforded opportunities to engage in reasoned augmentation (e.g., explaining and justifying their responses), but students also become more confident and self-regulated as they probe deeper into the topic and recraft their responses after considering alternative propositions (e.g., Wells & Arauz, 2006). Therefore, for talk to be considered exploratory, students' ideas need to be constructively challenged by inviting critique and getting them to clarify, to explain or to convince others' of their solutions, even if they have provided the right answer.

Research however suggests that exploratory talk is arguably a benign and less dominant representation of contemporary teacher-student dialogue. As early as the 1970s, Barnes (1976) observed the emergence of a second kind of "presentational" talk that contrasted directly with exploratory talk. In contrast to process-focused dialogue, the focus of presentational talk (for teachers) is a summative product- or performance-focused evaluation of whether students knew the right answer or how well they understood what had already been taught. In other words, when teachers use presentational talk, the instructional priority is in having students reproduce, orally, what and how much they have mastered a given topic, against an assessment criteria or performance standard determined by the teacher. Attempts to operationalize presentational talk often take the form of a monolingual three-part recitation sequence (e.g., Barnes, 2008; Mercer & Dawes, 2008; Sinclair & Coulthard, 1975): teacher *initiates* question, student *responds* and teacher *evaluates*. Although Sinclair and Coulthard (1975) and subsequently Barnes (1976) and Mehan (1979) were among the first to characterize the pervasiveness of the performative dialogue, in which the teacher prioritizes knowing the 'correct answer' over critical thinking, it remains omnipresent in contemporary classrooms. For instance, in the United Kingdom, several decades of classroom research have shown that the structure and quality of instructional dialogue remains relatively unaltered: "teachers dominate classroom interaction, talking most of the time, controlling topics and allocation of turns … [while] pupils talk much less than the teacher, for shorter durations and in most cases only in response to teacher prompts" (Lefstein & Snell, 2010, p. 2). Similarly, in the United States, Nystrand et al. (2003) found that teachers asked questions that fostered recitation of factual or previously acquired knowledge, and interactions with students primarily involved the use of highly codified closed-ended questions to promote procedural fluency. The authors concluded that the dominant profile of classroom discourse could be characterized as 1) little more than procedural reciprocity, 2) weak or almost absent coherence across subject topics, and 3) persistent questioning that involved factual recall on information previously taught and already known (see p. 138-139). Therefore, much of conventional classroom discourse appear to be located in two distinct facets: Performative talk (which focuses on the use of closed questioning to demonstrate, transmit and reproduce previously learned knowledge) and procedural talk (which focuses on the processes, rules, scripts or heuristics involved in solving a specific problem or understanding specific knowledge).

To explain the inconsistent theory-practice nexus, some scholars assert that performative (and procedural) talk may not always be hostile to working on understanding. Wells and Arauz (2006) point out that both monologic and dialogic talk may have their practical applications in the classroom. In later publications, Barnes (2008), despite earlier reservations, also acknowledged the utility of performative talk for assessment purposes. However, in spite of their neutrality on performative talk, they make clear that “learning is likely to be more effective when students are actively engaged in the dialogic co-construction of meaning about topics that are of significance to them” (p. 379). Therefore, process-oriented talk remains a crucial centerpiece of active learning and high-quality classroom conversations. For Wells and Arauz (2006), and others (e.g., Rojas-Drummond, et al., 2013), the dialogical transition from performative to exploratory talk is not simply using more open-ended questions. The articulation and expansion of students’ ideas through shared dialogue need to involve a range of collective, reciprocal, cumulative and purposeful strategies orchestrated by the teacher (Alexander, 2008). For instance, while classroom talk can be initiated with performative (e.g., what is the right answer?) and process (e.g., how did you get that?) probes, teachers should continue to scaffold the dialogue by following up with exploratory questions and interpretive prompts, such as asking students to *clarify* their answers (e.g., Chapin, O’Connor, & Anderson, 2009), to *explain* or justify the reasoning, meaning and how they arrived at their answers (e.g., Chi et al., 1994), and/or to make *connections* among different ideas that generate additional information and construct new knowledge (e.g., Krathwohl, 2002). In a recent review, Gillies (2015) noted that the types of exploratory talk that consistently represented the hallmarks of dialogic interactions typically involved the following sequences: *clarifying questions* "to challenge students' understanding and inconsistencies in their thinking", *explaining questions* "to provide reasons for their responses… to consider alternative explanations" and *connecting questions* that encourage students to "connect different lines of thought into cogent explanations" (p. 39). Subsequently, in this study, these concepts were used in the development of constructs underlying exploratory talk.

2.2. Scaffolding cooperative learning

What we know thus far about productive talk is, it is not simply a series of open-ended teacher questions nor is it about getting students to sit together, working on individual tasks with minimal interaction. Similarly, classroom dialogue without purposeful scaffolding is equally unproductive (Bakker, Smit, & Wegerif, 2015; Lin et al., 2015). Over the past two decades or so, a consistent strand of research has documented that cooperative learning is an effective instructional activity that promotes active student interactions, engagement and learning. It fosters socialization skills but also encourages cooperative rather than performative pursuits in learning (Johnson & Johnson, 2009; Slavin, 2014). When students work together, they explore and discover information together, clarify misconceptions and engage in co-regulated learning routines that promote peer mentorship, confidence and trust. Indeed, studies have also shown that the experiences gained from cooperative learning help students achieve better academic outcomes (Hattie, 2015), manage interpersonal conflicts (e.g., Johnson et al., 1997), hone communication skills and foster stronger orientations towards cooperative work (Tarim, 2009) – all essential traits of 21st Century outcomes. But unless teachers create a conducive and cooperative classroom culture for co-constructive learning, productive teacher-student and student-student interactions are unlikely to happen (Bruner, 1983; Gillies, 2014 Lefstein & Snell, 2014). Additionally, teachers also need to teach students how to engage in constructive dialogue.

Commenting on the importance of the teacher's role in fostering productive student interaction, Gillies (2017) argues that teachers must first guide students to work cooperatively. Only then will students feel sufficiently confident to express and solicit ideas, negotiate alternative propositions and reason cogently among themselves. Similarly, professional development authors of productive talk moves also echo the same message that, before proceeding with cooperative work, teachers ought to establish ground rules for active and respectful participation and communicate the activity goals clearly to students (Chapin et al., 2009). Otherwise, students may not know what is expected, will be reluctant to talk freely or may perceive the activity as a waste of time (Gillies, 2014; Rojas-Drummond et al., 2006; Wells & Arauz, 2006). Subsequently, discord among group members will set in, students will be confused, feel frustrated with the demands of the tasks and contributions among group members will be uneven (Johnson & Johnson, 2009). To avoid these pitfalls, Johnson and Johnson advocate five essential scaffolds for effective cooperative learning: promote positive interdependence, encourage individual accountability, promote interaction among members, build social skills and support students in processing information collectively.

There are strong evidence in support of teacher scaffold of cooperative learning and productive talk. For instance, while empirical studies have consistently revealed that a dialogic classroom promotes better reasoning and cognitive skills, its effects were also most pronounced when students are afforded opportunities to work collaboratively supported by appropriate teacher scaffolds (Howe et al., 2019; Lin et al., 2015; Slavin, 2013; Van Boxtel, Van der Linden & Kanselaar., 2000; see Gilles, 2014 for an extensive review). Through cooperative group work, students learn to share and listen to ideas, clarify misconceptions when they arise and become more willing to provide extended explanations that facilitate deeper understanding and continued engagement in the subject-matter – all hallmarks of academically productive discussions. Teachers who provide effective scaffolds to support students' co-constructive dialogue are more likely to engage students' in talk that is both productive and creative (Mercer & Littleton, 2007; Resnick et al., 2011). Put simply, when classroom talk is properly scaffolded to maximize learning, students are not only guided to "think aloud" (Barnes, 2008), but they also learn how to "think together" (Mercer & Littleton, 2007).

Recent studies continue to support these claims. Webb et al. (2014) found that teachers who interacted dialogically with students through collaborative group work not only witnessed higher levels of engagement, these students were also observed to be more receptive to new ideas as they built upon each other's ideas to generate new mathematical understandings. Although a positive link between exploratory talk and mathematics achievement was observed, the authors concluded that it was the quality of cooperative learning experiences (i.e., work together, sharing ideas) that really mattered. In a recent large-scale quantitative study that investigated teacher-student dialogue in whole-class teaching, Howe et al. (2019) reported that cooperative group work predicted the frequency of teacher elaboration and student participation in exploratory talk (i.e., expressing ideas and engaging with the ideas of others).

Despite the supportive evidence presented and the various propositions related to the teacher's role in directing productive classroom talk, however, scholars remain concerned about the state of productive talk in classrooms around the world, but more so in the West. Littleton and Howe (2010, p. 3) (also cited in Nind, Curtin, & Hall, 2016) content that "there is certainly a lot of talk going on in classrooms, but seemingly to little effect" (p. 3) and "… educational dialogue is typically less effective than it ought to be" (p. 10). Subsequently, they concluded that "much of the talk in collaborative activity in classrooms is unproductive" (p. 271). Perhaps as Howe and Abedin (2013) point out, the lack of progress stems from the limitations of qualitative research in responding to target-based assessments of dialogic practices, for example, whether there are optimal modes of talk with respect to different outcomes and student characteristics, whether there are conditional and sequential processes of talk practices (e.g., presentational-to-exploratory or exploratory-presentational, and do they depend on cooperative scaffolding), and whether student or contextual factors may influence the participation in and quality of classroom dialogue (Howe, 2010; Howe et al., 2019). Research in these areas could be pursued using advanced quantitative methods (Howe & Abedin, 2013) and should be prioritized (Bokhove, 2018).

2.3. Classroom talk in Singapore classrooms

Similar to the West, previous studies on classroom talk in Singapore have been primarily qualitative and relatively consistent with existing literature. For instance, in Singapore, teacher talk dominates classroom interaction (Kogut & Silver, 2009). Out of 345 activities in 28 lessons, talk related to classroom management and instruction were most frequent (over 65%), followed by teacher correction/answer checking (~35%), teacher exposition (~28%) and teacher questioning (~20%). While the last two categories may be deemed dialogic practices, the authors countered that these "teacher-led activities tend to be implemented in a traditional and didactic manner with limited opportunity for extended student response: Teacher questioning, Teacher exposition, and Teacher correction/Answer checking all tend to elicit short student responses…" (p. 14). Another large-scale qualitative study of 625 mathematics and English lessons extend the same compelling narrative that the 'recitation script' of short, closed-ended teacher questions continues to dominate (Hogan, Kwek et al., 2013). Coding a 3-minute sequence in a lesson as a phase (i.e., a 30-minute lesson would have 10 phases), the authors reported that in mathematics and English classrooms, teachers were more likely to ask closed rather than open questions in whole class contexts (68% versus 4% and 58% versus 14%, respectively). Extended student responses were a rarity (6% and 2%, respectively). The authors concluded that there was little evidence of exploratory talk or dialogical exchanges in Singapore classrooms.

3. The present study

Based on the literature reviewed, a key objective of this study is to investigate, using multilevel quantitative methods, the influence of cooperative scaffolding on different modes of classroom talk, and subsequently how it contributes additively to learning outcomes related to cooperative orientations (as a key 21st Century disposition) and academic achievement (Gillies, 2016; Littleton & Howe, 2010).

Three research questions are pursued. First, what is the structure of talk in Singapore mathematics and English classrooms? To guide our analysis, we draw on the discussion-based recitation sequence in mathematics (e.g., Chapin et al., 2009) and general subjects (e.g., Gilles, 2014). These models propose that the prototypical classroom dialogue should typically begin with performative probes, followed by exploratory prompts to elicit student-generated responses. In addition, we will also explore alternative models of talk. As Howe and Abedin (2013) point out, the ability to test alternative models is an important contribution towards target-based assessments.

Next, extending the validated model, our second question was: What influence do different facets of classroom talk have on student outcomes, specifically, student academic achievement and cooperative learning dispositions? In line with supporting evidence that teacher questions with predetermined answers dominate classroom interactions, we expect performative talk to predict better achievement outcomes. Although exploratory forms of talk may predict academic achievement (Mercer & Sams, 2006), some scholars suggest that the talk-achievement relationship may be mediated through cooperative learning dispositions (Slavin, 2013).

Third, a visible omission in current classroom discourse research is the consideration of student and contextual factors, with the exception of one recent study (Howe et al., 2019). In quantitative classroom effects research, however, compositional classroom characteristics of students are particularly meaningful (Harker & Tymms, 2004; Willms, 2010), and more so among educational systems with academic track placements such as Singapore. The importance of prior attainment was highlighted in Howe et al.'s (2019) study where they found that among the range of control variables examined, prior attainment featured prominently across several dialogue variables. Therefore, our third question was: How much do individual and contextual classroom factors influence the relationship between classroom talk and student outcomes? For the final analysis, students' prior attainment was included as an explanatory covariate to determine whether the influence of student and class-average prior attainment (i.e., the compositional effect) on the different modes of classroom talk. Arising from the literature review, a second contextual variable – *cooperative scaffolding* – was considered for this analysis to investigate the role of the teacher in promoting different modes of classroom talk in the context of cooperative learning (Gillies, 2014).

4. Method

4.1. Participants and research design

Data was drawn from a large-scale study that examined pedagogy and assessment practices in Singapore classrooms. Using a two-stage stratified design, the population of Secondary schools in Singapore was rank-ordered by academic standing and divided into three proportional clusters. Subsequently, a random sample of 30 schools was selected (10 from each cluster), representing 18% of the total population.

Within each school, all Secondary Three classes (students aged 15 years) were involved. This approach allowed us to make representative inferences that cover a wide spectrum of achievement levels as well as student and classroom characteristics. Within each class (*n*=30 – 35), a split-half approach (also known as the multi-matrix design; Munger & Loyd, 1988) was used in which approximately half were randomly assigned a mathematics test and a survey on teaching and learning in the same subject, while the other half were assigned English. This approach (random assignment) helped protect against selection bias. The final sample was 119 classrooms, of which 1166 (50.3% males) and 1027 (49.7% males) students completed mathematics and English, respectively.

4.2. Procedure

In this study, measures of whole-class instruction were rated by students. This is common and reliable approach in instructional effects literature (e.g., Fauth et al., 2014) as well as international assessments such as PISA. For these instructional variables, the referent is at the classroom where the response stem begins with "In your English class…", or "How often does your English teacher…". Using multilevel modeling nomenclature, these variables are treated as Level 2 (L2) variables which represent the shared agreement among students within the same classroom, rating the same teacher. On the other hand, student level variables (e.g., student background, dispositions, academic achievement) are known as Level 1 (L1) variables. In multilevel educational literature, L1 background variables, when aggregated to the classroom, represent L2 compositional variables (Opdenakker et al., 2002) and are conceptually distinct from the L1 measures (Morin, Marsh, Nagengast, & Scalas, 2014). These L2 variables examine the extent to which student outcomes can be explained by differences in classroom composition related to particular L1 student variables. For example, when student prior attainment is aggregated to the classroom, the new L2 variable now reflects class-average prior attainment.

Multilevel SEM (structural equation modeling) was then used to simultaneously examine the effects of all L1 (student) and L2 (classroom and compositional) variables on L1 outcomes (e.g., achievement). For L1 variables (prior attainment and achievement) that are modelled at L2, we utilized a latent aggregation approach that attends more reliably to measurement and sampling errors (Lüdtke et al., 2008). This is a preferred approach as opposed to the conventional aggregation where student responses are simply mean aggregated to the classroom level. Because samples across classrooms are often unbalanced, the mean aggregation approach would unfairly attribute the same weight to all classes, regardless of class size differences, thereby leading to estimation biases and reduced statistical precision. To facilitate discussion of our results (sub-sections 6 and 6.1), multilevel effect sizes are reported following the guidelines by Morin et al. (2014).

4.3. Measures

Academic achievement. The mathematics test consisted 28 questions that assessed topics in numbers, algebra, measurement, geometry and statistics, while the English test contained 70 questions covering reading comprehension, grammar and vocabulary. Prior to administration, instruments were piloted for empirical validity and face-validated for content validity by former teachers. A two-parameter item response analysis was performed to compute a standardized scaled score for each student. Reliabilities were strong at .84 and .92 for mathematics and English, respectively, and intraclass correlations (ICCs) ranged between .50 and .60.

Cooperative learning disposition. This variable consisted of seven items that assessed student's orientation towards cooperative learning. A higher-order factor comprising two subsidiary factors (group engagement and teamwork) was supported with excellent fit to the data (χ^2 = 16.44 (12), p=.17; CFI = .999; RMSEA = .017). Using the response stem "In my math/English" and response options "Almost Never to Almost Always", sample items for *group engagement* were: "I share my ideas during group work"; "I get involved in class discussions". Sample items for the *teamwork* were: "I work well with others"; "I get along with people who have different opinions". The response stem was "In your math/English class, how easy or difficult is it for you to…" and response options for this variable was "Very Difficult to Very Easy". Construct reliabilities were strong at .90 and .92 across both subjects. ICCs were small for mathematics (.014) but moderate for English (.114). To construct the L2 component of this construct, the same latent aggregation approach was applied (similar to prior attainment and achievement). At L2, this aggregated construct reflects how well learners in the same class relate to one another as well as their collective willingness to work and cooperate on common goals. In other words, it reflects the overall *cooperative learning climate* of the classroom. This is distinct from the L1 cooperative learning disposition which relates to the individual student's own assessment of him or herself with respect to working with and engaging others.

Cooperative scaffolding. Five student-rated items were used to assess teachers' active facilitation of cooperative learning in mathematics and English classrooms. Confirmatory factor analysis (at the individual level) indicated a good fit (mathematics: χ^2 = 3.55 (4), p=.471; CFI/TLI = 1.00/1.00); English: χ^2 = 19.0 (4), p=.001; CFI = .992/.981). Sample items were "My Math/English teacher … encourages us to discuss issues or problems in small groups; …provides lots of guidance when we work together in groups; ...encourages us to take on different roles in small group work". Construct reliabilities were strong at .85 across both subjects. Subsequently, group-mean aggregation was applied to create the L2 variable.

Classroom talk. Based on a 5-point scale (*almost never* to *almost always*), twelve items were developed to assess students' exposure to five facets of teacher questions, with slight variations to accommodate linguistic norms in mathematics and English (indicated by superscripts ^{m} and ^{e}). Two-items measured *performative talk*: "Is this answer right or wrong?" and "What is the correct answer to this ^{m}problem/^{e}question?"; *procedural talk*: "What's the best way to ^{m}solve/^{e}address this problem?" and "^{m}What is the next logical step to solve this problem/^{e}What's the most effective way of approaching this particular task?"; and *clarifying* *talk*: “What do you mean?” and “Could you explain what you mean?” Reliability estimates using the Spearman-Brown formula (for two-item measures) were [m].72/[e].72, [m].67/.[e]75, and [m].71/[e].74, respectively. Four items measured *explanatory talk*: “How did you come to that [m]answer/[e]conclusion?”, “Can you give me reasons for why you think that?”, “[m]How would you know whether that is true or not/[e]What reasons can you give to support your statement?” and “How good an explanation is that?” Construct reliabilities for mathematics and English were .82 and .84. Finally, *connecting talk* comprised a single item: “What is the relationship between [m]this mathematical idea and that one/[e]these two ideas?”. Similarly, a group-mean aggregation was applied to transform all classroom talk variables to L2.

Student background. Prior achievement was measured by students’ self-reports of their mathematics and English grade obtained in a national high-stakes examination during the last year of Primary schooling (approximately 12 years old). ICCs were .36 and .33 for mathematics and English, respectively, indicating strong evidence of compositional effects.

5. Results

Table 1 presents the class level (L2) descriptive statistics of the variables. Performative talk was the most dominant form of teacher questions in mathematics (M = 3.66, SD = 0.33) and English (M = 3.52, SD = 0.40). Further, achievement was most strongly related to performative (r=.48) and procedural (r=.40) talk. In English classrooms, achievement was not significantly related to any classroom talk variable. Unlike mathematics, however, cooperative classroom climate was significantly related to all talk measures. Not surprisingly, the strongest relations were found between compositional effects of prior attainment and academic achievement, indicating that classes with higher class-average prior achievement were associated with better achievement.

Table 1. *Descriptive statistics and classroom correlations*

English Math	1	2	3	4	5	6	7	8	9
1. Academic Achievement	--	.75**	.48**	.13	.02	.17	.14	-.03	.16
2. Prior Achievement	.62**	--	.37**	.15	-.00	.18	.04	-.05	.08
3. Cooperative Disp.	.35**	.23**	--	.33**	.32**	.41**	.42**	.33**	.41**
4. Cooperative Instr.	-.16	-.03	-.00	--	.38**	.48**	.44**	.51**	.58**
5. Performative Talk	.48**	.34**	.29**	-.06	--	.47**	.62**	.46**	.55**
6. Procedural Talk	.40**	.24*	.15	.15	.66**	--	.62**	.72**	.81**
7. Clarifying Talk	.07	.09	.20*	.30**	.42**	.60**	--	.65**	.76**
8. Connecting Talk	.31**	.14	.25**	.41**	.54**	.65**	.53**	--	.77**
9. Explaining Talk	.33**	.27*	.17	.28*	.60**	.84**	.71**	.74**	--
Math: Mean (SD)	.027 (.81)	4.73 (1.09)	3.60 (.22)	2.88 (.45)	3.66 (.33)	3.41 (.34)	3.21 (.34)	3.22 (.39)	3.26 (.33)
English: Mean (SD)	.115 (.81)	5.07 (.77)	3.64 (.34)	3.24 (.46)	3.52 (.40)	3.26 (.41)	3.39 (.41)	3.24 (.38)	2.37 (.32)

Note. ***p*<.01; **p*<.05; Correlations (L2) above and below the diagonal represent English and mathematics, respectively.

5.1. Structural model of classroom talk

Guided by the discussion-based recitation sequence (Chapin et al., 2009; Gillies, 2015), we specified a multilevel structural model that begins with presentational talk (performative and procedural) to exploratory talk (clarifying, explaining and connecting). Figure 1 shows the final model across mathematics and English. No student level covariates or outcomes were included. The aim of this analysis was to identify the best model that represents the structural and interdependent relations among the different modes of talk.

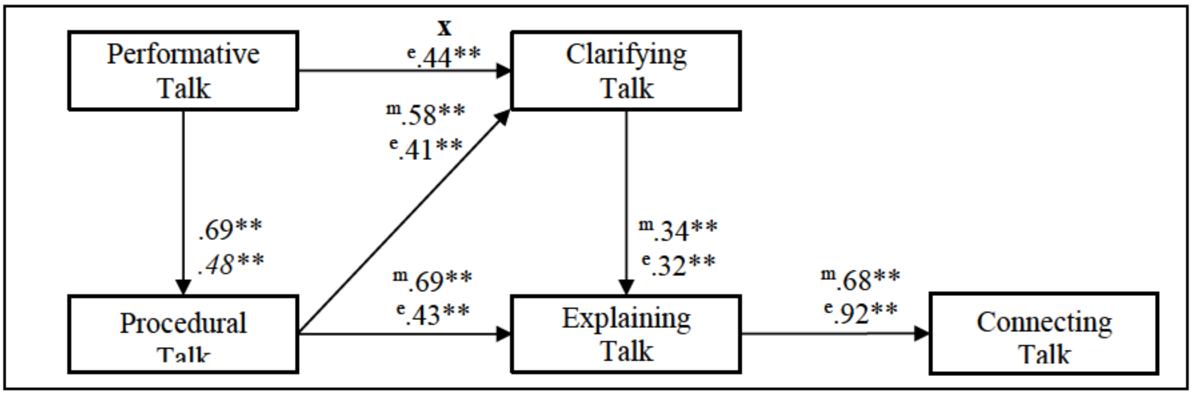

Figure 1. Structural model of classroom talk.
Note. All coefficients are statistical significant (p<.01). Superscripts m and e refer values for mathematics and English. "**x**" denotes paths not modelled.

In mathematics, all regression coefficients (unstandardized) were positive and significant, supported by good model fit statistics: χ^2=3.069 (5), *p*=.689; CFI/TLI=.1.00/1.00, RMSEA=.000. In English, the results were generally consistent with mathematics, with the exception of two. First, in line with expectations of a generative performative effect, an additional parameter was estimated from performative to clarifying talk which significantly improved the overall model (reduction of 22 X^2 units). Second, the coefficient between explaining and connecting talk was considerably larger. Model fit for the English model was also satisfactory: χ^2=7.84 (4), *p*=.10, CFI/TLI=.987/.968, RMSEA=.031. To ascertain the robustness of this model, several competing models were examined but rejected due to poor model fit (e.g., alternative models with explaining, performative or procedural talk specified as the endogenous/ outcome variable).

5.2. Classroom talk and student outcomes

Cooperative learning disposition, cooperative classroom climate and academic achievement

Figure 2 shows the expanded multilevel model regressing students' cooperative learning disposition (L1), cooperative classroom climate (L2) and academic achievement (L1 and L2) on connecting talk (the endogenous L2 variable) in mathematics and English. Overall, both models fitted the data well with the following fit statistics for mathematics: χ^2=22.82 (12), *p*=.03; CFI/TLI=.966/.938, RMSEA=.019; and English: χ^2=16.37 (12), *p*=.18; CFI/TLI=.987/.975, RMSEA=.019.

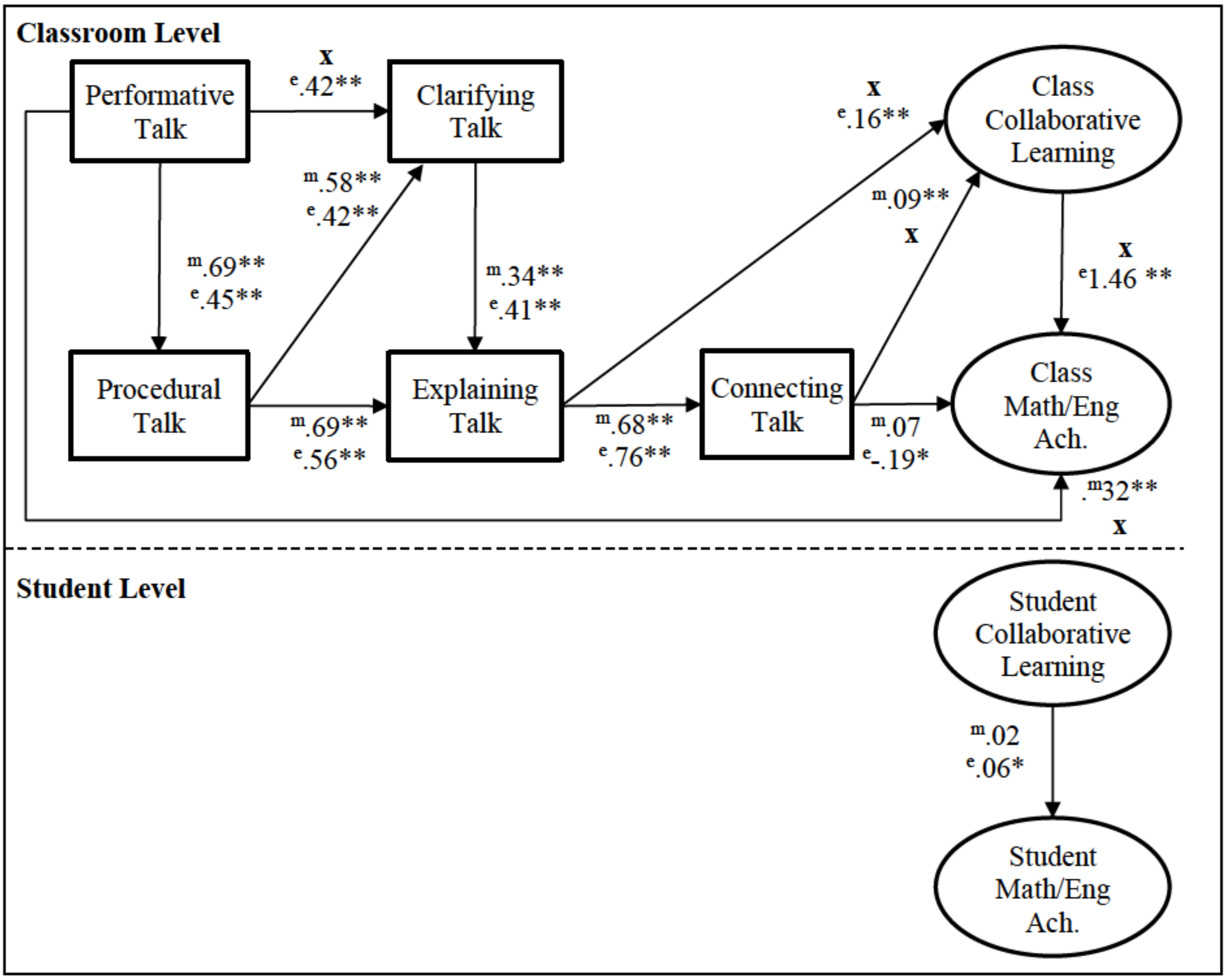

Figure 2. Multilevel structural equation analysis of classroom talk on student outcomes.
Note. Superscripts m and e refer values for mathematics and English. "**x**" denotes paths not modelled.

In mathematics, connecting talk positively predicted cooperative classroom climate (B=.09, p<.01), associated with an effect size of d=.18. Despite the modest absolute value, the Bayes-estimated 95% interval of .05 to .31 suggests a non-trivial effect given the size of the upper bound estimate. Modelling the same predictor on achievement however yielded no statistically significant relationship (B=.07, p>.05). Additionally, the path between the outcome variables was not modelled due to the lack of statistical significance. Finally, the empirical model supported an additional parameter between performative talk and achievement (B=.32, p<.01) associated with a very large effect size (d=.95; 95% CI = .48, 1.42). This relationship was not unexpected given the performative characteristic of mathematics.

In English, several unique relationships were revealed. Unlike mathematics, connecting talk had a significant negative direct effect on achievement (B=-.17, $p<.05$) with a moderate effect size (d=.57; 95% CI = -1.09, -.05), but no effect on cooperative classroom climate (B=-.06, $p>.05$). Second, unlike mathematics, a positive relationship was found between explaining talk and cooperative classroom climate (B=.19, $p<.01$) with a moderate effect size (d=.46; 95% CI = .17, .75). Third, unlike mathematics, cooperative classroom climate had a very large effect on English achievement (B=1.45, $p<.01$) (d = 1.38; 95% CI = .86, 1.96). Fourth, explaining talk generated a number of significant indirect paths into cooperative classroom climate: procedural talk (B=.14, $p<.01$, 95% CI = .05, .24), performative talk (B=.10, $p<.01$, 95% CI = .04, .17) and clarifying talk (B=.08, $p<.01$, 95% CI = .03, .14). Fifth, given the strong relationship between the outcome variables, cooperative classroom climate generated a number of significant indirect paths into achievement: explaining talk (B=.27, $p<.01$, 95% CI = .09, .45), procedural talk (B=.20, $p<.01$, 95% CI = .06, .34), performative talk (B=.14, $p<.05$, 95% CI = .04, .24), and clarifying talk (B=.11, $p<.01$, 95% CI = .03, .20). These indirect effects are not trivial, which suggest that the various facets of talk are well-connected empirically and characterize the interaction of teachers and students in English classrooms. A final noteworthy result was that, unlike mathematics, there was no direct relationship between performative talk and achievement.

5.3. Cooperative learning disposition, cooperative classroom climate and academic achievement, with prior achievement and cooperative scaffolding

The final analysis adds two important explanatory variables to the preceding structural models – prior achievement (L1 and L2) and cooperative scaffolding (L2). Figure 3 presents the results of the integrated models with acceptable model fit statistics for mathematics: χ^2=27.54 (18), p=.07; CFI/TLI=.976/.951, RMSEA=.022; and English: χ^2=39.06 (22), p=.01; CFI/TLI=.966/.941, RMSEA=.028.

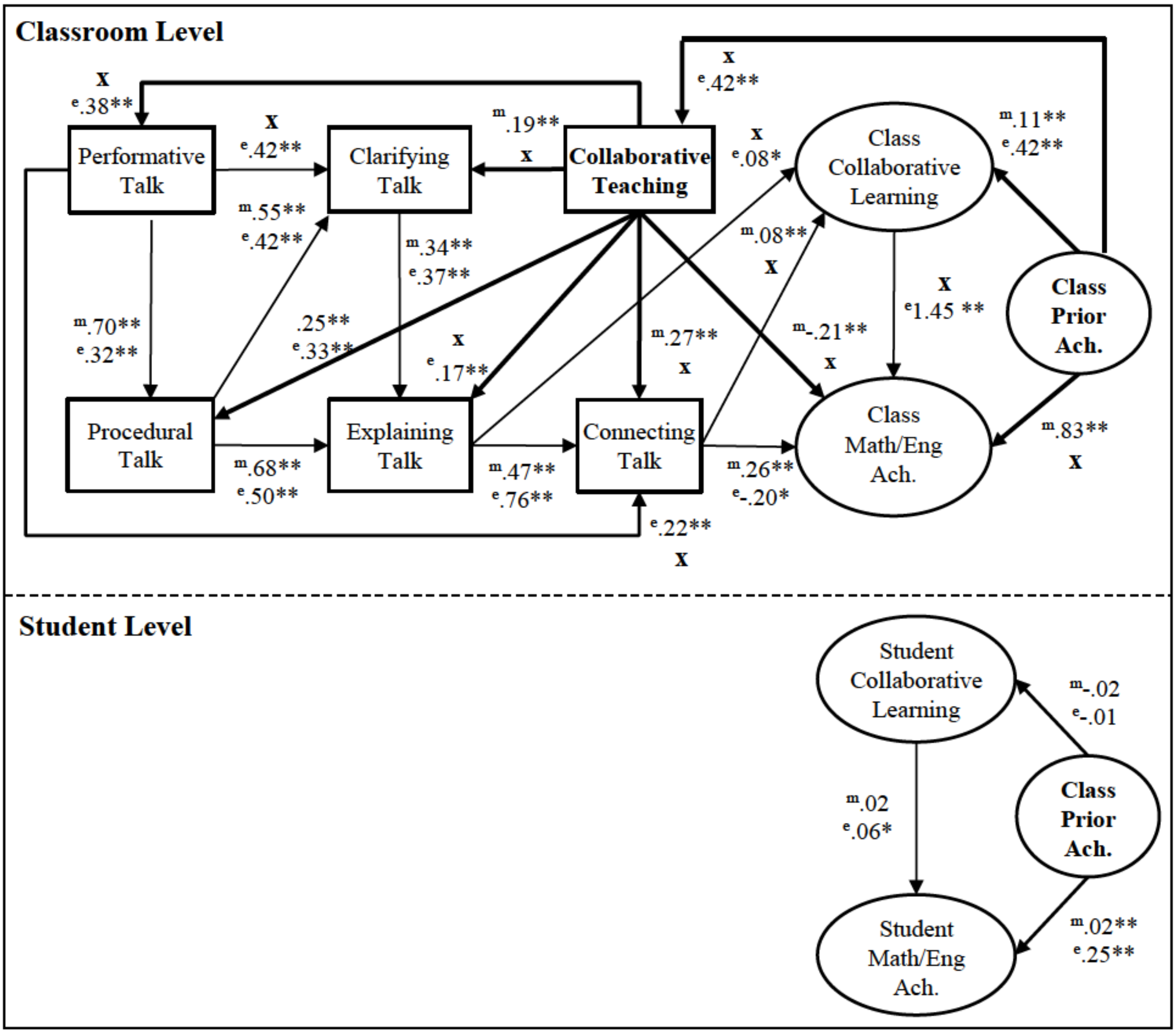

Figure 3. Multilevel structural equation analysis of classroom talk on student outcomes including student and classroom covariates.
Note. For facilitate comparison with the previous model, paths originating from new variables are boldfaced. Superscripts ^{m} and ^{e} refer values for mathematics and English. "**x**" denotes paths not modelled.

In mathematics, the presence of cooperative scaffolding produced three positive effects – procedural talk (B=.25, p<.05), clarifying talk (B=.19, p<.05), connecting talk (B=.27, p<.01). However, we were surprised to find a negative effect on mathematics achievement (B=-.21, p<.01). As this finding was consistent with the bivariate correlation (r=-.16), spuriosity was ruled out. Overall total effect of cooperative scaffolding on achievement was negative but only marginally given the imprecision of the point estimate (B=-.105, p<.05, 95% CI=.01, -.22). In spite of latter finding, several noteworthy findings need to be highlighted.

First, the inclusion of cooperative scaffolding strengthened the relationship between connecting talk and achievement; from B=.07 in the previous model to B=.26. Second, in the presence of cooperative scaffolding, there was an exogenous backward shift of performative talk away from achievement onto connective talk. In other words, the direct effect between performative talk and achievement (B=.32, from Figure 2) was completely mediated by classroom talk (specifically, connecting talk), producing a total indirect effect of B=.13 (p<.01, 95% CI=.07, .22).

In English, cooperative scaffolding also produced three positive effects – performative talk (B=.38, p<.01), procedural talk (B=.35, p<.01), explaining talk (B=.17, p<.01). Unlike mathematics, a direct relationship between cooperative scaffolding and achievement did not exist, but was mediated by cooperative classroom climate and, interestingly, explaining talk. Total indirect effect (sans connecting talk) was statistically significant (B=.11, p<.05, 95% CI=.04, .18).

With respect to prior ability, preliminary analyses without classroom variables indicated strong compositional effects between class-average prior achievement and student achievement (B=.890, d=1.52, for mathematics and B=1.095, d=.85, for English), but weaker relations with cooperative classroom climate (B=.116, d=.15 for mathematics and B=.233, d=.23, for English). The full model in Figure 3 clearly revealed that class-average prior achievement continue to extend a very strong influence on current achievement, in particular mathematics. Indeed, taking into account the preliminary and main analyses, the inclusion of classroom predictors produced a modest reduction (Δd=.10) in the compositional effects (of prior ability) on mathematics achievement (B=.830, d=1.42) but had no appreciable change (Δd=.02) for cooperative classroom climate (B=.112, d=.17).

For English, however, the direct link between prior and current achievement no longer existed but was mediated by cooperative scaffolding. Therefore, despite the strong association with prior and current achievement (*r*=.75), the total indirect effect from class prior achievement on current achievement, via classroom talk and cooperative scaffolding was B=.25 (*p*<.01; 95% CI=.07, .22), more than four times lower than the estimated direct effect of B=1.095 observed in our preliminary analysis. These findings indicate that the combined effect of classroom talk and cooperative scaffolding fully mediated the compositional effects of class prior achievement on achievement in English. Moreover, compared to mathematics, the inclusion of classroom predictors produced a much larger increase (Δd=.28) in the compositional effects on cooperative classroom climate (B=.830, *d*=51). One positive aspect of the English results suggest that the use of productive talk and cooperative scaffolding is quite successful in reducing within-class student background effects in learning. On the other hand, however, the findings also suggest that teachers use dialogic teaching lean towards higher ability classrooms.

6. Discussion

Quantitative studies of classroom talk are rare with the field largely dominated by qualitative classroom observations. Even when quantitative studies exist, the methodology employed typically involved computation of field coding frequencies followed by reports using basic inferential statistics (Howe & Abedin, 2013) or descriptive analysis of coded transcribed data (Mercer, 2010). This study therefore offers a novel quantitative perspective into classroom discourse theories proposed by leading qualitative scholars (Alexander, 2008; Barnes, 2008; Mercer, 2000) and attend to Howe and Abedin's (2013) plea to dispel "the widespread unease about quantitative analysis in relation to dialogue" (p. 345).

In her discussion of classroom discourse, Cazden (2001) made the observation that the true merit in recitation instruction lies "in the difference between helping a child somehow get a particular answer and helping that child gain some conceptual understanding from which answers to similar questions can be constructed at a future time" (p. 92). In general, our quantitative findings support this preposition and highlight that it not so much a contest of recitation versus exploration, but recitation *and* exploration.

Addressing the first research question, the classroom talk model (Figure 1) revealed a relatively tight coupling of various facets of classroom talk across mathematics and English classrooms. Performative teacher questions were not necessarily just talk about getting the right answer or determining whether a solution or an argument is true or false. Rather, these questions were very often followed up by a stable sequence of procedural and discussion-based questions that involved clarifying, explaining and making connections – all components of academically productive talk (Alexander, 2008; Mercer & Littleton, 2007; Resnick et al., 2011). Importantly, the magnitude of the regression parameters in the presented model coupled with the rejection of competing models suggest that mathematics and English teachers in Singapore do possess a rather high level of tacit and reasonably developed causal understandings of the types of relationships between the different kinds of questions that they ask and what prompts students to "work on understanding".

Support for our second research question was mixed as exploratory forms of classroom talk did not predict mathematics achievement and the relationship with cooperative classroom climate was modest (d=.18). In mathematics, only one aspect of talk predicted achievement – performative talk (d=.95, p<.01) – while the solitary link between connecting talk and cooperative classroom climate did not achieve statistical significance. In English, however, the same relationship was not observed. Rather, the only significant effect (and in the negative direction) was between connecting talk and achievement (d=.57). While this finding was unexpected given the importance of connective interpretations in literacy and language subjects, it was not an unreasonable account of the types of instructional choices teachers often have to make that competes between spaces for substantive knowledge work, on one hand, and curriculum coverage and exam preparation, on the other (Hogan, Chan et al., 2013). Moreover, while exploratory talk provides the richest and most valuable contribution to the quality of learning, teaching of this kind often requires considerable attention to student engagement and aptitude for learning and teacher skills in scaffolding techniques that are often not easy (Tomlinson, 2001).

By all accounts, a key finding in English classrooms is in the very large predicted effect between cooperative classroom climate and achievement (*d* = 1.38). This relationship is hugely important as it sets up the nomological sequences of dialogic exchanges beginning with performative talk to procedural to clarifying and finally to connecting talk through explaining talk. This suggests that in English classes, cooperative classroom climate is the key pedagogical hinge by which classroom talk and student achievement become instructionally and causally linked. Specifically, English teachers who asked questions that required extended explanations could only improve achievement if they promoted a cooperative climate, where students were involved in class discussions, where cooperative learning was supported and students were receptive to alternative opinions from others.

In further articulation of these findings, one might query why the limited influence of connecting talk in English classrooms; despite its above average occurrence (M=3.24), and given that the ability to make connections is an important characterization of dialogic talk. To explain this conundrum, we applied two respecifications to the English model: first, we removed both direct paths from explaining talk, and second, we allowed only connecting talk to predict cooperative climate and English achievement. Although a satisfactory model fit was obtained, the effect on cooperative climate was not statistically significant (B=.09), while the effect on achievement was significant but negative (B=-.16). As the respecified model also did not fit as well as the current model (AIC= 5899 versus AIC=5890), this supplementary test suggests that while connecting talk matters, it is less productive than explaining talk; as others have also found (e.g., Howe et al., 2007).

6.1. Classroom compositional and cooperative climate effects on classroom talk

In support of our final research question, we found very substantial compositional (class prior achievement) and climate (cooperative scaffolding) effects on classroom talk and student achievement. In mathematics, compositional effects of prior ability on achievement remained substantially strong (*d*=.83), taking into account other variables in the model. In English, however, the inclusion of classroom talk, through cooperative scaffolding, and particularly cooperative climate, fully replaced and accounted for the types of compositional effects observed in the mathematics model. This finding signals the pivotal role that cooperative scaffolding enacts in English classrooms.

Another important finding we wish to highlight extends from Figures 2 and 3, which showed that the pedagogical benefits derived from productive exploratory talk occurred only in the presence of effective teacher scaffolds (Bakker et al., 2015; Lin et al., 2015; Rojas-Drummond et al., 2013; van de Pol, Volman, & Beishuizen, 2010). While performative talk in mathematics (Figure 2) directly predicted achievement bypassing all other forms of classroom talk, in the integrated model (Figure 3) with cooperative scaffolding, performative talk no longer produced a direct link to achievement. Rather, performative talk assumes a more constructive role through connecting talk (note that this path did not exist in Figure 2). From a disciplinary practice and classroom discourse perspective, this finding, coupled with the increased effect of connecting talk to achievement (from B=.07 to B=.26) and the weakening of the performative paths, are hugely important and recognize the coexistence of performative and exploratory dialogue for working on and deepening understanding.

6.2. Conclusion

Our literature review reveals that despite repeated calls for teachers to engage in dialogic teaching, and despite evidence that dialogic teaching is beneficial, teachers continue to drill students on answers that are predominantly right or wrong, and students continue to engage in classroom routines that offer little opportunities for extended dialogue and meaningful exchange of ideas (Lefstein, & Snell, 2010; Nystrand et al., 2003). Our findings are broadly consistent with this exposition in regard to the dominance of performative teacher questions and the direct inferences we might draw about them. On the other hand, our findings also revealed that conventional recitation instruction can be educationally valuable when used skilfully with exploratory talk and especially when combined with teacher-facilitated cooperative scaffolding.

6.3. Limitations

This study contains several limitations, the most major of which relates to the facets of classroom talk investigated. While we value the significance of expanding the parameters of exploratory and presentational talk, they lack the conceptual depth necessary to articulate the different modes of productive talk moves which reflect how students express and solicit ideas, negotiate alternative propositions and reason cogently (Mercer, 2000). Future quantitative researchers may find Howe et al.'s (2019) work inspiring who, by their own account, conducted the first large-scale investigation of teacher-student dialogue on student outcomes, while also taking into account a comprehensive set of antecedent variables. However, the extensive use of classroom video recordings as the primary source of data may present practical challenges for quantitative researchers with limited financial resources and expertise on classroom videography. If survey research remains the preferred mode, we think that future studies ought to expand the conceptual coverage of the talk constructs by including more items that illuminate more nuanced aspects of productive teacher-student dialogue.

Another limitation of this study is the lack of comparable qualitative data to validate the results from the quantitative instruments. However, while we acknowledge this limitation, we also posit that the methodological design of this study makes qualitative comparisons improbable due to the large sample size involved (all Secondary Three classrooms in each school). Moreover, beyond sample size issues, one important methodological distinction is the temporal referent of classroom talk. In this study, students were asked to report the frequency of particular modes of talk in a particular school year, for particular subjects. Therefore, it is reasonable to infer that our data on classroom talk reflects an overall (or snapshot) perception of the student's engagement with the teacher throughout the year, and aggregated over multiple student raters within each class. This is unlike qualitative studies where observations are typically collected over several consecutive lessons (covering a lesson topic or unit) or random single lessons spread across the year. Although some researchers have tried to address this issue, they also experienced practical field challenges in data collection. For example, unavoidable scheduling conflicts meant that student surveys and classroom observations were often administered across different semesters, thus, affecting comparability of both rating methods (e.g., Van der Scheer, Bijlsma, & Glas, 2019). Notwithstanding the lack of qualitative data, we take comfort that some of our descriptive findings, particularly the higher mean scores for performative talk and weak relationship between exploratory forms of talk and academic achievement, were quite consistent with findings reported by related qualitative studies (see Sub-section 2.3). In sum, the validity and reliability of student ratings of classroom instruction remain a work-in-progress.

References

Alexander, R. (2008). Essays on pedagogy. London: Routledge.

Bakker, A., Smit, J., & Wegerif, R. (2015). Scaffolding and dialogic teaching in mathematics education: Introduction and review. *ZDM Mathematics Education, 47*(7), 1047–1065. https://doi.org/10.3102/0034654314551063.

Barnes, D. (1976). From communication to curriculum. Harmondsworth, England: Penguin.

Barnes, D. (2008). Exploratory talk for learning. In N. Mercer, & S. Hodgkinson (Eds.). *Exploring talk in school* (pp. 1–16). London: Sage Publications.

Bruner, J. (1983). Child's talk: Learning to use language. New York: Norton.

Bokhove, C. (2018). Exploring classroom interaction with dynamic social network analysis. *International Journal of Research and Method in Education, 41*, 17-37. https://doi.org/10.1080/1743727X.2016.1192116.

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

Chapin, S. H., O'Connor, C., & Anderson, N. C. (2009). *Classroom discussions: Using math talk to help students learn: Grades 1–6* (second edition). Sausalito: Math Solutions Publications.

Chi, M. T. H., Deleeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477. https://doi.org/10.1207/s15516709cog1803_3.

Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment, 4*, 37-73. https://doi.org/10.1207/s15326977ea0401_2.

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9. https://doi.org/10.1016/j.learninstruc.2013.07.001.

Gillies, R. M. (2014). Developments in classroom-based talk. *International Journal of Educational Research, 63*, 63-68. https://doi.org/10.1016/j.ijer.2013.05.002.

Gillies, R. M. (2015). *Enhancing classroom-based talk*. London: Routledge.

Gillies, R. M. (2016). *Enhancing classroom-based talk: Blending practice, research and theory*. London, UK: Routledge.

Gillies, R. M. (2017). Promoting academically productive student dialogue during collaborative learning. *International Journal of Educational Research, 97*, 200-209. https://doi.org/10.1016/j.ijer.2017.07.014.

Harker, R., & Tymms, P. (2004). The effects of student composition on school outcomes. School Effectiveness and *School Improvement, 15*(2), 177-199. https://doi.org/10.1076/sesi.15.2.177.30432.

Hattie, J. (2015). The applicability of Visible Learning to higher education. *Scholarship of Teaching and Learning in Psychology, 1*(1), 79-91. https://doi.org/10.1037/stl0000021.

Hogan, D., Chan, M., Rahim, R., Kwek, D., Aye, K. M., Loo, S.-C., et al. (2013). Assessment and the logic of instructional practice in Secondary 3 English and mathematics classrooms in Singapore. *Review of Education, 1*(1), 57-106. https://doi.org/10.1002/rev3.3002.

Hogan, D., Kwek, D., Towndrow, P., Rahim, R. A., Tan, T.-K., Yang, H.-J., et al. (2013). Visible learning and the enacted curriculum in Singapore. In Z. Deng, S. Gopinathan, & C. K.-E. Lee (Eds.). *Globalization and the Singapore curriculum: From policy to classroom* (pp. 121–150). Singapore: Springer.

Howe, C. (2010). *Peer groups and children's development*. Oxford: Blackwell.
Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education, 43*(3), 325-356. https://doi.org/10.1080/0305764X.2013.786024.
Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., Jessiman, E., & Donaldson, C. (2007). Group work in elementary science: Organizational principles for classroom teaching. *Learning and Instruction, 17*, 549-563. https://doi.org/10.1016/j.learninstruc.2007.09.004.
Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher-student dialogue during classroom teaching: Does it really impact on student outcomes? *Journal of the Learning Sciences, 28*, 462-512. https://doi.org/10.1080/10508406.2019.1573730.
Johnson, D., & Johnson, R. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365-379. https://doi.org/10.3102/0013189X09339057.
Johnson, D., Johnson, R., Dudley, B., Mitchell, J., & Fredrickson, J. (1997). The impact of conflict resolution training on middle school students. *The Journal of Social Psychology, 137*, 11-21. https://doi.org/10.1023/B:ADCO.0000027425.24225.87.
Kogut, G., & Silver, R. E. (2009). Teacher talk, pedagogical talk and classroom activities. *Proceedings of the Redesigning Pedagogy Conference*. June, Retrieved from https://repository.nie.edu.sg/bitstream/10497/16429/1/CRPP-2009-Galut%26Silver_a.pdf.
Krathwohl, D. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*, 212-218. https://doi.org/10.1207/s15430421tip4104_2.
Lefstein, A., & Snell, J. (2010). Classroom discourse: The promise and complexity of dialogic practice. In S. Ellis, & E. McCartney (Eds.). *Applied linguistics and primary school teaching* (pp. 165-185). Cambridge: Cambridge University Press.
Lefstein, A., & Snell, J. (2014). Better than best practice: Developing teaching and learning through dialogue. London: Routledge.
Lin, T.-J., Jadallah, M., Anderson, R. C., Baker, A. R., Nguyen-Jahiel, K., Kim, I.-H., et al. (2015). Less is more: Teachers' influence during peer collaboration. *Journal of Educational Psychology, 107*(2), 609-629. https://doi.org/10.1037/a0037758.
Littleton, K., & Howe, C. (2010). *Educational dialogues: Understanding and promoting productive interaction*. Abingdon, Oxon: Routledge.
Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*, 203-229. https://doi.org/10.1037/a0012869.
Mehan, H. (1979). Learning lessons. Cambridge, MA: Harvard University Press.
Mercer, N. (1992). Culture, context and the construction of knowledge in the classroom. In P. Light, & G. Butterworth (Eds.). *Context and cognition: Ways of learning and knowing* (pp. 28–46). London: Harvester Wheatsheaf.
Mercer, N. (2000). *Words and minds: How we use language to think together*. London: Routledge.
Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *The British Journal of Educational Psychology, 80*(1), 1-14. https://doi.org/10.1348/000709909X479853.
Mercer, N., & Dawes, L. (2008). The value of exploratory talk. In N. Mercer, & S. Hodgkinson (Eds.). *Exploring talk in school* (pp. 55-72). London, England: Sage.

Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education, 40*(4), 430-445. https://doi.org/10.1080/03054985.2014.934087.
Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning Culture and Social Interaction, 1*(1), 12-21. https://doi.org/10.1016/j.lcsi.2012.03.001.
Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking*. London: Routledge.
Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve mathematics problems. *Language and Education, 20*(6), 507-528. https://doi.org/10.2167/le678.0.
Mercer, N., & Hodgkinson, S. (Eds.). (2008). *Exploring talk in school*. London: Sage.
Morin, A. J., Marsh, H. W., Nagengast, B., & Scalas, L. F. (2014). Doubly latent multilevel analyses of classroom climate: An illustration. *Journal of Experimental Education, 82*(2), 143-167. https://doi.org/10.1080/00220973.2013.769412.
Munger, G. F., & Loyd, B. H. (1988). The use of multiple matrix sampling for survey research. *Journal of Experimental Education, 56*(4), 187-191. https://doi.org/10.1080/00220973.1988.10806486.
Nind, M., Curtin, A., & Hall, K. (2016). *Research methods for pedagogy*. London: Bloomsbury.
Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35*(2), 135-198. https://doi.org/10.1207/S15326950DP3502_3.
Opdenakker, M.-C., Van Damme, J., De Fraine, B., Van Landeghem, G., & Onghena, P. (2002). The Effect of schools and classes on mathematics achievement. *School Effectiveness and School Improvement, 13(*4), 399-427. https://doi.org/10.1076/sesi.13.4.399.10283.
Resnick, L. B., Michaels, S., & O'Connor, C. (2011). How (well structured) talk builds the mind. In R. Sternberg, & D. Preiss (Eds.). *From genes to context: New discoveries about learning from educational research and their applications* (pp. 163-194). New York: Springer.
Rojas-Drummond, S., Mazón, N., Fernández, M., & Wegerif, R. (2006). Explicit reasoning, creativity and co-construction in primary school children's collaborative activities. *Thinking Skills and Creativity, 1*(2), 84-94. https://doi.org/10.1016/j.tsc.2006.06.001.
Rojas-Drummond, S., Torreblanca, O., Pedraza, H., Vélez, M., & Guzmán, K. (2013). Dialogic scaffolding: Enhancing learning and understanding in collaborative contexts. *Learning Culture and Social Interaction, 2(*1), 11-21. https://doi.org/10.1016/j.lcsi.2012.12.003.
Simpson, A., Mercer, N., & Majors, Y. (2010). Editorial: Douglas Barnes revisited: If learning floats on a sea of talk, what kind of talk? And what kind of learning? *English Teaching Practice & Critique, 9*(2), 1-6.
Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils.* London: Oxford University Press.
Slavin, R. (2013). Cooperative learning and achievement: Theory and research. In R. W. Miller, & G. I. Weiner (Eds.). *Handbook of psychology* (pp. 199-212). (2nd ed.). Hoboken, NJ: Wiley.
Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de psicología, 30*(3), 785-791. https://doi.org/10.6018/analesps.30.3.201201.

Smith, F., Hardman, F., Wall, K., & Mroz, M. (2004). Interactive whole-class teaching in the national literacy and numeracy strategies. *British Educational Research Journal, 30,* 395-411. https://doi.org/10.1080/01411920410001689706.

Tarim, K. (2009). The effects of cooperative learning on preschoolers' mathematics problem solving ability. *Educational Studies in Mathematics, 72,* 325-340. https://doi.org/10.1007/s10649-009-9197-x.

Tomlinson, C. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Upper Saddle River, NJ. Pearson: Merrill Prentice Hall.

Van Boxtel, C., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*(4), 311-330. https://doi.org/10.1016/S0959-4752(00)00002-5.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–Student interaction: A decade of research. *Educational Psychology Review, 22*(3), 271-296. https://doi.org/10.1007/s10648-010-9127-6.

Van der Scheer, E. A., Bijlsma, H. J. E., & Glas, C. A. W. (2019). Validity and reliability of student perceptions of teaching quality in primary education. *School Effectiveness and School Improvement, 30*(1), 30-50. https://doi.org/10.1080/09243453.2018.1539015.

Walshaw, G., & 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. https://doi.org/10.3102/0034654308320292.

Webb, N. M., Franke, M. L., Ing, M., Wong, J., Fernandez, C. H., Shin, N., et al. (2014). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research, 63,* 79-93. https://doi.org/10.1016/j.ijer.2013.02.001.

Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. The Journal of the Learning Sciences, 15(3), 379–428. https://doi.org/10.1207/s15327809jls1503_3.

Wilkinson, I. A. G., & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research, 37*, 425-447. https://doi.org/10.1016/S0883-0355(03)00014-4.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record, 112*(4), 1008-1037.

Wolf, M., Crosson, A., & Resnick, L. (2006). *Accountable talk in reading comprehension instruction. CSE technical report 670.* Learning and Research Development Center. University of Pittsburgh.