

Dialogic Practices in Primary School Classrooms

Language and Education, in press

Maria Vrikki^{a*1}, Lisa Wheatley^a, Christine Howe^a, Sara Hennessy^a, Neil Mercer^a

^a Faculty of Education, University of Cambridge, 184 Hills Road, CB2 8PQ, Cambridge, UK.

*Corresponding author: maria.vrikki@outlook.com; Telephone: +44 1233 767569

Abstract

Research into classroom dialogue suggests that certain forms are especially productive for students' learning (Howe and Abedin, 2013). Despite the large number of studies in this area, there is inadequate evidence about the prevalence of the identified forms, let alone their productivity. However, scarcity is widely presumed. The overall aim of the study reported in this paper was to examine the extent to which the forms are embedded within current practice in English primary schools. Video-recordings of two lessons from each of 36 classrooms formed the database, with two subjects from mathematics, English and science covered in each classroom. Each lesson was coded per turn for the presence of 'dialogic moves' and rated overall for the level of student involvement in specified activities. Results revealed that the supposedly productive forms were not always as scarce as sometimes presumed, while also highlighting huge variation in their relative occurrence. They also point to the role of professional development for teachers in promoting use of some forms.

Keywords:

Classroom dialogue; Classroom interaction; Primary school education; Teacher professional development

¹ Author changed affiliation at the time of revision. The new affiliation is the Department of Education, University of Cyprus, Dramas 11-13, Theophanides Building, Nicosia 1077.

1. Introduction

1.1. Productive Classroom Dialogue

Classroom dialogue has been heavily researched in recent years due to its perceived role in student learning. Influenced by socio-cultural perspectives, authors in this field view learning as a social activity, mediated through dialogue. Specifically, dialogue is perceived as the intermediary between collective and individual thinking (Vygotsky 1962). Its quality, therefore, becomes particularly important as it determines the quality of collective thinking and, through this, individual progress. These views have resulted in research which aims to identify forms of dialogue that promote higher order thinking and, thus, are optimal for learning. Thanks to this research, there is now a fair degree of consensus over which forms are especially productive (Littleton and Mercer 2013).

The characteristics of optimal classroom dialogue proposed by Alexander (2008) have proved particularly influential. According to Alexander, classroom dialogue should be: 1) collective with participants reaching shared understanding of a task; 2) reciprocal with ideas shared among participants; 3) supportive with participants encouraging each other to contribute and valuing all contributions; 4) cumulative, guiding participants towards extending and establishing links within their understanding; and 5) purposeful, that is directed towards specific goals.

Similar forms of dialogue have been highlighted in the context of student-student interaction. Mercer and colleagues (e.g. Littleton and Mercer 2013) have identified three types of student-student talk: disputational, cumulative and exploratory. Characterised by disagreement and individualised decisions, disputational talk was

thought to be the least educationally productive. Some educational value was attributed to cumulative talk, as it was characterised by general acceptance of ideas, but lack of critical evaluation. Exploratory talk was observed less frequently; yet, it was regarded as the most educationally effective. It involved participants engaging critically with ideas and attempting to reach consensus. Initiatives, like the “Thinking Together” programme (Dawes, Mercer, and Wegerif 2003; Mercer and Littleton 2007), aimed to promote primary school children’s use of exploratory talk, and showed a positive impact on students’ problem solving, mathematics and science attainment/learning. Likewise, “accountable talk” has been promoted as the most academically productive classroom talk (Michaels, O’Connor, and Resnick 2008). It encompasses accountability to: 1) the learning community, through listening to others, building on their ideas and expanding propositions; 2) accepted standards of reasoning, through emphasis on connections and reasonable conclusions; and 3) knowledge, with talk that is based on facts, texts or other publicly accessible information and challenged when there is lack of such evidence.

Working in secondary classrooms, Nystrand and colleagues (e.g. Nystrand, Gamoran, Kachur, and Prendergast 1997) characterised dialogic instruction via three key discourse moves that teachers might make: 1) authentic questions, which are questions with no predetermined answers; 2) uptake, which occurs when previous answers are incorporated into subsequent questions; and 3) high-level evaluation, which occurs when teachers elaborate or ask follow-up questions in response to students’ replies, instead of giving a simple evaluation, such as “Good” or “OK” (Nystrand, Gamoran, Zeiser, and Long 2003).

While there are differences between these approaches, there are also marked commonalities, regardless of whether the research refers to whole class or small group contexts. Shared features include:

- Invitations that provoke thoughtful responses (e.g. authentic questions, asking for clarifications and explanations);
- Extended contributions that may include justifications and explanations;
- Critical engagement with ideas, challenging and building on them;
- Links and connections;
- Attempts to reach consensus by resolving discrepancies.

For these features to occur, a generally participative ethos is important, with participants respecting and listening to all ideas. This necessitates making the discourse norms accessible to all (Michaels, O'Connor, and Resnick 2008). Changing the classroom culture in this manner might be a challenge for any teacher.

1.2. Professional Development on Classroom Dialogue

The characteristics of productive classroom dialogue have been widely disseminated in practitioner publications and have also formed the basis of professional development (PD) initiatives. These initiatives have typically been intervention programmes, involving workshops that promote target features and discussion meetings with research teams around specific experiences (e.g. video-recorded lessons). Typically,

the success of the programmes is indexed through comparing use of target dialogue during pre- and post-intervention lessons. The outcomes have been mixed, with some studies reporting increases in all target features, and others reporting partial or no success in changing practice.

Studies showing limited success include Pehmer, Gröschner, and Seidel (2015). Their Dialogic Video Cycle programme resulted in teachers' feedback becoming more focused on students' learning processes and self-regulation. Yet, no change was observed for teachers' questions and students' talk. Similarly, Wells and Arauz's (2006) seven-year programme led to an increase in the number of discussion-type sequences. However, the proportion of these sequences remained low. Lefstein and Snell's (2014) one-year programme promoting interactional awareness assessed teachers' questions (e.g. open, closed, uptake), teachers' feedback (e.g. elaborated, non-elaborated), and students' contributions (e.g. response to teacher, spontaneous contribution, choral response). The sole increase was openness in teachers' questions. Finally, Ruthven et al.'s (2017) epiSTEMe intervention placed strong emphasis on dialogue in small group and whole-class settings. A range of markers was assessed, including teachers asking for explanations, clarifications and reasoning, as well as students providing reasons, and taking extended turns. While some teachers implemented some target features, the programme was not successful for all features and all participants.

Other interventions seem, however, to have been more successful. Sedova, Sedlacek and Svaricek (2016) found that in seven out of eight classrooms their action research programme (including workshops, video-recorded lessons and reflective interviews) boosted students' talk with reasoning, teachers' use of open questions,

teacher uptake (i.e. building on students' contributions), and open discussion. Similarly, Chinn, Anderson, and Waggoner (2001) supported four teachers in using a collaborative reasoning technique through half-day workshops followed by discussions. They reported increases in the amount of student talk, students' elaborated utterances with evidence, and the proportion of authentic teacher questions. Working with a single teacher, Haneda, Teemant, and Shearman (2017) reported evidence for joint inquiry, open exchange of ideas, and engagement with multiple perspectives. In an intervention promoting inquiry dialogue, Wilkinson et al. (2017) found that scores on their Argument Rating Tool, which measured the "quality of teacher facilitation and student argumentation" (Wilkinson et al. 2017, 71), significantly increased. Hennessy, Dragovic, and Warwick (2017) explored their PD programme's impact on teachers' practice through video-stimulated discussions and a multimedia resource bank. Interviews with teachers indicated increases in understanding and use of target dialogue around interactive whiteboards. Finally, Alexander, Hardman and Hardman (2017) offered a substantial PD programme of eleven cycles of mentoring and self-evaluations to improve the quality of classroom talk. They reported a positive impact on several indicators of teachers' and students' talk.

Despite the positive outcomes of some programmes, there is an issue of scalability (Howe and Mercer 2017). In most of the seemingly successful programmes (but also many of their less successful counterparts), there was huge investment of time and effort from researchers and teachers. Wilkinson et al. (2017) offered two 6-hour workshop days, biweekly meetings with teachers, and monthly individual coaching (30-40 minutes each). Haneda et al. (2017) offered a 30-hour summer workshop and seven cycles of individualised coaching in classrooms. Alexander et al. (2017)

undertook 20 weeks of intensive intervention, and Sedova et al. (2016) offered a one-year programme. Therefore, the potential for scaling these programmes up for larger groups of teachers is questionable.

Another issue is sustainability. Despite the intensive support provided by these programmes, their long-term impact has seldom been measured. Exceptionally, Hennessy et al. (2017) observed two lessons (English and science) ten weeks after the end of their programme. Field notes and materials from observed lessons illustrated that teachers continued to pose open-ended questions, construct shared interpretations and encourage students to justify and build on others' ideas. However, the follow-up sample was small owing to resource limitations and it is unknown whether all participants sustained their practices beyond the intervention. Apart from this study, the long-term impact of PD on the quality of classroom dialogue has not been investigated.

1.3. Prevalence of IRF Pattern

Indeed, observational studies give the strong impression that features of productive classroom dialogue are not firmly embedded in current practice (Howe and Abedin 2013). Instead, the dominant form in teacher-student interactions is thought to remain the traditional Initiation-Response-Feedback (IRF) format, first noticed by Sinclair and Coulthard (1975) and subsequently reported in classrooms across the world (Nystrand et al. 1997; Wells and Arauz 2006). This format involves teachers asking mostly closed questions with “low cognitive demand” (Sedova et al. 2016, 14), students producing short and simple answers, and teachers evaluating those answers based on their correctness.

Without doubt, the ubiquity of the IRF format is well established. For instance, in their analysis of mathematics lessons, Berry and Kim (2008) found that teacher talk was “chiefly recitational” (323), with the two main types of question, eliciting and incremental, both closed and leading. Such questions impose tight control over student participation, a finding endorsed through Bleicher, Tobin and McRobbie’s (2003) analysis of talk during a chemistry class. Similarly, Pontefract and Hardman (2005) found that teacher-led recitation, rote and repetition dominated classroom interactions with little focus on student understanding. Moreover, in mathematics classrooms, Sepeng (2011) found that triadic dialogue prevailed even when knowledge was dialogically co-constructed.

1.4. Focus of the Paper

Yet while the cumulative evidence suggests that the IRF format is extremely common, the frequency of other forms has not been thoroughly examined. It remains possible that there are ‘pockets of excellence’ in some classrooms, perhaps (but not necessarily given the aforementioned issues of scalability and sustainability) related to teachers’ prior PD around dialogue. Using data drawn from a larger project² which assessed the implications of classroom dialogue for student outcomes, the main aim of the study reported here was to assess the incidence of the forms pinpointed earlier as productive.

² The ESRC-funded project ‘Classroom Dialogue: Does it really make a difference for student learning?’, led by Howe, Hennessy and Mercer, ran from 2015-2017. See <http://tinyurl.com/ESRCdialogue>.

To the extent that ‘pockets of excellence’ were detected, a subsidiary aim was to examine how far they were teacher-driven (as opposed, say, to being dependent on students or even the subject of study). For reasons of manageability, the work was restricted to teacher-student dialogue (i.e. teacher-whole class, teacher-small group, teacher-individual, but not student-student/s). Thus, the main research question for the present paper was:

- 1) To what extent does teacher-student dialogue involve forms that are widely seen as productive?

In addition, two supplementary questions were formulated:

- 2) Can any variation in key forms be attributed to the participants, and if it can be, what contribution do teachers make as opposed to students?
- 3) To the extent that teachers play a key role, is their previous participation in professional development relating to dialogue likely to be important?

2. Methodology

2.1. Sample

Seventy-two lessons comprised the sample for this paper. These involved 36 teachers in 28 primary schools located in Cambridgeshire (42%), London (22%) and other northern, central and eastern areas of England, jointly representing a diverse geographical area. Teachers were predominantly female (67%).

Each teacher contributed two lessons covering two core subjects from the primary curriculum (Department for Education 2013), i.e. two of mathematics, English and science. Power statistics indicated that a minimum of 22 lessons per subject was needed³, and the selected teachers were the first 36 from the larger project's sample permitting compliance with that minimum while also allowing all possible pairs of mathematics-English, English-science and mathematics-science to be sampled equally.

To address the PD research question, teachers were asked whether they had received any PD relevant to classroom dialogue. While many reported not receiving any such PD, others gave a variety of responses, including self-guided research, input during initial or in-service teacher education, their school being involved in relevant research projects, or receiving relevant short staff training. Although these PD experiences varied in length and content, it seemed safe to assume that teachers who had received some PD had increased awareness of the meaning of productive classroom dialogue, in comparison with teachers who had no exposure whatsoever. Teachers

³ A priori power analysis was conducted using G*Power, with an effect size estimate of 0.40, and power of 0.80. The sample size required was 66 lessons across subjects, thus a minimum of 22 lessons per subject.

therefore were divided into two groups: 1) those with prior professional development on classroom dialogue (designated PriorPD, $N=18$); 2) those with no such prior professional development (designated NoPD, also $N=18$).

As for the classes, these had a mean of 28 students ($SD=2.78$), from diverse socio-economic backgrounds (ranging from 0-100% of students eligible for free school meals, $M=15.58\%$ eligible, $SD=21.34$). The classes also varied greatly over the number of students with English as an additional language (ranging from 0-97%, $M=16.47\%$ EAL, $SD=20.55$), although most students ($M=97.08\%$, $SD=7.71$) were reported by their teachers to be fluent in English. The classes ranged from being 0% to 100% minority ethnic ($M=33.48\%$, $SD=32.49$), although only five classes had more than 75% minority ethnic students. A small number of students were registered with Special Educational Needs ($M=2.83$, $SD=2.10$).

2.2. Data Collection

Data were derived from video-recorded lessons. Schools were initially approached via email and telephone and interested schools were sent more information, as well as consent forms for teachers and the students' parents. Visits for the video-recordings were then agreed at mutually convenient times. During video-recordings, a camera attached to a tripod was placed in an unobtrusive area of the classroom and two microphones were used for high quality audio: one for the environmental sound and the other attached to the teacher. The teachers were asked to conduct their lessons as normal, and the students were encouraged to ignore the camera. Students with no consent were taken out of class during recordings or seated out of camera range. All 72

lessons were professionally transcribed in verbatim form using a subset of the Jefferson (1984) notation.

2.3. Dialogue Analysis

2.3.1. Coding Dialogic Moves

The extent of approximation to target forms of dialogue was charted using an adapted version of the Scheme for Educational Dialogue Analysis (SEDA: Hennessy, Rojas-Drummond et al. 2016)). The adapted version, called Cambridge Dialogue Analysis Scheme (CDAS), comprised 10 ‘dialogic move’ codes, which are detailed in Table 1 and believed to reflect current views about productive forms. Specifically, the ELI and REI categories captured authentic questions that provoked thoughtful answers (e.g. Nystrand et al. 1997). The EL, RE and Q categories captured core features of exploratory talk (e.g. Littleton and Mercer 2013) and accountable talk (Michaels et al. 2008); namely building on ideas, justifying and challenging respectively. The CI category addressed invitations to synthesise ideas, while SC and RC addressed responses to such invitations, the difference between RC and SC being that RC draws on evidence, theory or a mechanism for justification (Felton and Kuhn 2001; Osborne, Erduran, Simon and Monk 2004). Establishing links and identifying connections, stressed by Alexander (2008) and Michaels et al. (2008), were represented by the RB and RW categories, which focus respectively on prior knowledge or beliefs and the wider context.

Table 1

Descriptions of Dialogic Moves Codes and Agreement Levels

CODES	DEFINITION	COHEN'S KAPPA
Elaboration invitations (ELI)	Invites building on, elaboration, evaluation, clarification of own or another's contribution. E.g. 'I agree with you it makes a strong picture, but what do you picture?'	.62
Elaboration (EL)	Builds on, elaborates, evaluates, clarifies own or other's contribution. E.g. [In reply to 'It's sort of describing how you do it] 'Yes, it's got a good emphasis and a good use of vocabulary'	.63
Reasoning invitations (REI)	Explicitly invites explanation, justification of a contribution or speculation (new scenarios), prediction or hypothesis. E.g. 'Why do you think the bottle floats?'	.73
Reasoning (RE)	Provides an explanation or justification of own or another's contribution, or speculates, predicts, hypothesizes with grounds given. E.g. [After 'He came back'] 'because he made a promise'.	.80
Co-ordination invitations (CI)	Invites synthesis, summary, comparison, evaluation or resolution based on two or more contributions. E.g. 'Would anyone like to summarize the ideas we've been hearing?'	NC
Simple co-ordination (SC)	Synthesises or summarises collective ideas (including own and others' ideas). E.g. 'Some of you are talking about weight and some are talking about size; both matter – things float when they're light for their size'.	.76
Reasoned co-ordination (RC)	Compares, evaluates, resolves two or more contributions in a reasoned fashion. It includes all SC descriptors plus a counter-argument, reasoned rebuttal, two partial truths, e.g. drawing on evidence, theory or a mechanism. E.g. 'We've	NC

	been arguing about how much of personality is inherited; twin studies show conclusively it's 50%'.	
Agreement (A)	Explicit acceptance of or agreement with a statement(s). E.g. 'Brilliant', 'Good'.	.69
Querying (Q)	Doubting, full/partial disagreement, challenging or rejecting a statement. E.g. 'Do you really think these angles are the same?'	.62
Reference back (RB)	Introduces reference to previous knowledge, beliefs, experiences or contributions (includes procedural references) that are common to the current conversation participants. E.g. 'Can anyone remember which of the animals we saw at the zoo are nocturnal?'	.62
Reference to wider context (RW)	Making links between what is being learned and a wider context by introducing knowledge, beliefs, experiences or contributions from outside of the subject being taught, classroom or school. E.g. 'It's like in Macbeth where the storm builds into it'.	.58
Other Invitations (OI)	Invitations of all kinds of verbal contributions (e.g. opinions, ideas, beliefs), except for those coded as EL, REI or CI. This includes invitations on a new topic if this does not fall in another invitation code, and procedural questions.	.72

Two further codes are not directly mappable onto current conceptions of productive dialogue: A and OI. Nevertheless, in combination with ELI or EL, A represents Nystrand et al.'s (2003) high-level evaluation. Nystrand et al. highlight "simple evaluation plus elaboration" and "simple evaluation plus follow-up question" as high-level teacher evaluations of student responses. In our coding system, the first example is captured through the combination of A and EL and the second through the combination of A and ELI. As for OI, this category was included to contrast ELI, REI

and CI with less productive invitations. Twelve lessons were independently coded by two coders, who were drawn at random from the four-strong coding team. Cohen's Kappa values, also presented in Table 1, show acceptable levels of agreement ($> .60$) for all but the RW code, which approached the desirable level.

Binary coding was used to determine the presence or absence of the 12 codes in each turn⁴. Each code could only be used once per turn, regardless of the number of utterances in which it appeared. Coding rules stipulated that, if both EL and RE appeared in the same utterance, then RE would trump EL. If both RW and RB appeared in the same utterance, RW would trump RB. Four codes, namely RB, RW, A and Q, could in principle occur in the same utterance as any of the other codes, and if this happened they were still noted. For the four invitational codes, there was a further distinction between whether the invitation received a reply that was relevant (code 'R') or whether the invitation was ignored (code 'X'). Finally, all turns not represented via the 12 codes were recorded as uncoded (UC).

2.3.2. Rating scales of participation

In addition to the turn-level coding, rating scales represented a lesson's dialogic ethos. The scales are defined in Table 2 and cover student and teacher participation in specified lesson activities. One of these activities, namely 'Student Participation', captures the participative ethos, which involves listening and respecting others' ideas. Each lesson was rated across three levels (0 to 2), with the lowest level indicating that

⁴ A turn is defined here as any contribution that begins and ends with a speaker switch or audience change.

this dimension was not evident, the middle level that it occurred but was teacher-led, and the highest level that there was some student input. Each lesson was judged holistically after viewing the video-recording. Table 2 presents the percentage agreement between coders.

Table 2

Participation Dimensions

DESCRIPTORS	DEFINITION	AGREEMENT (%)
Aims and objectives	The extent to which the lesson aims and objectives are explicit, teacher-led and/or student focused.	75
Monitoring and guidance	The extent to which monitoring and guidance is provided by the teacher throughout the lesson, and the quality of this.	83
Reflection on learning process	The extent to which reflection on learning processes takes place, either being reported by the teacher or discussed with the students.	75
Focusing on talk rules	The extent to which a focus on rules for talk (if present) is introduced by the teacher or negotiated with the students.	92
Student participation	The extent to which students are given the opportunity to express their ideas publicly and engage with other ideas.	92

3. Results

The findings are presented separately here for codes and rating scales.

3.1. Codes

3.1.1. Occurrence of Productive Forms (Move Codes)

Addressing the main research question, Table 3 presents the average frequencies for all codes across the 72 lessons, after correcting for lesson duration (dividing raw frequencies by lesson duration in minutes, and then multiplying by 65.4, the mean duration in minutes of all 72 lessons). The table also presents the average frequencies for teachers and students separately, which will be discussed in Section 3.1.2.

Table 3

Descriptive Statistics for Productive and Non-productive Forms of Dialogue

	Overall		Teachers		Students	
	Mean	SD	Mean	SD	Mean	SD
ELI	29.18	17.66	28.35	17.28	.76	1.45
EL	77.78	37.20	37.92	18.66	39.75	22.42

REI	18.54	13.90	17.41	13.54	1.13	1.65
RE	53.46	20.89	26.05	12.52	27.33	14.86
CI	.01	.10	.01	.10	.00	.00
SC	.47	.80	.46	.80	.01	.12
RC	.12	.42	.06	.25	.06	.35
Q	18.04	12.88	13.98	9.25	4.19	5.20
RB	6.09	6.37	4.79	3.95	1.29	2.95
RW	4.02	4.69	2.94	3.71	1.05	1.94

Non-dialogic forms

	Mean	SD	Mean	SD	Mean	SD
A	77.20	37.27	69.10	33.48	7.11	8.48
OI	135.04	43.42	99.53	39.24	33.10	18.59
UC	246.68	92.70	88.83	46.93	158.80	54.86

Turns involving the supposedly non-dialogic codes (OI+UC) dominated the lessons. In fact, when calculating the percentage frequency of these two codes against the total of nine codes (ELI, EL, REI, RE, CI, SC, RC, OI and UC, i.e. excluding the

potentially crosscutting A, Q, RB and RW), the usage of these two varied greatly but they were dominant in all lessons (min=34.92%, max=89.89%).

Regarding the dialogic codes, the three coordination codes were seldom used, suggesting that synthesis of ideas is rare. However, other dialogic forms were far from insignificant. In particular, ELI and EL (thereafter Elaborated) and REI and RE (thereafter Reasoned), including their equivalent invitation codes, are associated with relatively high frequencies, while (from the *SDs*) clearly also varying enormously in frequency across lessons.

3.1.2. Contribution of teachers versus students (Move Codes)

The 72 lessons covered mathematics, English and science, and subject matter affected code frequencies (although not lessons ratings). One-way ANOVAs revealed significant effects of subject on the frequencies of: 1) ELI, $F(2,69)=4.78$, $p=.011$; 2) EL, $F(2,69)=5.02$, $p=.009$; 3) REI, $F(2,69)=21.01$, $p<.001$; and 4) RE, $F(2,69)=22.19$, $p<.001$. ELI and EL were significantly more frequent in mathematics lessons (for ELI, $M=35.57$, $SD=16.36$; for EL, $M=85.20$, $SD=33.35$) than science (for ELI, $M=20.93$, $SD=13.81$; for EL, $M=59.28$, $SD=30.20$). REI and RE were significantly more frequent in mathematics lessons (for REI, $M=30.47$, $SD=14.67$; for RE, $M=71.75$, $SD=20.27$) than science (for REI, $M=13.78$, $SD=8.93$; for RE, $M=45.69$, $SD=15.21$) or English (for REI, $M=11.38$, $SD=8.71$; for RE, $M=45.69$, $SD=15.21$). However, the consistently large standard deviations suggested substantial differences between participating classes even with subject variation taken into account; and addressing the second research question there was good reason to regard the teachers as the driving forces as regards those differences. In particular, as Table 3 illustrates, virtually all Elaborated and

Reasoned invitations were produced by teachers (97.16% of ELI, and 93.91% of REI). Moreover, most invitations received replies, with only 6.25% of ELIs, and 7.73% of REIs ignored (coded as the 'X' variation).

Additionally, a strong relationship was found between the types of teachers' invitations and students' replies. Specifically, the correlation between teacher ELI and student EL was .81, $p < .001$. Similarly, the correlation between teacher REI and student RE was .86, $p < .001$. These relations suggest that, even when Elaborated and Reasoned dialogue occurred, this was mainly within the classic IRF format. The following examples illustrate this.

Excerpt 1. Reasoned talk in an IRF format from a science lesson

Teacher: So why do we have scientific symbols instead of just the pictures of whatever it is we're using? Samantha?

Samantha: Cos, like if you have a diagram then like you can show the symbols instead of having to draw out the actual like proper light bulb or something, you could use symbols and then write underneath it light bulb.

Teacher: Ok. Do we even need to write underneath it 'light bulb'? Because we know what the symbol is don't we, we did that last time. Ok, so why might then we want to just draw the symbol and not the realistic lifelike drawing of the component itself? Laura?

Laura: It would take too long.

Teacher: It would take too long wouldn't it? So what we're saying is it's much...?

Class: Quicker.

Teacher: Quicker, easier, simpler clearer because we're not all the best

artists, are we? [...]

Excerpt 1 comes from a science lesson on circuits. It begins with a teacher asking students for a reason why scientific symbols are used when drawing a circuit (coded REI). A student, Samantha, provides a reason during the following turn (coded RE). She explains that symbols can be used instead of actual drawings and that the word 'light bulb' can be underneath the symbol. The teacher queries the idea of writing underneath the symbol (coded Q) and continues with another invitation to reason: why not just draw a lifelike drawing of the symbol? (coded REI). Laura explains that that would be too time-consuming (coded RE) and the teacher shows her agreement by repeating what Laura said (coded A).

Excerpt 2 presents an IRF sequence involving Elaborated talk.

Excerpt 2. Elaborated talk in an IRF format from an English lesson

Teacher: OK. So you're using- so direct is directly to one person, whereas indirect might be to a group of people. Interesting. Anybody got anything to add to that? ((Some hands raised))⁵ Jack?

Jack: So direct is like... 'I like cheese,' said Perseus, and then indirect is like, 'Perseus stated that he liked the cheese.'

Teacher: OK, interesting.

Jack: It's more actually saying that- it's not putting anything in speech marks, indirect, whereas with direct you are.

Teacher: Brilliant. Anything else? Can anybody add to what Jack said or to what Chris said? Anybody agree, disagree? What do you think? [Indicating to Chris]

Chris: Indirect is like when the narrator says it or something.

Teacher: Tell me more.

⁵ Brackets indicate non-verbal action

The excerpt starts with the teacher paraphrasing a student's misconception (student's turn was inaudible) about what direct and indirect speech is. Instead of challenging it, the teacher asks if anyone would like to add to that (coded ELI). Jack responds by providing examples of direct and indirect speech (coded EL), which the teacher accepts (coded A). She then invites building on what Jack has said (coded ELI). Chris responds (coded EL) and the teacher asks for more detail (coded ELI).

3.1.3. Role of Prior Professional Development (Move Codes)

The third research question was concerned with the role of prior professional development in promoting productive dialogue. As described in Section 2.1, our teachers were equally divided into two groups, designated PriorPD when they had received prior professional development and NoPD otherwise.

Independent-samples t-tests compared the two groups in terms of the four main dialogic move codes. A difference approaching statistical significance was found for REI, M for PriorPD=21.69, SD =10.36; M for NoPD=15.40, SD =8.53, $t(34)=-1.99$, $p=.055$. A significant difference was found for RE, M for PriorPD=59.01, SD =15.01, M for NoPD=47.91, SD =13.51; $t(34)=-2.33$, $p=.026$. However, non-significant differences were found for ELI and EL. The findings suggest, therefore, that PriorPD may have an impact on Reasoned talk (including inviting reasoning), but not on Elaborated talk.

3.2. Rating Scales

3.2.1. Occurrence of Productive Forms (Rating Scales)

Table 4 presents the mean ratings for each of the five scales across the 72 lessons. As described in Section 2.3.2, ratings ranged from 0 to 2 reflecting the extent to which certain activities took place or not (0), teachers led (1) or students were actively involved (2).

Table 4

Percentages of distribution of ratings

	0	1	2
Aims & objectives	13.89	80.56	5.56
Monitoring & guidance	2.78	29.17	68.06
Reflection on learning process	52.78	18.06	29.17
Talk rules	88.89	6.94	4.17
Student participation	16.67	52.78	30.56

The first scale suggests that setting lesson aims and objectives was largely teacher-led (80.56%). The ‘Monitoring & Guidance’ scale showed more variation. The majority (68.06%) of lessons demonstrated student-involvement, meaning that teachers offered help with student work, but without taking over. Yet in a good proportion (29.17%) of lessons, monitoring was teacher-led, with teachers observing students working and offering suggestions and evaluations. Variation also occurred with the scale addressing reflection on the learning process. In approximately half of the lessons,

no reflection took place. In 18.06% of the lessons reflection was driven by the teacher, and in 29.17% of the lessons students were involved. Regarding ,‘Talk Rules’, these were not mentioned in the majority of the lessons (88.89%). The few remaining lessons were split between teachers reporting on talk rules (6.94%) and discussing them with students (4.17%). Finally, more variation is seen with the Student Participation scale. In more than half of the lessons, students expressed their ideas publicly and at length, and in nearly one-third of the lessons, students were also engaged with each other’s ideas.

3.2.2. Contribution of Teachers versus Students (Rating Scales)

The contribution of teachers and students to the variation in the ratings is reflected in the distributed frequencies of ratings (see Table 4). Teachers clearly contributed more to setting out aims and objectives. There was more balanced initiation across teachers and students for monitoring and guidance, reflection on the learning process and student participation. As noted, there was little focus on talk rules.

3.2.3. Role of Prior Professional Development (Rating Scales)

Lesson ratings across the five scales were compared between teachers with PriorPD and teachers with NoPD (see Table 5). Because two scales failed normality tests (Aims, Talk Rules), non-parametric tests were used (Mann Whitney U tests) and showed no significant difference between the two groups for any scales.

Table 5

Comparison of PriorPD Group and NoPD Group for Rating Scales

Scales	Median	U	Z	p
Aims	1	154	-.30	.766
Monitoring	2	140.5	-.77	.443
Reflection	.75	125.5	-1.19	.233
Talk rules	0	156	-.28	.783
Student participation	1	126.5	-1.17	.243

4. Discussion

4.1. Presence of Productive Forms of Dialogue

The study's main objective was to investigate the occurrence of dialogue forms that are widely regarded as productive. The results indicate relatively high usage of many such forms in primary classrooms, contradicting the impression often given by observational studies in this field: remembering that the mean duration of lessons was 65.4 minutes, EL and RE were both used on average around once per minute. Indeed, while from the turn coding, OI+UC were clearly ubiquitous, the frequency of the target features can never have been lower than 10.11% of total turns and sometimes must have been as high as 65.07%. These figures are particularly significant when scholars in this area argue for maximising productive dialogue where appropriate, not for all turns to involve such forms. Having categorised teacher-student talk along two dimensions in science

classrooms, namely dialogic-authoritative and interactive - non-interactive, Aguiar, Mortimer and Scott (2010) argued that, in exploring ideas, “transitions between dialogic and authoritative interactions [are] fundamental to supporting meaningful learning of disciplinary knowledge” (178). According to them, effective classroom practice does not preclude the occurrence of teacher-centred interactions (based on closed question IRFs and authoritative presentations) but rather involves judicious use in conjunction with more ‘dialogic’ interactions in order to allow significant student involvement in meaning making.

Looking at the dialogic moves more closely, the codes with the highest average frequency (see Table 3) were Elaborated (ELI, EL), Reasoned (REI, RE) and Querying (Q). This finding resonates with other research that pinpoints these as key features of productive dialogue. Building on ideas, providing reasons or evidence, and challenging ideas have, for instance, been highlighted repeatedly via such constructs as ‘exploratory talk’ and ‘accountable talk’ (Littleton and Mercer 2013; Michaels et al. 2008). The results from the frequencies of the ‘Monitoring & Guidance’ and ‘Student Participation’ scales also support this result.

An interesting finding was the low occurrence of certain forms of dialogue. The three coordination codes (CI, SC, RC) were rarely used. As synthesis and connection of ideas are marked as important features in the literature because they capture accumulation (Alexander 2008; Hennessy et al. 2016), they were expected to occur especially after brainstorming activities. Potential sources of challenge may be keeping track of multiple ideas from students that do not necessarily occur in sequence. Similarly, the reference back and reference to the wider context codes (RB, RW) also rarely occurred. Particularly referring to the latter, the concept of “semantic waves”

represents the key to cumulative development of educational knowledge over time, as it refers to the shifting between ‘context-dependent and simplified meanings’ (equivalent to references to the wider context) and ‘decontextualized and condensed knowledge’ that students need for assessment (Maton 2013, p. 9). The highly sophisticated functions of making connections emerging here may need to be boosted in PD programmes. More research would shed light on this issue.

As regards the rating scales, at first glance the frequencies of ‘Aims & Objectives’ and ‘Talk Rules’ suggest a non-dialogic environment. However, these findings are unsurprising because teachers are required to set lesson objectives by the national inspection agency for England (Office for Standards in Education). Regarding talk rules, some teachers may not be familiar with the recent ‘initiative’ (e.g. in Littleton et al. 2010) of setting ground rules for talk. In addition, even if teachers do use talk rules, it could be that over time students get accustomed to the dialogic ethos and might not need to be continually reminded of rules.

4.2. Contribution of Teachers & the Role of Professional Development

Having discussed which dialogue forms were frequent on average, this section focuses on variation across the 36 classrooms. Features that were low on average, such as the coordination codes, were of consistently low frequency across classrooms. The frequency of the more prevalent features, however, like elaboration, showed considerable variation. Looking at the turn coding, the total percentage of OI+UC varied from 34.93% to 89.89% in a single lesson. This suggests that the frequency of the target features ranged from 10.11% to 65.07% of total turns, hinting at the discovery of what we called ‘pockets of excellence’. This suggests, therefore, that teachers are the

driving forces as regards the key forms of dialogue, not only because they produce the vast majority of Elaborated and Reasoned invitations, but also because the types of student reply are highly correlated with the types of teacher invitations. Consistent with Macbeth (2011), this points to embedding within IRF sequences, a conclusion supported through the qualitative examples in Section 3.1.2. Whatever the case, the finding certainly highlights the teacher's power to shape classroom dialogue. When teachers use "model" forms, they are likely to trigger "model" dialogue.

Teachers are major drivers despite subject differences, but the driver of their behaviour was at best only partially PriorPD. As shown by the final research question, PD may have contributed to the variation of Reasoned but it certainly did not bear on the variation of Elaborated. One possible interpretation is that uptake of elaboration is harder for teachers in response to PD, perhaps because it is not such a 'self-defining' act as reasoning; reasoning is more strongly associated with cue words (e.g. 'because') and so it is arguably more salient (Hennessy et al. 2016). The higher saliency of Reasoned over Elaborated also became evident to us during the inter-coder reliability process, as agreement for the RE category was the highest of all our codes (see Table 2). Alternatively, this finding may suggest that current PD programmes have a more explicit focus on reasoning (e.g. 'Thinking Together' programme, Mercer and Littleton 2007). More emphasis may need to be placed on elaboration in such programmes, given its importance in the literature. Nevertheless, the variation of Elaborated, regardless of PD, could be due to teachers' beliefs about how children learn. Specifically, it might be the case that some teachers believe in more elaborated responses to their students as a more engaging technique for learning.

Whatever the case, however, the important finding here is that good practice

can occur in the absence of PD. More evidence is needed in order to consolidate this paradoxical relationship. Only then, issues of scalability and sustainability of PD programmes can be resolved.

4.4. Conclusions

Our rigorous, systematic analysis of reasonably representative data from a diverse sample of English primary schools allows some important conclusions to be drawn. The data showed that the talk commonly had a significantly dialogic component, with high frequencies of elaborated and reasoned talk. Forms with high frequencies, however, also showed considerable variation across classrooms, revealing ‘pockets of excellence’ in our data. This finding makes a significant and original contribution to the field, as it contradicts the impression that observational studies have given to date. Teachers played an important role in enabling such dialogue and so there are important implications for teacher education and teacher professional development in helping practitioners understand the importance of their own use of high quality talk during teaching. More research is required, however, on the exact effects of varied PD programmes on the use of dialogue.

Acknowledgements

We are grateful to the teachers and students who participated in this study, and to the anonymous reviewers for their helpful comments on an earlier draft. This work was supported by the ESRC under Grant ES/M007103/1.

References

- Aguiar, O. R., E. F. Mortimer and P. Scott. 2010. "Learning from and responding to students' questions: The authoritative and dialogic tension." *Journal of research in science teaching* 47 (2): 174-193.
- Alexander, R. 2008. *Towards dialogic teaching: Rethinking classroom talk*. Cambridge: Dialogos.
- Alexander, R., F. Hardman, J. Hardman, with T. Rajab and M. Longmore. 2017. Changing talk, changing thinking: Interim report from the in-house evaluation of the CPRT/UoU dialogic teaching project. Accessed 10 September 2017. <http://www.robinalexander.org.uk/wp-content/uploads/2017/07/Alexander-et-al-EEF-in-house-interim-report-final-170714.pdf>
- Berry, R.A.W., and N. Kim. 2008. "Exploring teacher talk during mathematics instruction in an inclusion classroom." *The Journal of Educational Research* 101 (6): 363–378.
- Bleicher, R.E., K.G. Tobin, and C. J. McRobbie. 2003. "Opportunities to talk science in a high school chemistry classroom." *Research in Science Education* 33 (3): 319–339.
- Chinn, C. A., R. C. Anderson, and M. A. Waggoner. 2001. "Patterns of discourse in two kinds of literature discussion." *Reading Research Quarterly* 36 (4): 378-411.
- Dawes, L., N. Mercer, and R. Wegerif. 2003. *Thinking Together: a programme of activities for developing speaking, listening and thinking skills for children aged 8-11*. Birmingham; Imaginative Minds Ltd.

- Department for Education. 2013. *The national curriculum in England: Key stages 1 and 2 framework document*. Report no. DFE-00178-2013.
- Felton, M. & D. Kuhn. 2001. The development of argumentive discourse skill, *Discourse Processes* 32 (2-3): 135-153.
- Haneda, M., A. Teemant, and B. Sherman. 2017. "Instructional coaching through dialogic interaction: helping a teacher to become agentic in her practice." *Language and Education* 31 (1): 46-64.
- Hennessy, S., T. Dragovic, and P. Warwick. 2017. "A research-informed, school-based professional development workshop programme to promote dialogic teaching with interactive technologies." *Professional Development in Education*, 1-24.
- Hennessy, S., S. Rojas-Drummond, R. Higham, A. M. Márquez, F. Maine, R. M. Ríos, R. García-Carrión, O. Torreblanca, and M. J. Barrera. 2016. "Developing a coding scheme for analysing classroom dialogue across educational contexts." *Learning, Culture and Social Interaction* 9: 16-44.
- Howe, C., and M. Abedin. 2013. "Classroom dialogue: A systematic review across four decades of research." *Cambridge Journal of Education* 43 (3): 325-356.
- Howe, C. and N. Mercer. 2017. "Commentary on papers." *Language and Education* 31 (1): 83-92.
- Jefferson, Gail 1984. "Transcription notation." In *Structures of social interaction*, edited by J. Maxwell Atkinson and J. Heritage, ix-xvi. New York: Cambridge University Press.
- Lefstein, A. and J. Snell. 2014. *Better than best practice: Developing teaching and learning through dialogue*. London: Routledge.

- Littleton, K. and N. Mercer. 2013. *Interthinking: Putting talk to work*. Routledge.
- Littleton, K., N. Mercer, L. Dawes, R. Wegerif, D. Rowe, and C. Sams. 2010. "Talking and thinking together at Key Stage 1". *Early Years* 25 (2): 167-182.
- Macbeth, D. 2011. "Understanding understanding as an instructional matter". *Journal of Pragmatics* 43: 438–451.
- Maton, K. 2013. "Making semantic waves: A key to cumulative knowledge-building." *Linguistics and Education* 24 (1): 8-22.
- Mercer, N. and K. Littleton. 2007. *Dialogue and the Development of Children's Thinking*. London: Routledge.
- Michaels, S., C. O'Connor, and L. B. Resnick. 2008. "Deliberative discourse idealized and realized: Accountable talk in the classroom and civic life." *Studies in Philosophy and Education* 27 (4): 283–297.
- Nystrand, M., A. Gamoran, R. Kachur, and C. Prendergast. 1997. *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Teachers College Press.
- Nystrand, M., A. Gamoran, S. Zeiser, and D.A. Long. 2003. "Questions in time: Investigating the structure and dynamics of unfolding classroom discourse." *Discourse Process* 35 (2): 135-198.
- Osborne, J., S. Erduran, S. Simon, and M. Monk. 2004. "Enhancing the quality of argument in school science." *School Science Review* 82 (301): 63-70.
- Pehmer, A. K., A. Gröschner, and T. Seidel. 2015. "Fostering and scaffolding student engagement in productive classroom discourse: Teachers' practice changes and reflections in light of teachers professional development." *Learning, Culture and Social Interaction* 7: 12-27.

- Pontefract, C. and F. Hardman. 2005. "The discourse of classroom interaction in Kenyan primary schools." *Comparative Education* 41 (1): 87–106.
- Ruthven, K., N. Mercer, K. S. Taber, P. Guardia, R. Hofmann, S. Ilie, S. Luthman, and F. Riga. 2017. "A research-informed dialogic-teaching approach to early secondary school mathematics and science: the pedagogical design and field trial of the epiSTEMe intervention." *Research Papers in Education* 32 (1): 18-40.
- Sedova, K., M. Sedlacek, and R. Svaricek. 2016. "Teacher professional development as a means of transforming student classroom talk." *Teaching and Teacher Education* 57: 14-25.
- Sepeng, P. 2011. "Triadic dialogue: An analysis of interactions in multilingual mathematics primary classrooms." *US-China Education Review* 8: 412–418.
- Sinclair, J.Mc.H. and M. Coulthard. 1975. *Towards an analysis of discourse: The English used by pupils and teachers*. Oxford: Oxford University Press.
- Vygotsky, L. 1962. *Thought and language*. Cambridge MA: MIT.
- Wells, G. and R. M. Arauz. 2006. "Dialogue in the classroom." *The Journal of the Learning Sciences* 15 (3): 379-428.
- Wilkinson, I. A. G., A. Reznitskaya, K. Bourdage, J. Oyler, M. Glina, R. Drewry, M. Kim, and K. Nelson. 2017. "Toward a more dialogic pedagogy: changing teachers' beliefs and practices through professional development in language arts classrooms." *Language and Education* 31 (1): 65-82.