Learning, Culture and Social Interaction 43 (2023) 100742

Contents lists available at ScienceDirect

Learning, Culture and Social Interaction

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

Disciplinary dialogues: Exploring the association between classroom dialogue and learning outcomes within and between subjects in English primary schools

Annabel Amodia-Bidakowska *, Sara Hennessy, Paul Warwick

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

ARTICLE INFO

Keywords:
Disciplinary context
Subject teaching
Dialogue
Coding
Primary school

ABSTRACT

There is a large literature on the types of classroom dialogue that are theorised to be productive for learning; however, there is very little evidence for whether these forms of dialogue are more common, applicable or impactful in certain subjects. This mixed-methods study explored how dialogue unfolds in different curriculum contexts. Classroom dialogue was examined through naturalistic observations of English, mathematics and science lessons involving children aged 10–11 in primary schools in England. The Cambridge Dialogue Analysis Scheme (Vrikki, Wheatley, Howe, Hennessy, & Mercer, 2019) was used to investigate the frequency of key dialogue features in 72 lessons. Statistical analysis identified that reasoned dialogue is more frequent in mathematics and both English and mathematics have a proclivity towards more elaborated dialogue (clarifying/building on ideas) in comparison to science. Furthermore, relationships between student attainment and subject dialogues were examined in 63 lessons through multilevel modelling, revealing that elaboration of own and others' ideas (in conjunction with reference to the wider context) in English was associated with student attainment in reading and spelling, punctuation and grammar. No other significant associations between subject dialogues and student attainment emerged. The findings have highlighted that teacher professional development on dialogue should be sensitive to the disciplinary context.

1. Introduction

In any lesson (in the context of schooling in England), instances of no interaction are extremely rare. Throughout the lesson, teachers are communicating with their whole class, small groups or individual students. Additionally, students are often engaging in discussions without the presence of the teacher, either in groups or dyads. All of these interactions that occur in lessons shape the individual's learning experience. Research into educational dialogue aims to examine how to amplify these interactions to optimise students' learning outcomes; thus, it has for many decades been a central focus of attention in the field of educational research (Mercer & Dawes, 2014).

A systematic review of 225 studies on classroom dialogue identified a substantive body of research examining optimal forms of dialogue in group work settings but a lack of rigorous, empirical research in whole-class contexts; that is, teachers interacting with

* Corresponding author.
E-mail addresses: annabelamodia-bidakowska@cantab.net (A. Amodia-Bidakowska), sch30@cam.ac.uk (S. Hennessy), ptw21@cam.ac.uk (P. Warwick).

https://doi.org/10.1016/j.lcsi.2023.100742
Received 18 February 2023; Received in revised form 2 July 2023; Accepted 22 July 2023
Available online 18 September 2023
2210-6561/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

their whole class of students (Howe & Abedin, 2013). Furthermore, existing research has emphasised the need to investigate dialogue through a contextually sensitive approach (Littleton & Howe, 2010). We need not only to understand the dialogue features that are particularly pertinent to enhanced attainment in disciplinary dialogues, but also how they may be part of the complex web of contextual factors that impinge on attainment in primary classrooms. This article considers one of these important contextual factors in particular, namely the disciplinary context. How might disciplinary norms and traditions be shaping the interactions we observe in lessons? How are subject curriculae impacting on the nature of dialogue? How are disciplinary dialogues in turn linked to student attainment? These are some of the questions that align with a contextually sensitive approach to the study of dialogue and deserve attention immediately.

Research on disciplinary differences in dialogue is remarkably scarce. To date, we know very little about how dialogue unfolds in different subject lessons and how this might be linked to attainment. This article addresses this important research gap. The work aims to support teachers to develop dialogic teaching in curriculum subject areas by developing understanding of what is appropriate and beneficial in each discipline. This was examined through naturalistic observations of 72 English, mathematics and science lessons with teachers and students aged 10–11. Systematic turn-by-turn dialogue coding of lesson transcripts enabled the transformation of recorded classroom interaction to quantitative data that could be subjected to rigorous statistical analysis in order to test specific hypotheses, account for confounding factors (e.g., prior attainment) and maximise generalisability.

This study makes a significant contribution to the field by providing rigorous quantitative evidence that the disciplinary context of the lesson impacts on the dialogue that takes place at this age, which, in turn, has implications for student attainment. This research ultimately informs more contextually relevant teacher professional development (TPD) guidance concerning dialogic teaching.

2. Literature review

The last 45 years have seen a surge of educational research on classroom interaction (Mercer & Dawes, 2014), but the notion of 'dialogue' is not clear-cut. Whilst some in the field of education conceptualise dialogue as any exchange between two or more individuals, the majority present dialogue as a specific form of interaction that goes beyond talk, conversation, discussion etc. Thus, educational dialogue can be conceptualised as an exchange between two or more individuals where participants are able to shift perspectives to understand others' views (Wegerif, 2011), to respond and elaborate on ideas, and to co-construct and negotiate meaning (Cook, Warwick, Vrikki, Major, & Wegerif, 2019). Educational dialogue may also be conceptualised as a shared enquiry (Wells, 1999) as it is cumulative and cooperative (Alexander, 2008), resulting in the emergence of new ideas that are shaped by multiple participants (Littleton & Mercer, 2013) and it is purposeful as it can be used to achieve key learning goals (Alexander, 2008). These perspectives on dialogue underpin the work presented here.

A number of different features of dialogue are typified as optimal for learning (as reviewed by Howe, Hennessy, Mercer, Vrikki, & Wheatley, 2019), including evaluating competing opinions (Lefstein, 2010; Mercer & Littleton, 2007), querying, challenging and presenting counter-arguments (Larrain, Howe, & Cerda, 2014; Michaels, O'Connor, & Resnick, 2008; Omland, Ludvigsen, & Rødnes, 2022), making extended responses (Alexander, 2008; O'Connor, Michaels, & Chapin, 2015), coordinating different perspectives (Alexander, 2008; Reznitskaya & Gregory, 2013), open-ended questioning (Alexander, 2008; Nystrand, Wu, Gamoran, Zeiser, & Long, 2003), and reasoning (Mercer, Dawes, Wegerif, & Sams, 2004). Proficient teachers not only model these optimal features but also encourage these forms of dialogue among their students (Boyd & Kong, 2017).

Sociocultural perspectives (e.g. Mercer, 1995; Vygotsky, 1978) and dialogic theories (e.g. Wegerif, 2011, drawing on Bakhtin, 1986) have offered some insight into how dialogue may impact upon student learning processes. However, evidence to suggest that specific forms of dialogue lead to accelerated learning outcomes in the core subjects is woefully limited. To examine the extent to which the research field has achieved this goal, Howe and Abedin (2013) carried out a systematic review of 225 studies on classroom dialogue which revealed that the forms of dialogue that were productive for learning were generally not isolated from other potentially confounding factors. Furthermore, whilst a number of small-scale studies had explored dialogue interventions in group-work settings, there was limited research regarding the potential benefits of whole-class dialogue on student learning (Howe & Abedin, 2013).

This has since been addressed by Howe et al. (2019) who examined the link between teacher-student dialogue and curriculum mastery at primary level through a naturalistic study. Overall, the authors reported that when taking into account other factors of potential influence (such as prior attainment, demographics, teaching practices etc.), two core dialogue features, elaboration (including clarification/building) and querying (i.e. "doubting, full/partial disagreement, challenging or rejecting a statement", p. 51) acquired through turn based coding, were positively associated with curriculum mastery when (and only when) combined with a third element: a classroom ethos of high student participation (i.e. "opportunities for students to express ideas and engage with the ideas of others", p.19). Likewise, Alexander (2018) randomised control trial revealed modest gains in 9- and 10-year-old students' attainment (in mathematics, science and English) after receiving an intervention that promoted 61 indicators of dialogic teaching (including key features of dialogue and the contextual factors/conditions that enable dialogue); students who had received the intervention over a 4.5-month period exhibited a 2-month attainment gain compared to controls.

Furthermore, Howe and Abedin (2013) systematic review highlighted that whilst a number of additional factors were explored in relation to variation in dialogue (such as gender, school type etc.), none of these had been examined in great depth; with the exception of a few small-scale studies (e.g., Nassaji & Wells, 2000, discussed in 2.2), the curriculum context was an area that had been particularly neglected in the field. The authors highlighted the need for methodological change in the field by incorporating quantitative methods to statistically examine dialogue at scale in relation to outcomes. They additionally flagged up the need for a contextually sensitive approach to studying educational dialogues; since the content of the subject matter is of particular "significance and consequence" (Littleton & Howe, 2010, p.84), this study focusses on a key contextual factor pertaining to educational dialogue:

curriculum subject.

The review of literature presented below explores the extent to which components of classroom dialogue have the same applicability, emphasis and use in three core subject areas: English, mathematics and science. This focus reflects the scope of the dataset collected by Howe et al. (2019) that formed the basis for the study reported in this article.

2.1. Classroom dialogue in science

The nature of classroom dialogue has been examined across a number of different curriculum contexts. However, science has been given particular attention by educational researchers, perhaps due to science curriculum policy pressure to move away from the traditional presentational style of science teaching that previously dominated (Mortimer & Scott, 2003). For example, in Newton, Driver, and Osborne (1999) study, a total of 34 biology/chemistry/physics lessons were observed and they found that only 2 % of these science lessons were dedicated to discussion.

Argumentation and reasoning in the teaching of science have been examined in great depth by psychologists, developmentalists and educationalists, and a large body of research suggests that they should be promoted in the teaching of science (Asterhan & Schwarz, 2007; Driver, Newton, & Osborne, 2000; Erduran, Ardac, & Yakmaci-Guzel, 2006; Howe, 2010; Howe et al., 2007; Howe et al., 2015; Howe & Tolmie, 2003; Larrain, Freire, & Howe, 2014; Larrain, Howe, & Freire, 2018; Zohar & Nemet, 2002). Lemke (1990) explains that the discussion of science is not merely a discussion about scientific facts but, instead, "talking science means observing, hypothesising, describing, comparing, classifying, analysing" (p.1). Argumentation is a central component of these skills, helping students to think critically and provide reasons and explanations (Osborne & Chin, 2010). Additionally, researchers assert that argumentation aids students' conceptualisation of scientific ideas (Kuhn, 2010; Von Aufschnaiter, Erduran, Osborne, & Simon, 2008). The complexity of scientific concepts requires in-depth discussion including hypothesising, predicting and analysing through reasoning and interpretation, which are crucial for learning science (Hennessy et al., 2007). Ludvigsen and Steier (2019) state that "Educational dialogues need to be connected to epistemic orientations for students to appropriate an understanding of complex tasks in and across knowledge domains" (p.419). Therefore, for students to have authentic learning experiences, the teaching of a subject should, to some extent, replicate the subject culture and mirror its epistemic practices. It can be argued that, in science communities, the culture is driven by providing evidence, clarifications and elaborations, and stating reasoning or justifications for claims; therefore, a culture that is inherently dialogic.

Due to the focus on teacher-student dialogue in this study, an example of a whole-class dialogue intervention where epistemic practices are modelled and scaffolded is highlighted here. Larrain et al. (2014) administered an intervention to optimise whole-class argumentation with 11 teachers and their classes (aged 10–11) and compared attainment gains to a control group. Teachers asked for justifications for student predictions and students displayed scientific reasoning; for example, "on the smooth surface because on the other the wheels have more traction" (p.8). Overall, the results indicated that providing justifications and counter-arguments predicted students' science learning gains on a delayed post-test (one month after the intervention).

The literature on classroom dialogue has also emphasised the importance of students' ability to elaborate/build on their own and others' contributions across disciplinary contexts. However, close examination of the concept of 'elaboration' in science shows that it is intrinsically linked to reasoning, i.e. student elaboration in science involves students "elaborating on their reasoning" (e.g. Schindler, Gröschner, & Seidel, 2015, p.13). Therefore, it is difficult to distinguish whether the benefits of elaboration are attributable to building on ideas or the use of reasoning.

2.2. Classroom dialogue in mathematics

In the last 20 years research has delved into how argumentation can aid mathematics activities by encouraging mathematical reasoning (Schwarz, Hershkowitz & Prusak,

2010). The National Council of Teachers of Mathematics (2000) emphasises the need for students of all ages to engage in mathematical argumentation. Transcripts of mathematics lessons have shown that argumentation and reasoning take central positions in mathematics teaching as they aid students to make sense of mathematical problems and make conjectures, leading to enhanced mathematics learning (Kyriacou & Issitt, 2008; Schwarz, Hershkowitz, & Prusak, 2010). Often, to gauge a student's understanding in mathematics, teachers ask students to provide explanations of their reasoning (Barmby, Harries, Higgins, & Suggate, 2009).

Given the importance associated with argumentation and reasoning in mathematics teaching it is crucial to examine what mathematics argumentation looks like in practice and whether the incorporation of this type of dialogue leads to accelerated learning outcomes. Mercer and Sams (2006) employed the Thinking Together programme with 406 students aged 9–10 years old to encourage productive forms of dialogue in mathematics. The programme successfully aided students to use language as a tool for reasoning in mathematics lessons. For example, when students were required to hypothesise about the functions applied to numbers, students applied reasoning, e.g. "*I think it's halving because we had 6, and it ended up 3. Now we've got 4 and it ended up in 2*" (ibid., p.523). Overall the programme led to increased reasoning during group discussions, which in turn resulted in increased individual gains in mathematics performance. Mercer and Sams (2006) concluded that the teacher plays a pivotal role in modelling effective language use for mathematical reasoning in the classroom.

A recent study by Webb et al. (2019) closely examined how teachers foster student participation and collaboration. The researchers focused specifically on the nature of mathematics thinking in students aged 8–10 years and examined links between the content of interactions and student attainment. Analysis revealed that when controlling for prior knowledge, students' explanations of their own and others' ideas were predictive of learning outcomes. The findings indicated higher achievement among students who provided

complete, correct and unambiguous explanations. Furthermore, the authors discussed the centrality of "challenging others with mathematical reasons" (p.180); critically engaging with others' ideas by questioning/rejecting/challenging, alongside reasoning to justify standpoints, was predictive of learning outcomes.

The importance of elaboration when discussing mathematical problems with peers is also apparent in the literature, particularly when identifying errors in others' solutions to problems (Yackel, Cobb, & Wood, 1991). Elaboration of one's own idea in mathematics could prompt students to think about specific elements of a mathematics problem and internalise rules for problem solving (Webb, 2009). In turn this may help individuals identify areas of misconception and gaps in understanding which could be filled by other participants' elaboration (Veenman, Denessen, Van Den Akker, & Van Der Rijt, 2005; Webb & Farivar, 1999). Veenman et al. (2005) implemented a teacher training programme on promoting effective helping behaviours among students (aged 11–12) during a collaborative mathematics task. The frequency of elaboration during the task was significantly correlated to performance on the mathematics task. However, it is important to consider that studies on elaboration in mathematics often incorporate reasoning. For example, in the above study, their examples of requesting explanations or providing extended explanations often included reasoning, e.g. "It will balance that way because, here, exactly in the middle, are 3 kilograms and there are 6" (Veenman et al., 2005, p.130). Therefore, like in science, it cannot be said whether building on ideas in the absence of reasoning positively impacts students' learning in mathematics.

2.3. Classroom dialogue in English

The nature of classroom dialogue has many similarities in science and mathematics. However, the literature highlights a difference in the emphasis placed on argumentation/reasoning and elaboration in English, particularly in literacy lessons where it may be more akin to "possibility thinking" (Craft, Cremin, Burnard, Dragovic, & Chappell, 2013). One example is a study by Maine (2013) who analysed recordings of 8 pairs of primary school children to qualitatively explore the nature of dialogue when discussing non-written texts. She argues that when children are discussing texts this form of language is tentative and involves questioning and inviting "modification and surmicse" (p.9). Teachers encourage questioning and elaboration, in contrast to invitations for one correct answer that limits further discussion (Maine, 2013). This was evident in students' responses as students expressed themselves creatively to justify their views and were not driven to identifying the 'correct' answer. Students used questioning and made suggestions with tentative phrases such as 'I wonder', 'maybe' and 'I think', which creates an open dialogue where children were able to explore diverse ideas and opinions freely (Boyd, Chiu, & Kong, 2019; Boyd & Kong, 2017). This study provided valuable insight into the nature of dialogue which incorporates elaboration and justification of views within an open and creative discourse during the discussion of texts.

Reasoning may be less applicable in literacy. A study that examined reasoning in relation to learning outcomes in literacy was carried out by Rojas-Drummond, Mazón, Fernández, and Wegerif (2006) who explored whether an 'Exploratory Talk' intervention (designed to promote students' use of reasoning) was beneficial across activities and educational contexts including a reasoning task and a literacy task. A comparison of pre- and post-discussions revealed that the intervention resulted in greater use of student reasoning in the reasoning task, which in turn resulted in performance gains; however, students did not display any reasoning in the pre- or post-literacy task. Instead the students showed greater use of elaboration, reformulating and building on the information presented (Rojas-Drummond et al., 2006). The authors concluded that reasoning was advantageous for the reasoning task as it required one correct answer to complete each question of the task but reasoning was not applicable to the literacy task, which was more open-ended and creative. Additionally, they stated that the promotion of reasoned contributions may hinder discussions in creative open-ended tasks as students may be unwilling to present ideas that are not grounded in evidence. Therefore, dialogue intervention that may in fact hinder student contributions in certain activities is problematic. For example, writing stories may require flights of elaborative imagination relying on the internal logic of the story which is a stark difference from task requirements in science and mathematics. As such, the authors concluded that future research in this area should focus on which interaction styles are productive for learning in different tasks, fields of knowledge and educational contexts.

This section has provided a brief overview of literature on dialogue in science, mathematics and English, and highlighted that argumentation and reasoning are apparent across all three disciplines but to different degrees. Argumentation and reasoning are at the core of discussions in science education as they aid the learner with identifying and evaluating evidence to support or refute a theory. Argumentation and reasoning are also central to dialogic practices in mathematics as a way of helping the learner make sense of mathematical problems and reason about the best strategies to solve them. Although argumentation and reasoning are also viewed as optimal in English discussions (such as sourcing evidence from text), other factors, such as creativity, are equally prominent. Here, an individual can express his/herself freely in open discussion without being overburdened to find the 'correct' answer with evidence to back it up, as one would expect in science and mathematics. Elaboration and building on ideas take a more central role in literacy and, unlike with mathematics and science, effective elaboration is not dependent on the inclusion of reasoning. In fact, as discussed above, reasoning may not be applicable in more open-ended literacy tasks (Rojas-Drummond et al., 2006). This therefore indicates that features of productive dialogue in different curriculum contexts may vary.

2.4. Cross-subject comparison of classroom dialogue

While there is some research on optimal forms of dialogue *within* each of the core disciplines, direct comparison of dialogue *between* subjects has rarely been given attention. In the review on classroom dialogue by Howe and Abedin (2013), 30 % of studies on dialogue did not specify which lesson subject was examined. The authors emphasised the need for future research to examine the curriculum context in relation to classroom dialogue variation. The few key articles that have discussed disciplinary differences in dialogue are

presented below.

The first study, by Nassaji and Wells (2000), examined whole-class discourse in science and arts subjects (predominantly literature but some history lessons) in a range of classes (students aged 7–14). The study developed a coding scheme for analysing dialogue containing 173 categories, categorising different aspects of classroom talk and the context they occurred. The types of teacher invitations and follow-ups were compared between curriculum contexts. When examining specific activity orientations such as 'reviewing' or 'constructing,' the frequency of 'Known Information Questions' ('fact, rule-governed answer', p.40) was higher in science than in arts lessons. However, 'Negotiatory Questions' ('opinion and explanation', p.40) were more frequent in arts than in science lessons (Nassaji & Wells, 2000). Overall, the authors concluded that there were no substantial differences in the styles of discourse between science and arts subjects (Nassaji & Wells, 2000; Wells & Arauz, 2006). However, due to the small sample of teachers ($n = 9$) and episodes observed ($n = 44$), in addition to the inequitable proportions of arts episodes (23 %) in comparison to science (77 %), there is a strong case for further research into subject differences in classroom dialogue.

Wells and Arauz (2006) argued that the ways in which teachers create opportunities for discussions to be dialogic differ according to disciplines. The authors reported on the work of O'Connor (2001), who emphasised the importance of position-driven discussion, where teachers present a statement or question and students are encouraged to state whether they agree or disagree and provide a justification for their position. Wells and Arauz (2006) argue that this could be appropriate in some science or mathematics activities, where one correct conclusion can be reached based on analysis of evidence. However, the authors also note that a 'position-driven' approach may be less common in the arts subjects, where a single correct conclusion may not be identifiable. In practice, this distinction may not always be as clear cut as discussed above; indeed, position-driven discussions may be inappropriate in many mathematics activities and there are some activities in English (such as distinguishing between fact and opinion statements) where a position-driven discussion may be appropriate. This research does point out, however, that there may be, to some degree, disciplinary differences in the types of discussions promoted by teachers. As such, Wells and Arauz (2006) state that future research needs to identify subject-appropriate approaches to encouraging dialogic interaction.

Furthermore, an examination of whole-class interaction in primary literacy and numeracy lessons was carried out by Smith, Hardman, Wall, and Mroz (2004). They set out to examine differences in teachers' discourse strategies between the two subjects by exploring the rate of various discourse moves. Their findings indicated that literacy lessons contained significantly more uptake questions and longer teacher explanations compared to numeracy lessons. Additionally, open questions took up a larger proportion of literacy lessons than numeracy lessons which contained more discourse moves coded as 'direction'. However, these findings should be interpreted with caution as uptake and open questions were very rare across the whole sample (constituting only 4 % and 10 % of questioning exchanges respectively).

Finally, a study by Muhonen, Pakarinen, Poikkeus, Lerkkanen, and Rasku-Puttonen (2018) qualitatively explored the quality of whole-class dialogue during 11 language arts or physics/chemistry lessons with teachers and students aged 12 years old. Teacher-initiated and student-initiated dialogue was rated as either high quality or moderate quality based on Muhonen, Rasku-Puttonen, Pakarinen, Poikkeus, and Lerkkanen (2016) patterns of dialogic teaching criteria. High quality dialogue was characterised by more open-ended questioning, scaffolding, student participation in formulating own initiatives/questions and teachers' expansion and co-ordination of student ideas. Overall, the dialogue observed in physics/chemistry lessons was rated higher quality than dialogue during language arts lessons which contained less teacher support for student participation in discussion. Dialogue in physics/chemistry lessons were rated as high quality due to teachers' use of authentic questions, the incorporation of student contributions in discussions and teachers allowing students to shape the direction of whole-class interaction. This study provides a useful first step in examining potential disciplinary differences in the quality of dialogue through qualitative methods. However, once again these subject differences must be interpreted with caution as only two physics/chemistry lessons were explored in this analysis (compared to nine language arts lessons).

The research on cross-subject differences in dialogue, although limited, provides an indication that classroom dialogue, in particular the nature of teacher invitations and student responses, may be partially disciplinary; however, as very little research has examined the impact of subject type on any aspects of dialogue no firm conclusions can be drawn, highlighting the need for the existing study. Our first research question (RQ1) aimed to tackle this gap by investigating: ***To what extent does the frequency of dialogue features differ between English, mathematics and science lessons?*** The second research question (RQ2) examined the link between subject dialogues and attainment which has also been identified in the field as an area for further investigation: "it remains unclear whether some modes of dialogic organisation 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)" (Chan, 2020, p.2). Therefore, this project also explored the following research question (RQ2): ***Are there subject-specific features of dialogue associated with high attainment?*** Sub-questions were also formulated, which appear in Section 4.

3. Methodology

3.1. Sample

This study was framed within a large-scale ESRC-funded project (Howe et al., 2019) which recruited a sample of 72 teachers and their classes (students aged 10–11: the final year of primary school) from 48 state primary schools in England. The sample was demographically diverse; it contained classes with 0–100 % of students eligible for free school meals ($M = 19.3$ %) and 0–96 % from minority ethnic backgrounds ($M = 32.6$ %). The sample was also geographically diverse, spanning seven counties in the North, East and Southeast of England. All included classes were mixed-sex with an average of 28 pupils (range: 20–36) per class (Howe et al., 2019).

3.2. Procedures

3.2.1. Phase 1

Two researchers administered the National Foundation for Educational Research (NFER, Centre for Assessment, 2016) literacy and mathematics assessments[1] to pupils at the start of a school year. The collection of prior attainment scores in English and mathematics enabled the inclusion of prior attainment as a potential confounding variable. Previous research has indicated that dialogue features may be positively associated with curriculum mastery when taking into account other factors such as prior attainment (Howe et al., 2019). Half of the class were randomly assigned to complete the literacy assessment and the other half of the class completed the mathematics assessment, in order to minimise test fatigue. This provided a set of baseline achievement scores. Teachers were asked to complete a questionnaire (see Section 3.3).

3.2.2. Phase 2

Prior to the observation sessions each researcher attended a video recording training session. The researchers were instructed to video as much of the class as possible during whole-class discussions and follow the teacher during group work.

Each participating teacher and his/her class were videotaped for two/three one-hour lessons in at least two different subjects (English, mathematics or science), depending on teacher preferences. As this was a naturalistic study, the researchers did not specify what teachers should cover during lessons. Observation data included lessons with a range of foci matching the national curriculum and standardised test content in Year 6. Teachers were aware that the project examined whole-class interaction and thus it is possible that teachers could have altered their teaching practices; however, this was unavoidable if we were to ensure teachers were aware of the nature of the research with no level of deception.

3.2.3. Phase 3

End-of-year national Statutory Assessment Tests[2] (SATs) in English and mathematics and a SATs-style science test (see Howe et al., 2019) were administered by class teachers and anonymised data were shared with the research team.

3.3. Instrumentation

3.3.1. Cambridge Dialogue Analysis Scheme (CDAS)

This coding scheme built upon the Cam-UNAM SEDA scheme (Scheme for Educational Dialogue Analysis: Hennessy et al., 2016), developed by a large team of researchers in the UK and Mexico. SEDA is a comprehensive analytical tool that enables the analysis of a wide variety of dialogic features in different sociocultural contexts. SEDA is grounded in a range of theoretical perspectives and consists of 33 codes relating to what the authors argue are core features of classroom dialogue. The SEDA codes are clustered into 8 categories: invite elaboration/reasoning, express/invite ideas, make reasoning explicit, build on ideas, position and coordinate, connect, reflect on dialogue, and guide direction of dialogue.

Howe et al. (2019) used SEDA as a basis for developing a new dialogue coding scheme that aligned with the aims of the study. SEDA was adapted, and in particular, condensed, to ensure that: 1) the codes could be reliably applied to a large sample of lessons, and 2) micro-level coding could be employed (in addition to macro-level dimensions) to enable quantification of mutually exclusive codes through frequency counts, as described in some details by Hennessy, Howe, Mercer, and Vrikki (2020). Firstly, the number of original SEDA codes ($n = 33$) was reduced to increase the inter-coder reliability of the scheme, as Hennessy et al. (2016) reported improvements in inter-coder reliability when coding 8 clusters in comparison to 33 codes. The research team selected 18 codes that the literature theorises as optimal dialogue features and reformulated them into a set of 12 broader codes by combining SEDA codes (for instance, 'invite building on' and 'ask for clarification' were subsumed under "invite elaboration"). The final adapted coding scheme, the Cambridge Dialogue Analysis Scheme (CDAS), consists of 12 codes that were applied to each turn within a lesson (summarised in Table 1).

CDAS enabled the examination of various core features of classroom dialogue in each subject area. As outlined in the literature review, educational dialogue is cumulative and reciprocal in nature, therefore it is important not only to examine dialogue features within a specific turn but also to examine how it fits within the exchange as its meaning is shaped by what proceeds and follows that turn (Wegerif, 2006). The coding process was sensitive to the utterances surrounding each turn. Coders watched the video footage alongside reading the transcript to ensure the wider context was taken into consideration when assigning codes to each turn.

Cohen's kappa was used to examine levels of inter-rater reliability between four coders which resulted in substantial reliability on 11 codes (*Cohen's kappa* = 0.62 to 0.80) and marginal reliability for the Reference to Wider Context code (*Cohen's kappa* = 0.58).

3.3.2. Participation dimensions and FFT/MET rating scales

Based on the Frameworks for Teaching (FFT) (Danielson, 2013) and Measures of Effective Teaching (MET) project (Kane & Staiger, 2012) that examined observable determinants of effective teaching, Howe et al. (2019) developed the FFT/MET Scale (Appendix A) to

[1] The National Foundation for Educational Research (NFER, Centre for Assessment, 2016) mathematics and literacy tests (www.nfer.ac.uk/schools/nfer-tests).

[2] SATs are externally-set, national assessments in English and mathematics taken at the age of 11 in every state-funded school as part of the national curriculum (www.gov.uk/education/primary-curriculum-key-stage-2-tests-and-assessments).

Table 1
Cambridge dialogue analysis scheme.*, *

CDAS categories	Codes	Definition
Category 1: Elaboration, reasoning and coordination	Invite Elaboration (ELI)	Invites building on, elaboration, evaluation, clarification of own or another's contribution. E.g. 'Have you noticed anything else that the poet uses?'
	Elaboration (EL)	Builds on, elaborates, evaluates, clarifies own or other's contribution. E.g. [After 'It's sort of describing how you do it'] 'Yes, it's got a good emphasis and a good use of vocabulary'
	Invite Reasoning (REI)	Explicitly invites explanation/justification of a contribution or speculation (new scenarios) /prediction/hypothesis. E.g. 'Why do you think the bottle floats?'
	Reasoning (RE)	Provides an explanation or justification of own or another's contribution, or speculates, predicts, hypothesises with grounds given. E.g. [After 'He came back'] 'because he made a promise'.
	Invite Co-ordination (CI)	Invites synthesis, summary, comparison, evaluation or resolution based on two or more contributions. E.g. 'Would anyone like to summarise the ideas we've been hearing?'
	Simple Co-ordination (SC)	Synthesises or summarises collective ideas (including own and/or others' ideas). Compares, resolves or evaluates different opinions, perspectives and beliefs. 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'.
	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. 'We've been arguing about how much of personality is inherited; twin studies show conclusively it's 50 %'
Category 2: Agreement/querying	Agreement (A)	Explicit acceptance of or agreement with a statement(s). E.g. 'Brilliant', 'Good', 'Yeah', 'Okay', 'I agree with X…'.
	Querying (Q)	Doubting, full/partial disagreement, challenging or rejecting a statement. Includes a simple 'no' response when it shows rejection of an idea; not when in response to a question. E.g. 'Do you really think these angles are the same?' *(A is not intrinsically dialogic but can be when used in conjunction with other codes)*
Category 3: Refer beyond	Reference Back (RB)	Introduces reference to previous knowledge, beliefs, experiences or contributions (includes procedural references) that are common to the current conversation participants. Includes inviting reference back. E.g. 'Can you share with us what we were just having a quick chat about please?'
	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. Includes inviting reference to wider context. E.g. 'It's like in Macbeth where the storm builds into it'.
Category 4: Other invitations *(non-dialogic)*	Other Invitations (OI)	Invitations of all kinds of verbal contributions (e.g. opinions, ideas, beliefs), except for those coded as ELI, REI or CI. This includes invitations on a new topic if this does not fall in another invitation code, and procedural questions.

* Reproduced from Howe et al. (2019), p.476.
* CDAS code names are capitalised throughout this article.

incorporate other codes from the original SEDA that appear infrequently in lessons (such as 'reflect on learning process') and could thus be rated based on whole lessons (Howe, 2016a,Howe, 2016b). Multiple revisions led to the final version, which included 4 dimensions of effective teaching: 1) clarity of goals for student learning, 2) quality of feedback, 3) caring in teacher-student relations, and 4) effectiveness of behaviour management. These dimensions were chosen by Howe et al. (2019) because they represent a measure of Hattie (2009) highly ranked factors influencing student learning outcomes and they were used in the current study as potential confounds.

A key feature of dialogic pedagogy is encouragement of students' active involvement in different aspects of the lesson. So, in addition to 'dialogicality' as measured by CDAS, each whole lesson was rated in terms of student participation. A participation dimensions scale (Appendix B) was developed by Howe et al. (2019) by using 9 codes and 7 broader descriptors from SEDA (Howe, 2016a,Howe, 2016b). This scale provides an indication of student involvement in 1) aims and objectives, 2) monitoring and guidance, 3) reflection on the learning process, 4) focusing on dialogue, and 5) overall student participation.

3.3.3. Student attainment measures

At the end of the school year all participating students completed SATs in (a) mathematics, (b) reading and (c) spelling, grammar and punctuation. Results from these assessments provided a measure of achievement for each class. Whilst mathematics SATs include both arithmetic and reasoning, teachers only had access to students' overall scores in mathematics. Therefore, amalgamated scores of the two mathematics sections were shared with the team and used for analysis in this project. Furthermore, as science SATs have now been eliminated, the team developed a SATs-style science test in conjunction with local teachers (see Howe et al., 2019).

3.3.4. Teacher questionnaire

Howe et al. (2019) developed a 23-item teacher questionnaire (Appendix C) to assess further potential confounding variables. Items related to (1) class demographic information and class activities (e.g. homework); (2) technology use in the classroom; (3) specific teaching practices such as 'I use worked examples'. Teachers were asked to rate the frequency of each item on a five-point scale.

3.4. Confounding variables

Based on Hattie (2009) meta-analysis, Howe, 2016a,Howe, 2016b) identified confounds that are intrinsic to the link between dialogue and learning (such as feedback and quality of teaching) and variables that are extrinsic to this relationship (such as prior achievement and socio-economic status). This research informed the selection of potential confounds for this study. Overall, 30 potential confounds (Appendix D) were derived from measures listed above. For further details on the selection of potential confounds see Howe et al. (2019), pp.481–482. It is important to note that whilst Howe et al. (2019) carried out a thorough examination of all potential factors linked to dialogue and attainment in the literature, any researcher's awareness of potential confounds is necessarily limited. The list of potential confounds in this research is comprehensive but is, of course, bound by what is already known in the literature thus far; therefore, it is possible that other unknown factors associated with the variables of interest would not have been accounted for (Stylianides & Stylianides, 2020).

3.5. Research design

The research design for this study has been typified as a sequential mixed methods design (Teddlie & Tashakkori, 2009, p.151). Combining quantitative and qualitative approaches has been made possible through data conversion (Teddlie & Tashakkori, 2009), a mixed methods approach that quantifies narrative (qualitative) data. To achieve this, dichotomous variables were created and applied to transcripts of lessons (Creswell & Plano Clark, 2011); for example, every time reasoning emerged within the narrative, a '1' was coded for the relevant turn. This turn-by-turn coding allowed contextual factors to be examined within the process of converting the narrative data (in the form of videos/transcripts) into quantitative outcomes. This approach also has the benefits associated with rigorous quantitative results, including statistical analysis to establish reliability among coders, and the ability to test specific hypotheses, make predictions, account for contributing variables and generalisability. This method also enabled the movement back and forth between numerical and narrative data to provide illustrative excerpts of numerical information.

3.6. Research ethics

This research complies with the Economic and Social Research Council (2012), BERA Ethical Guidelines for Educational Research (British Educational Research Association, 2011) and the BPS Code of Ethics and Conduct (British Psychological Society, 2009). The study received ethical approval from the University of Cambridge Education Ethics Committee on 6th May 2015. Details of the ethical procedures taken can be found in Amodia Bidakowska (2021).

4. Results

4.1. Subject dialogues

RQ1: To what extent does the frequency of dialogue features differ between English, mathematics and science lessons?
Three sub-questions were formulated:

A. Are there subject differences in the frequency of dialogue features?
B. Are there subject differences in the frequencies of teachers' and students' dialogue features?
C. Are there subject differences in participation dimensions?

The full sample consisted of 144 lessons; however, the majority of teachers selected mathematics and English lessons for observations, resulting in fewer science lessons overall. To enable comparability between subjects the sample needed to consist of an equal number of lessons in each subject. As there were a total 24 science lessons, a maximum of 24 lessons in each of English and mathematics were also selected. It is also important to consider the possibility that teacher variation may impact upon findings (for example, teachers in the mathematics lessons group could be more dialogic by chance). Therefore, to ensure any subject differences were not due to teacher variation, each teacher provided two lessons from two different subjects. The final sample used to answer Research Question 1 consisted of 12 teachers providing one English and one mathematics lesson, 12 teachers providing one English and one science lesson, and 12 teachers providing one mathematics and one science lesson (a total of 72 lessons from 36 teachers). Whole lessons were filmed and analysed (including whole-class and group work); the camera followed the teacher at all times. Each lesson varied in duration (range = 30–102 min, mean = 65 min). A total of 4669 min of lesson time were analysed for this research.

To examine whether the frequencies of CDAS codes differed significantly between English, mathematics and science lessons, an analysis of variance was carried out (controlling for lesson duration). There was a significant effect of subject on the frequencies of: Invite Elaboration $F(2, 69) = 4.78$, $p < .05$, $\omega = 0.30$; Elaboration $F(2, 69) = 5.02$, $p < .01$, $\omega = 0.30$; Invite Reasoning $F(2, 69) = 21.01$, $p < .001$, $\omega = 0.60$; and Reasoning $F(2, 69) = 22.19$, $p < .001$, $\omega = 0.60$. Overall, mathematics lessons contained significantly more instances of Reasoning and Invitations for Reasoning than English and science lessons. Mathematics lessons also contained the highest proportion of Invitations for Elaboration, significantly more than science lessons. English and mathematics lessons contained equitable proportions of Elaboration, and both were significantly more frequent than science lessons.

Furthermore, a non-parametric Kruskal-Wallis test revealed that there was a significant effect of subject on the frequency of Agreement, $H(2) = 15.1$, $p < .001$. Agreement was significantly more frequent in mathematics lessons than in science lessons. There

was also a higher proportion of Querying in mathematics compared to English and science, $H(2) = 6.6, p < .05$. The same analyses were conducted using CDAS codes as proportions of turns to rule out the potential impact of turn length (turns were significantly longer in English lessons than in mathematics and science); patterns of subject differences mirrored the above results generated from raw frequencies.

There were no significant differences in the frequencies of Invitations for Coordination, Simple Coordination and Reasoned Coordination codes. These codes appeared very infrequently across the whole sample (e.g. range of Invite Coordination was 0–1 times per lesson) and were thus excluded from all further analysis. There were also no significant differences in the frequencies of Reference Back and Reference to Wider Context codes between subjects.

As the frequencies of Reasoning and Elaboration differed significantly according to subject, teacher and student contributions were examined separately. In accordance with the patterns presented above there was a significant effect of subject on the frequencies of teachers' Invitations for Reasoning ($F(2, 69) = 12.99$, $p < .01$, $\omega = 0.5$) and students' Reasoning ($F(2, 69) = 9.89$, $p < .001$, $\omega = 0.4$), in favour of mathematics over English and science. Therefore, the higher proportion of teachers' Invitations for Reasoning in mathematics seems to have resulted in a higher proportion of students' overall Reasoning, as we might expect. However, teachers' Reasoning did not significantly differ between subjects. Fig. 1 presents a visual representation of subject differences in students' and teachers' reasoning.

There was a significant effect of subject on teachers' Invitations for Elaboration ($F (2, 69) = 5.29$, $p < .01$, $\omega = 0.3$) and students' Invitations for Elaboration ($F (2, 69) = 3.12$, $p = .05$, $\omega = 0.2$). Although mathematics lessons contained the most Invitations for Elaboration overall, when examining teacher and student invitations separately, English lessons contained the most student Invitations for Elaboration, significantly more than science. Both English and mathematics contained significantly more teacher Elaboration than science lessons. Student Elaboration was highest in English and significantly more frequent than in science lessons. Fig. 2 displays the subject differences in teachers' and students' Elaboration.

Additionally, there was a significant effect of subject on teacher Agreement ($H(2) = 7.6, p < .05$) and Querying ($H(2) = 7.8$, $p < .05$), in favour of mathematics, but there was no difference in student Agreement or student Querying between subjects. All significant findings relating to teacher and student contributions are presented in Fig. 3.

When looking at correlations between teacher invitations and student responses, similar patterns were identified across all three subjects (Table 2). Overall, teachers' Invitations for Reasoning were highly correlated with student Reasoning in English, mathematics and science. Similarly, teachers' Invitations for Elaboration were highly correlated with student Elaboration in all three subjects. Again, these correlations show that students are responding in expected ways to teacher invitations.

Finally, subject differences were examined in relation to lesson ratings of participation dimensions. Analysis for this research question was limited to Fisher's exact test as the expected frequencies for this sample were too low (> 5) in some categories to carry out a chi-squared test. The findings revealed no significant subject differences on any of the rating scales.

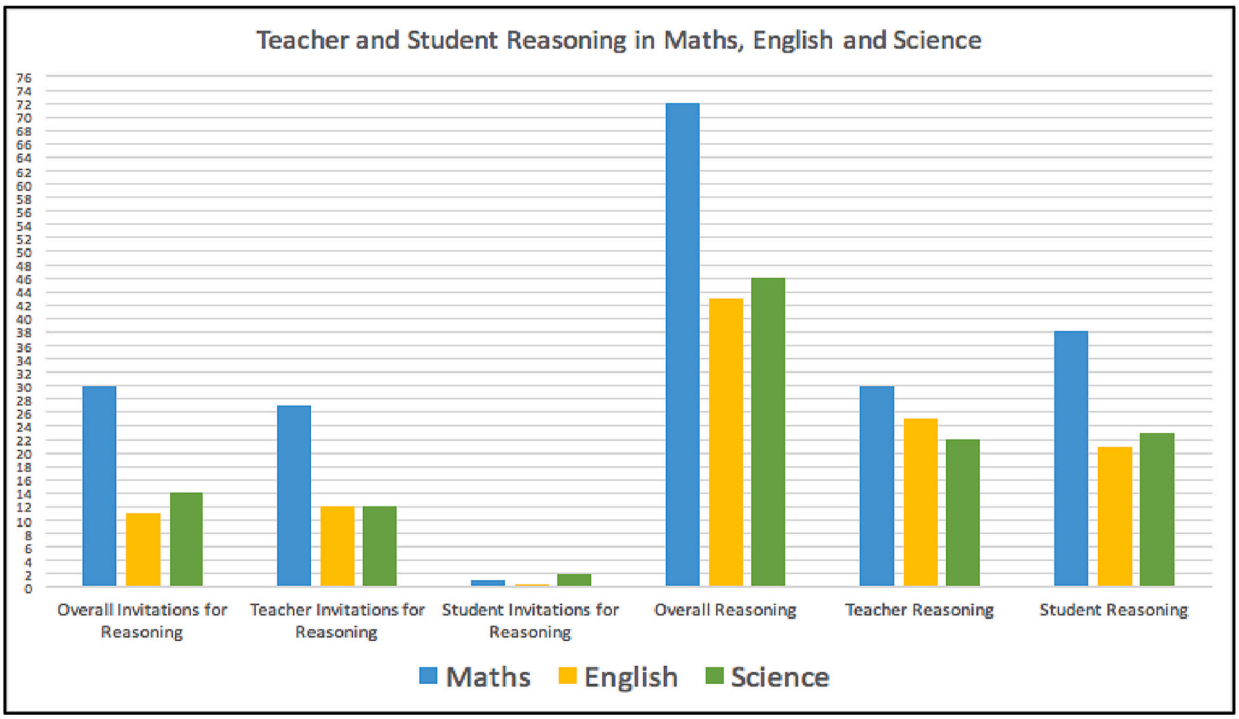

Fig. 1. Mean frequencies/lesson of teacher and student Reasoning and Invitations for Reasoning.

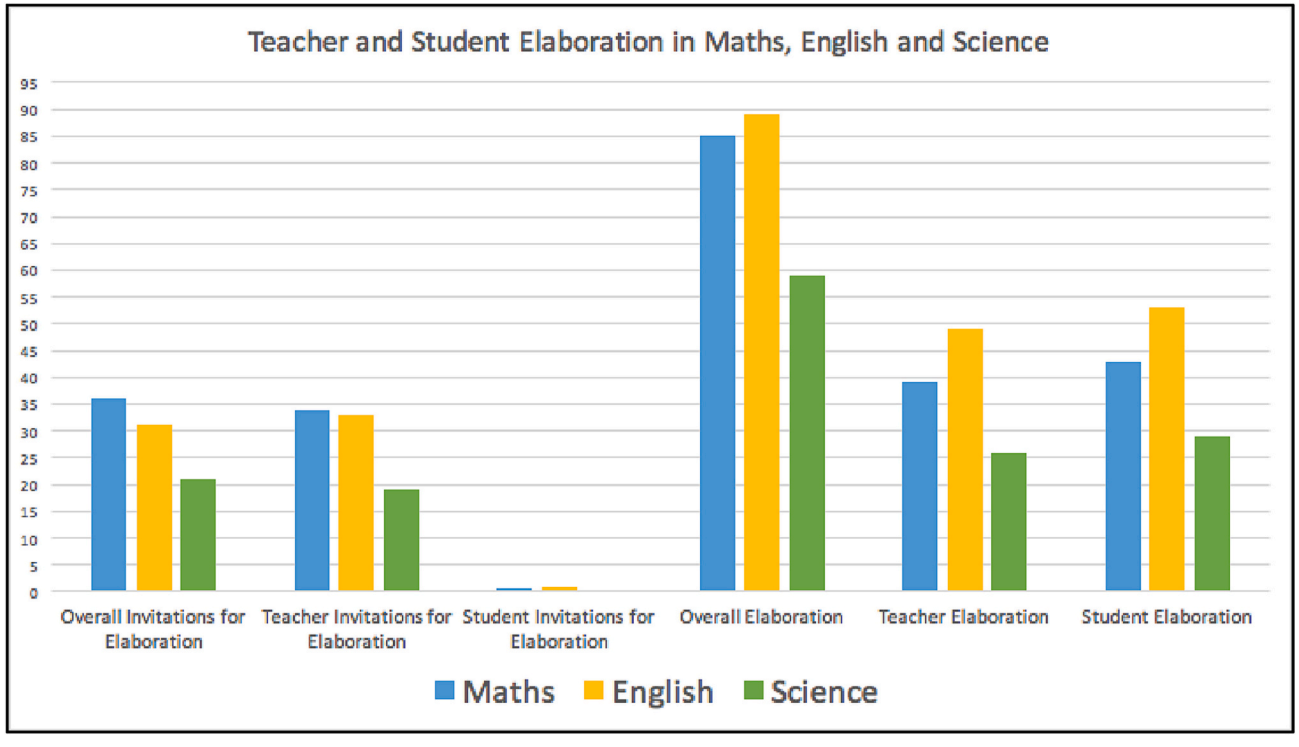

Fig. 2. Mean lesson frequencies of teacher and student Elaboration and Invitations for Elaboration.

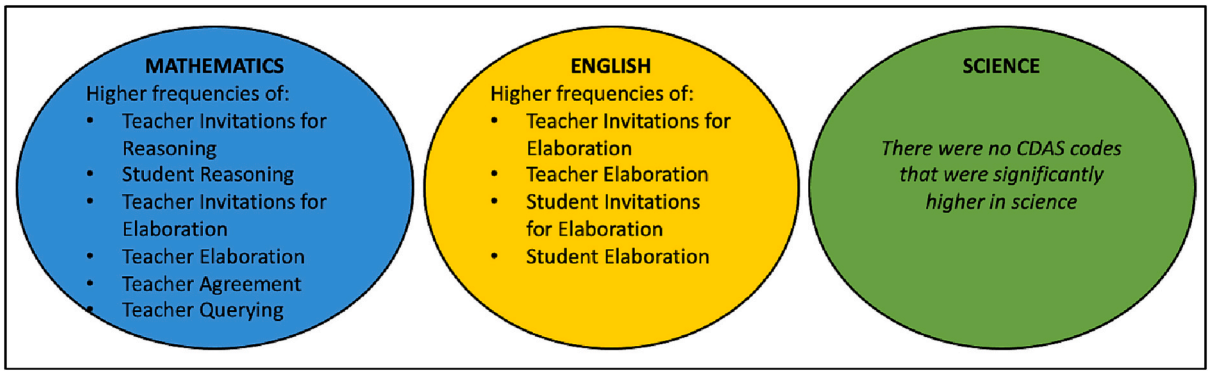

Fig. 3. Overview of significant subject differences in CDAS frequencies for teacher and student contributions.

Table 2
Correlations between teacher questioning and student responses.

	Mathematics	English	Science
Teacher invitations for reasoning & student reasoning	0.82**	0.82**	0.67**
Teacher invitations for elaboration & student elaboration	0.92**	0.75**	0.82**

* $p < .05$, ** $p < .01$.

4.2. Subject dialogues and student attainment

Multi-level modelling (MLM) was used to examine RQ2: Are there subject-specific features of dialogue associated with high attainment?

To address RQ1, the maximum number of science lessons ($n = 24$) was extracted and subsequently 24 lessons were randomly selected from English and mathematics. However, the same sample could not be used to address RQ2, as availability of attainment data also needed to be taken into consideration. For RQ2 a different sample was selected to: 1) ensure a balanced proportion of each subject; 2) obtain one lesson per teacher to meet the MLM assumption of outcome variable independence (i.e. independence across classes); and 3) optimise inclusion of the classes where attainment data were available. As discussed, science lessons were filmed in 24 classes; however, only 21 of these classes had attainment data available, therefore the final sample consisted of 63 separate teachers, each contributing one lesson in English ($n = 21$), mathematics (n = 21) or science (n = 21).

MLM is a regression-based statistical method used to analyse hierarchical data with nested structures. The dataset included two levels of data. Firstly, it contained CDAS variables and lesson ratings which were measured at the *class* level. These variables were used to predict attainment in each subject, which was collected at the *individual* level. As students are nested within classes, MLMs were used to analyse the class level and individual level data simultaneously to explore associations between CDAS variables/lesson ratings within each subject and attainment. Separate MLMs were carried out for each attainment outcome: Mathematics SATs, Reading SATs, SPAG SATs and Science Test. The links between core dialogue features in each subject and student attainment were explored whilst accounting for a broad range of potential confounds (such as prior attainment). Potential confounds were included in all subsequent analyses to ensure that any associations that arose from the analysis of attainment data remained significant over and above any effects

Table 3
Significant associations between predictors and potential confounding variables in mathematics.

Predictor & potential confound	Pearson Correlations
Elaborated dialogue & Inductive	−0.496*
Reasoned dialogue & Integration	0.442*
Querying & Spaced	−0.584**
Reference Back & PASS	−0.618**
Reference to Wider Context & Worked Examples	−0.491*
Reference to Wider Context & Lecture	0.485*
Aims & Goals	0.557**
Aims & Lecture	−0.471*
Aims & Mapping	−0.452*
Aims & Trips	0.486*
Monitoring & Feedback	0.633**
Student Participation & Lecture	−0.459*

* $p < .05$, ** $p < .01$, *** $p < .001$.

of other variables associated with both dialogue and attainment in each subject. Overall, 30 potential confounds (Appendix D) were derived from the teacher questionnaire (Appendix C), student questionnaire (Appendix E), lesson ratings using a framework for teaching rating scale (Appendix A), the Pupil Attitude to Self and School (PASS) survey (GL Assessment, 2013).

In order to explore the link between subject-specific features of dialogue and attainment through multilevel modelling, analyses were first completed with potential confounding variables. Firstly, a correlation matrix was examined to identify any significant associations between 25 potential confounds and dialogue features/lesson ratings within each subject. This resulted in a large number of potential confounds per subject (see associations in Tables 3, 4 and 5).

As displayed in the tables above, there were subject differences in associations between predictors and potential confounds. For example, the amount of peer group work (no teacher present) was positively associated with the frequency of Elaborated dialogue in science, but not significantly associated with the frequency of Elaborated dialogue in mathematics or English lessons. Overall, 10 potential confounds were significantly associated with predictors in mathematics, 7 were associated with predictors in English and 10 were associated with predictors in science.

The second step examined whether these same potential confounds were also related to attainment in the relevant subject area. For example, if group work in science lessons was associated with the frequency of Elaboration in science, was group work also associated with science attainment? Thus, the above potential confounds, prior attainment measures and dialogue predictors were then entered into a MLM for each of the four attainment outcomes. Any significant associations that emerged from these MLMs highlighted which variables should be entered as predictors and covariates in the final analysis (Table 6).

The above analysis helpfully reduced the number of confounds in each subject case to be entered as covariates in the final MLMs. In mathematics, 'Trips' was the only confound related to mathematics SATs attainment. In English, 'Feedback' was related to Reading SATs attainment and 'Goals' and 'Feedback' were related to SPAG SATs attainment. Finally, 'Free School Meal Eligibility' and 'Video Animation' were significantly associated with attainment on the science test.

The above covariates were entered into the four final MLMs to explore the extent to which the frequencies of dialogue features and lesson ratings (at the class level) explain variance in subject attainment (at the student level). The first MLM included mathematics SATs as the outcome variable, revealing that Reference to the Wider Context was negatively associated with mathematics attainment, $t(543) = -2.4$, $p = .017$; this suggests that when teachers and students make frequent references to the wider context in mathematics lessons, this negatively impacts on students' attainment on the mathematics SATs (Table 7). The final three MLMs (SPAG SATs, Reading SATs and Science Test) revealed no main effects for any dialogue variables on attainment in English and Science.

Based on the above results, further MLMs were run to explore interactions between predictors; for example, how does the combination of high Elaborated dialogue and high Reference to Wider Context in English lessons impact upon Reading attainment? Due to the large number of predictors, there are multiple combinations of interactions that could be tested; thus it is crucial to form initial predictions regarding the possible combinations of variables that may impact upon attainment in each subject.

Firstly, results from Howe et al.'s (2019) study indicated that features of dialogue (Elaborated Dialogue and Querying) *across* subjects were positively related to attainment as long as students' overall participation (as indicated by lesson ratings) was high. Thus, the first interactions explored in this study included combinations of student participation ratings and the core dialogic codes Elaborated/Reasoned. The intention was to ascertain whether combinations of high Elaborated/Reasoned dialogue, coupled with high student participation in *each* subject, were also associated with attainment.

Next, interactions including Reference to the Wider Context were carried out to explore inconsistencies between the current study findings *within* each subject and Howe et al.'s (2019) findings *across* subjects. As discussed above, this study found that Reference to the Wider Context was the only dialogue variable that had a direct negative effect on attainment in mathematics. However, Howe et al. (2019) found no significant findings relating to Reference to the Wider Context across subjects. If Reference to the Wider Context *was* positively associated to English or science, this may explain why it was non-significant across subjects in Howe et al.'s (2019) study; that is, divergent subject differences could be cancelling out the effect across subjects. As shown above, Reference to the Wider Context in the current study did not have a direct effect on attainment in English or Science, so interactions including Reference to the Wider Context and the core dialogue variables (Elaborated/Reasoned) were examined.

To run interactions within MLMs, the use of centred variables - i.e. data points subtracted by the mean of all values (Field, 2009) - is recommended to reduce multicollinearity between predictors used in interactions. Centred variables were created by transforming predictor variables into deviations around the mean (Field, 2009). Centred predictors were then entered in combination into the model in addition to the covariates associated with each predictor. Table 7 summarises significant and marginally significant findings in each subject area.

Table 4
Significant associations between predictors and potential confounding variables in English.

Predictor & potential confound	Pearson correlations
Reference Back & Video Animation	0.435*
Monitoring & Inductive	−0.655**
Monitoring & Goals	0.487*
Monitoring & Feedback	0.766**
Reflection & Spaced	0.580**
Dialogue & Inductive	0.441*
Student Participation & English Prior Attainment SD	−0.586**

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 5
Significant associations between predictors and potential confounding variables in science.

Predictor & potential confound	Pearson correlations
Elaborated dialogue & Video Animation	−0.474*
Elaborated dialogue & Group	0.519*
Querying & Relations	−0.505*
Querying & FSM	−0.475*
Reference Back & Mathematics Prior Attainment SD	0.530*
Aims & FSM	0.471*
Monitoring & Feedback	0.655**
Monitoring & Mapping	0.451*
Reflection & Goals	0.532*
Dialogue & Spaced	−0.449*
Student Participation & Mathematics Prior Attainment SD	0.623**

* $p < .05$, ** $p < .01$.

Table 6
Significant associations between predictors and attainment outcomes.

Predictor & attainment	Significance
Frequency of School Trips & Mathematics SATs	$t(543) = 3.74$, $p < .001$
Feedback & Reading SATs	$t(543) = 2.1$, $p = .037$
Feedback & SPAG SATs	$t(543) = 2.08$, $p = .037$
Goals & SPAG SATs	$t(543) = 1.81$, $p = -0.071$
Free School Meal Eligibility & Science Test	$t(543) = 3.83$, $p < .001$
Video Animation & Science Test	$t(543) = 3.13$, $p = .002$

Table 7
MLM results examining the link between dialogue and attainment.

Subject	Covariates	Direct/interaction effect	Predictor	Outcome variable	Significance
Mathematics	Trips	Direct effect	Reference to wider context	Mathematics SATs	$t(543) = -2.4$, $p = .017$
Mathematics	Trips	Interaction effect	Reasoning * Reference to wider context	Mathematics SATs	$t(542) = 2.93$, $p = .003$
English	Goals feedback	Interaction effect	Elaborated * reference to wider context	SPAG	$t(543) = 2.798$, $p = .005$
English	feedback	Interaction effect	Elaborated * reference to wider context	Reading SATs	$t(543) = 2.18$, $p = .03$

As displayed in Table 7, Elaborated Dialogue was significantly positively associated with English attainment as long as Reference to the Wider Context was also high; this was significant for both SPAG and Reading SATs. Secondly, whilst Reference to the Wider Context had a direct significant negative effect on mathematics attainment, there was a significant positive association between the combination of high Reasoned Dialogue and high Reference to the Wider Context and mathematics attainment. A summary of these findings is presented in Fig. 4. There were no interactions significantly associated with science attainment. Additionally, the lesson ratings, such as 'Student Participation' did not yield any significant interaction effects in any subject.

5. Discussion

5.1. Disciplinary differences in dialogue

This study firstly explored how teacher-student dialogue unfolds in three core disciplines at primary school level. The rationale behind this research question stems from the large body of literature that promotes certain forms of dialogue for student learning. Whilst there is little empirical evidence that directly shows disciplinary differences in dialogue, a review of the literature identified a proclivity towards certain features of dialogue in different subject areas and thus gave rise to question whether TPD on classroom dialogue should be disciplinary. Comparisons between subjects revealed a number of interesting and significant subject differences.

Firstly, the frequencies of teachers' Invitations for Reasoning were significantly higher in mathematics than English and science, as were instances of students' Reasoning. Students' Invitations for Reasoning were fairly infrequent across all subjects and did not significantly differ; this aligns with findings indicating that 93 % of Invitations for Reasoning are uttered by teachers (Vrikki et al., 2019). The overall finding, suggesting a proclivity towards reasoned dialogue in mathematics, falls in line with the literature reviewed above highlighting the centrality of reasoning in mathematics at primary and secondary level (e.g., Mercer & Sams, 2006; Schwarz et al., 2010). Additionally, as discussed above, the benefits of reasoned dialogue have been shown to be task-dependent; for example, Rojas-Drummond et al. (2006) found that the use of reasoning leads to performance gains in a problem-solving task but not, for example, in an open-ended literacy task which instead was seen to benefit from the use of elaboration and reformulation. This research may help to explain why English showed significantly higher levels of Elaboration than science in the current study, a finding that concurs with Nassaji and Wells' (2000) reports of higher 'Negotiatory Questions' (i.e., 'opinion and explanation') in literacy than in

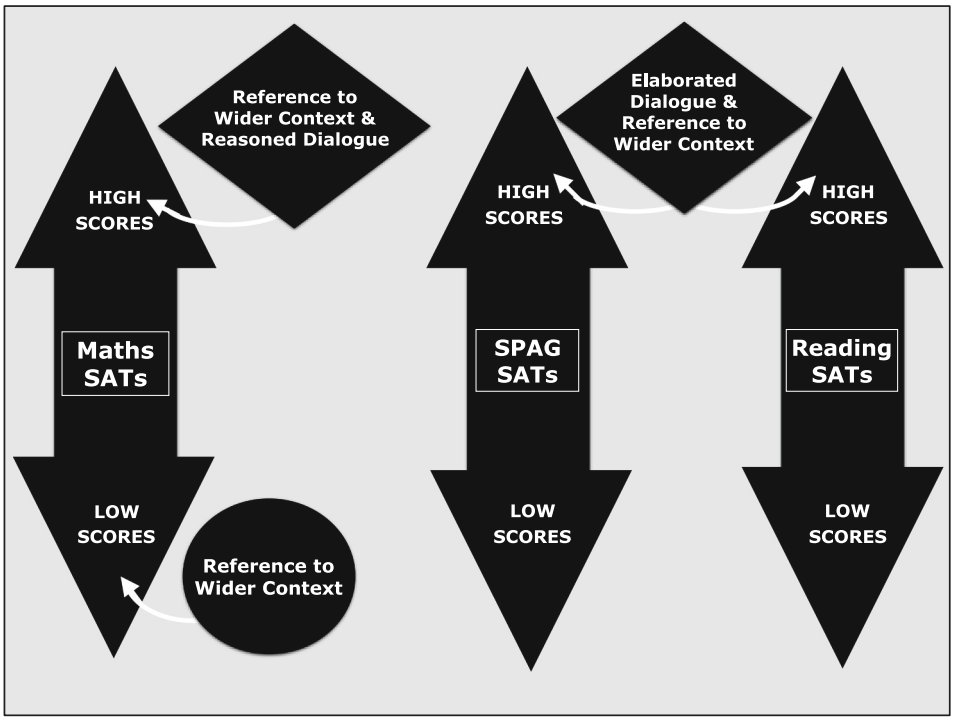

Fig. 4. Significant associations between dialogue features and attainment in mathematics, SPAG & Reading SATs.

science lessons.

Other significant differences that emerged from the dataset included those relating to Agreement and Querying. Agreement was significantly more frequent in mathematics than science and Querying was significantly more frequent in mathematics than both English and science. The use of Agreement and Querying may be more prominent during tasks that require the eventual identification of one correct answer. Such tasks may lend themselves to more agreement and querying compared to a more open-ended task that draws on multiple perspectives. Wells and Arauz (2006) argue that position-driven discussions that begin with a statement and follow up with justifications of agreement/disagreement could be more common in mathematics and science activities where one correct answer is required, but less common in arts subjects. Often in mathematics, it seems that agreement is required before students can move on to the next step in the process.

The findings thus far fall in line with predictions informed by the literature review. One area, however, that did not align with expectations was in relation to the results from science classrooms. The literature indicates that reasoning is also a crucial feature of discussions in science (Asterhan & Schwarz, 2007; Driver et al., 2000; Howe et al., 2007; Howe & Tolmie, 2003; Larrain et al., 2014; Zohar & Nemet, 2002). Researchers argue that reasoning and argumentation (including challenging, rejecting and counter-argument) are entwined in the core elements of science, including hypothesising based on available information, the selection of methods to test hypotheses and finally the analysis and interpretation of results (Lemke, 1990; Osborne & Chin, 2010). Indeed, studies have shown the link between scientific reasoning and attainment (e.g. Larrain et al., 2014). The results generated from Research Question 1 (indicating lower reasoning and querying in science lessons) therefore raise the question of whether the elimination of formalised end of year science assessments in UK primary schools may account for these findings. Although SATs are assumed to lead to transmissive 'teaching to the test' with limited scope for dialogue, since the elimination of science SATs in 2010, the number of hours dedicated to science lessons have dramatically reduced at primary level. Leonard, Lamb, Howe, and Choudhoury (2017) report an average of only 1 h and 24 min of science per week; less time dedicated to science may result in teachers rushing through the curriculum with limited time to implement dialogic techniques (supported by teacher interviews carried out by Amodia Bidakowska, 2021. This argument assumes that a reduction in curriculum time (whilst retaining a high expectation in terms of content coverage within the curriculum) could have led to conditions that discourage a dialogic approach in science classrooms. Furthermore, there is some literature that indicates that whilst the research field emphasises the importance of argumentation and reasoning in scientific learning, argumentation during student-student discussions in science does not occur in practice as often as expected (Christodoulou & Osborne, 2014; McNeill, 2011) and primary school teachers struggle to encourage scientific argumentation among their students (Aydeniz & Ozdilek, 2016).

The analysis undertaken in this paper did not identify any significant subject differences among subjects in terms of the coordination codes ('Invite Coordination', 'Simple Coordination' & 'Reasoned Coordination'). In fact, the coordination codes were very rare in every subject and thus no meaningful comparisons could be made. This is an important finding as there is a large literature dedicated to the importance of coordinating ideas and perspectives for student learning (Alexander, 2008; Reznitskaya & Gregory, 2013). The

findings from this study bring to question whether coordination is being taught in English primary schools.

Overall, subject trends in mathematics and English align with the disciplinary trends in the literature. However, findings only indicate a propensity towards certain forms of dialogue in English and mathematics. To test the usefulness of the coding scheme outputs further, CDAS data should be statistically examined in relation to outcomes of interest. As such, the next crucial step in this study was to examine whether the dialogue features common in each discipline (e.g. Elaboration in English) were associated with attainment in that particular subject i.e. do these trends exist because they have a particular advantageous outcome for student learning in that subject area? This question contributes to the debate surrounding why subject differences in dialogue may have emerged from this dataset. Findings relating to science lessons challenged expectations for primary science; however, these results only indicated that optimal forms of dialogue were not as frequent as expected; as mentioned above, they do not inform us on the quality of those forms of dialogue. Therefore, the next RQ explored whether optimal forms of dialogue associated with science in the literature (e.g. reasoning), though not as frequent as expected, still had a beneficial impact on science attainment.

5.2. Direct effects

The only direct effect of dialogue on attainment identified through MLM was Reference to the Wider Context in mathematics lessons. However, in contrast to the literature which emphasises the importance of linking the curriculum to the wider context for problem solving and application (e.g. Carpenter & Lehrer, 1999), results from this study showed that when students and teachers make frequent references to the wider context in mathematics lessons, this is associated with low scores on mathematics SATs. This apparent contradiction is explored below. When Reference to the Wider Context was examined in English or science alone, no direct effects were identified, which thus brings to question whether the importance of referring to the wider context could be disciplinary.

Reference to the Wider Context involves references to life outside the school context (such as story contexts or references to films/places etc.) or bringing in ideas, knowledge and examples from the everyday life, such as the use of experiences to contextualise the content taught in school. This thinking falls in line with constructionism (Ackermann, 2001), which postulates that optimal learning arises from the construction of links between taught content and experiences of the world (Berns & Erickson, 2001). Researchers suggest that whilst this is emphasised as ideal, in practice the successful application of mathematical concepts to real life occurs infrequently (Gainsburg, 2008). Some researchers emphasise that mathematical learning *should* be taught in the abstract, as this method does not require the learner to process additional extraneous details that they deem irrelevant to the comprehension of mathematical concepts (Belenky & Schalk, 2014; Kaminski, Sloutsky, & Heckler, 2008). There is an ongoing debate among educationalists surrounding the teaching of mathematics in abstract form versus contextual instruction. A comprehensive discussion of this debate can be found in Fyfe, McNeil, Son, and Goldstone (2014) systematic review

Consistent with Howe et al.'s (2019) study across subjects, the present study did not identify any further direct effects of the remaining dialogue variables on attainment in any subject. Therefore, whilst Elaborated Dialogue, Reasoned Dialogue, Querying, Reference Back and Reference to the Wider Context are theorised to be productive for learning, these features on their own did not solely contribute to attainment in any of the core subjects in this sample. However, further analysis has shown that a combination of these dialogue variables in each subject link to student attainment gains. Importantly, these significant interactions occurred whilst controlling for a number of key variables which overcomes some of the limitations associated with previous research: Howe and Abedin (2013) systematic review revealed that the majority of studies in their sample had not isolated their exploration of dialogue and learning from other potentially confounding factors.

5.3. Interaction effects

The examination of interactions through MLM revealed a number of interesting disciplinary differences. Primarily, the combination of high Reference to the Wider Context and high Elaborated Dialogue was significantly positively associated with students' attainment on SATs Reading. Additionally, the same combination of dialogue features was significantly positively associated with SPAG SATs. These findings therefore emphasise the importance of linking English lesson content to students' experiences and other subject matter, and, in line with the literature (e.g., Maine, 2013; Rojas-Drummond et al., 2006), building on and elaborating others' (and one's own) ideas. Drawing on the cognitive elaboration perspective from the work of O'Donnell (2006), Webb (2009) states that students elaborate on their ideas as a means of communicating their knowledge or ideas to their discussion partners. In doing so, "cognitive reconstructing" takes place where the speaker's own knowledge is reconstructed by clarifying and enhancing their ideas to provide a more articulate version of their knowledge (Webb, 2009, p.2; Veenman et al., 2005; Bargh & Schul, 1980). The combination of these dialogue features in the current study showed to provide a positive learning environment for both elements of English assessment: literacy and language. This interaction was not identified in Howe et al.'s (2019) study when examining across subjects; thus, it seems that it is specifically the type of Elaborated Dialogue and Reference to the Wider Context that occurs in English lessons that contributed to attainment.

The final significant interaction that arose from the analysis was the combination of Reference to the Wider Context and Reasoning in mathematics only. At initial glance, these results appear contradictory to those above. One plausible explanation is that regular references beyond the core mathematic concepts taught may be meaningless on their own if the connection has not been explicitly made. However, when references to the wider context are coupled with reasoning to explain and justify links to the outside world, their exploratory value becomes evident and thus may explain why a significant positive interaction between Reasoning and Reference to the Wider Context arose from the mathematics data. In a recent study by Lehesvuori et al. (2021), teachers' suggestions for increasing dialogicality and student engagement in their mathematics classrooms included making connections between learning content and

everyday contexts.

These results are of particular significance as researchers and teachers consider the effect on attainment of combinations of dialogue features in teaching and learning.

5.4. Non-significant results

It was predicted that interactions including ratings of student participation (in conjunction with elaboration/querying) would link to student attainment (as indicated by Howe et al.'s, 2019 study), since in order for teacher-student dialogue to be displayed, students must be fully committed to discussions and frequently contribute; this factor is unlikely to be disciplinary but relevant across all subjects. However, the examination of interactions in the present study did not identify any significant effects of Student Participation ratings when separated by subject. This could be explained by the dilution of associations when dividing by subjects (i.e., instead of using the whole sample of 72 lessons as in Howe et al.'s, 2019 study, this study split by subject resulting in a smaller sample size of 21 lessons in each subject).

Another area that did not draw any significant findings was science. A lower prevalence of Reasoning in science (as indicated in RQ1) does not necessarily mean a lack of quality in individual lessons. It was postulated that perhaps the Reasoning that did occur in science lessons, though not as frequent as expected, could be of high value and thus may contribute to science attainment. However, aligned with Howe et al.'s (2019) findings across subjects, the current study did not identify any direct or interaction effects of any dialogue variables on science attainment. Howe et al. (2019) stated that their lack of significant findings relating dialogue to science attainment could not be due to measure design, as it is robust in terms of its psychometric properties. Howe and colleagues do, however, speculate that the lack of findings relating to science attainment may be a result of test difficulty. Although the test content matched that of the curriculum, teachers reported low topic coverage on evolution and inheritance, which may have explained low overall student performance on this measure (Howe et al., 2019).

6. Conclusions

Overall, this study has revealed the importance of the disciplinary context on classroom dialogue. Primarily, it has contributed to the field by highlighting the following key findings and associated recommendations:

1. Dialogue features theorised to be productive for learning are, to some extent, linked to attainment, but disciplinary differences exist.

This study is the first to explore subject dialogues at a larger scale and provides evidence that there are indeed major differences in the forms of dialogue observed among this sample's primary school teachers and students in each subject. But are these differences beneficial for learning? Should subject dialogues be encouraged? In fact, RQ2 indicated that disciplinary trends in dialogue appear to be linked to learning. This research indicates that a focus on specific features of dialogue in subjects may be important. Therefore, there is an argument for increasing generalist primary school teachers' awareness of these nuances. Guidance for teachers resulting from this study has now been added to the Toolkit for Systematic Dialogue Analysis (T-SEDA), an extensive professional learning resource supporting practitioner inquiry into dialogue in their settings (T-SEDA Collective, 2023). A careful balance between empowering teachers to implement and model a broad range of dialogic techniques in addition to using that knowledge to tailor an approach for the disciplinary context could be a valuable outcome of this area of research. The current state of research on subject dialogues and attainment is not yet sufficient to draw any firm conclusions on how to shape classroom dialogue for the disciplinary context or indeed what optimal subject dialogue is. What has been provided, however, is a strong case to begin to question existing practice, asking how teacher professional development on dialogue could be sensitive to the disciplinary context of the lesson. A replication of this study at a larger scale and with different age groups would be valuable for the field to explore whether subject differences in dialogue are more pronounced at secondary level.

2. Classroom dialogue is most prevalent in primary mathematics

A core message generated from this study is the proclivity towards more dialogic forms of interaction in mathematics lessons. Not only were Reasoning and Invitations for Reasoning more frequent in mathematics but also Elaboration, Invitations for Elaboration, Agreement and Querying. So how can we mirror the high levels of subject-specific dialogue seen in mathematics lessons in other subjects such as science? Clearly the curriculum plays a key role by outlining exactly what is required of teachers; reasoned dialogue is an expectation outlined in the mathematics curriculum. Efforts need to be put in place to ensure the same quality of subject dialogues are outlined in curricula of other subjects such as science, in addition to associated teacher professional development on dialogue in science which is currently lacking.

3. The coordination of ideas does not occur in upper primary lessons in England.

Coordination (including synthesis, summary, comparison, evaluation or resolution of two or more contributions) did not take place in this sample of Year 6 lessons in *any* subject. This finding is pivotally important for the field as there is a strong emphasis on the importance of this dialogue feature for learning in the literature (e.g. Alexander, 2008; Reznitskaya & Gregory, 2013). It will be crucial

for future research to examine the frequencies of coordination in older age groups to explore if, among older students, the coordination of ideas is more common or still scarce. Secondly, an intervention study would be of use to examine whether student learning is benefited when coordination is enforced (through modelling and invitations for coordination).

Funding details

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Rather, an existing research dataset was drawn upon (Howe et al., 2019); this research was supported by the Economic and Social Research Council of Great Britain (ESRC: Grant ES/M007103/1).

Declaration of competing interest

None.

Data availability

The authors do not have permission to share data.

Appendix A. FFT/MET scale

(Howe et al., 2019)

	LEVEL 0	LEVEL 1	LEVEL 2
Clarity of GOALS for student learning	Goal(s) for student learning are either not clear or are stated as student activities. Goal(s) do not permit viable methods of assessment.	Goal(s) are only moderately clear or include a combination of goals and activities. Some goals do not permit viable methods of assessment.	Goal(s) are mostly clear but may include a few activities. Goal(s) are framed in terms of student learning and mostly permit viable methods of assessment.
Quality of FEEDBACK	Feedback is either not provided or is of uniformly poor quality [‘High quality’ feedback is task relevant, formative guidance for student learning, especially focused on learning process and self-regulation].	Feedback is inconsistent in quality: Some elements of high quality are present; others are not. Less than half of the feedback given includes formative guidance.	Feedback is consistently high quality: Over half of the feedback given is formative guidance.
Caring in teacher-student RELATIONS	Teacher interaction with at least some students is negative, demeaning, sarcastic or inappropriate to the age or culture of the students. Students exhibit disrespect for teacher.	Teacher-student interactions are generally appropriate but may reflect occasional inconsistencies, favouritism, or disregard for students’ culture. Students may exhibit some disrespect for teacher.	Teacher-student interactions are friendly and demonstrate general warmth, caring and respect. Such interactions are appropriate to developmental and cultural norms. Students exhibit respect for teacher.
Effectiveness of BEHAVIOUR MANAGEMENT	Student behaviour is not monitored, and teacher seems unaware of what students are doing.	Teacher is generally aware of student behaviour, but may miss the activities of some students.	Teacher is alert to student behaviour at all times. Monitoring by teacher is largely subtle and preventative.

Note. Effectiveness of Behaviour Management should be rated *by observers after the lesson* as the video record cannot always capture this dimension. Observers should also take note of subtle measures like seating plans that split up troublesome students.

Appendix B. Participation dimensions

(Howe et al., 2019)

	0 Not evident	1 Teacher-led	2 Teacher-led with student involvement
Aims and objectives	Aims (goals, objectives, plans etc.) are not made explicit	Teacher ‘announces’ lesson aims, with at most non – interactive, e.g. written, responses from students	Aims (goals, objectives, plans etc.) are to some extent negotiated with students or by students themselves *(does not include negotiation of dialogue rules – that is measured in ‘Focusing on dialogue’ scale)*
Monitoring and guidance	Monitoring and guidance are not apparent	Teacher watches while students carry tasks out and offers suggestions and/or evaluations	Teacher offers help without taking over. Monitoring and guidance is to some extent tailored through dialogue towards students’ needs (i.e. scaffolding or focusing, e.g.

(*continued on next page*)

(*continued*)

	0 Not evident	1 Teacher-led	2 Teacher-led with student involvement
			teacher strategically balances dialogic and authoritative)
Reflection on learning process	Reflection on the learning process or trajectory does not take place	Teacher reports on what has been achieved or makes learning process/ trajectory explicit, perhaps inviting reflection	Reflection on learning process or trajectory is discussed with students or by students themselves. This reflection can include a joint understanding of what has been achieved, evaluation of how activities have gone and an understanding of where learning is headed. Includes students' own metacognitive reflection on learning process itself.
Focusing on dialogue	No explicit focus on dialogue is apparent	Teacher introduces, models or reminds students of target dialogic practices, e.g. ground rules to be followed, inclusive turn taking etc.	Teacher and students or students themselves negotiate target dialogic practices, e.g. ground rules, perhaps along with reminders/ modelling. It may also include students taking responsibility for managing the dialogue, as well as students being involved in evaluating effectiveness of dialogic practices.
Student participation	Public exchanges consist of questioning and succinct students' contributions or Students do not have opportunities to discuss their ideas publicly	Students express their ideas publicly at length in whole-class situation and group work (consider what is audible in the video), but **they don't engage** with each other's ideas	Students express their ideas publicly at length in whole-class situation and group work (consider what is audible in the video) **AND** In doing so, they **engage with each other's' ideas** (e.g. 'It's a bit like what Shootle said but….', 'Sam had such a great idea, look [demonstrates]')

Appendix C. Teacher questionnaire

(Howe et al., 2019)

Section 1 – Your Class

How many children are in your class?

☐

How many of them are eligible for free school meals?

☐

For how many is English NOT the ffrst language?

☐

How many of those children are fiuent or almost fiuent in English?

☐

How many are registered with special needs?

☐

How many others would you say have special needs?

☐

For the following questions please tick the box that best represents what happens.

How often do the children in your class do the following?

	At least once a week	Several times a term	Once or twice a term	Once or twice a year	Never
Do homework					
Go on out-of-school trips					
Take physical exercise					
Take written tests					

Section 2 – Use of resources

How often do the children use the following in class?

	At least once a week	Several times a term	Once or twice a term	Once or twice a year	Never
Calculators in maths					
Computers (including portable devices)					
Interactive whiteboard or interactive display screen					
Video, animation/simulation or other visual media					
Concept mapping or mind mapping					

In your class, how often do you:

	At least once a week	Several times a term	Once or twice a term	Once or twice a year	Never
Teach generic thinking skills					
Integrate two or more areas of the curriculum					

Section 3 – Your teaching

Choose the answer that best reflects your level of agreement

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I present new topics to the class through lecture-style presentation					
We complete one topic before moving on to the next					
I start a topic with general theories and move on to examples					
I use worked examples					
We revisit previously taught topics					
I start a topic with concrete examples and then teach the theory					

Appendix D. Confounding variables

Information taken from: Howe, C. (2016). *Potential confounds*. Unpublished ESRC project design document

Confound category	Variable name	Description	Data source	Included in RQ2
Teacher practice	Homework	Frequency of homework	Teacher Questionnaire	X
	Trips	Frequency of trips	Teacher Questionnaire	√
	Exercise	Frequency of physical exercise	Teacher Questionnaire	X
	Tests	Frequency of written tests	Teacher Questionnaire	√
	Calculators	Frequency of calculator use in lessons	Teacher Questionnaire	√
	Computers	Frequency of computer use in lessons	Teacher Questionnaire	X
	IWB	Frequency of IWB use in lessons	Teacher Questionnaire	X
	VideoAmin	Frequency of video/simulation/animation/visual media use in lessons	Teacher Questionnaire	√
	Mapping	Frequency of concept mapping/mind mapping	Teacher Questionnaire	√
	Thinking	Frequency of taught generic thinking skills	Teacher Questionnaire	X
	Integration	Frequency teacher integrates two or more areas of the curriculum	Teacher Questionnaire	√
	Lecture	Extent teacher uses lecture-style presentation	Teacher Questionnaire	√
	WorkedEG	Extent teacher uses worked examples	Teacher Questionnaire	√
	Spaced	Extent teachers exhibit spaced practice (e.g. revisit previously taught topics)	Teacher Questionnaire	√
	Inductive	Extent teachers display inductive teaching (e.g. start with concrete examples then teach theory).	Teacher Questionnaire	√

(*continued on next page*)

(*continued*)

Confound category	Variable name	Description	Data source	Included in RQ2
	Group	Collaborative group work observed in lessons.	Group work observation	√
	Peer Tutoring	Peer tutoring observed in group work.	Group work observation	X
	Goals	Ratings of the clarity of lesson goals	Lesson ratings	√
	Feedback	Ratings of the quality of feedback students received in the lesson	FFT rating scale	√
	Classroom relations	Rating of the quality of teacher-student relations	FFT rating scale	√
	Behaviour management	Ratings of teachers' awareness of student misbehaviour.	FFT rating scale	√
Student characteristics	NFER Maths	Prior attainment score of National Foundation of Educational Research Maths Test.	NFER pre Test	√
	NFER English	Prior attainment score of National Foundation of Educational Research Reading Test.	NFER pre Test	√
	PASS	Pupils attitudes to school test (GL Assessment).	Student assessment	√
	PI Talk	Parental involvement (frequency of talk with parent about learning).	Student Questionnaire	√
	PI Visit	Parental involvement (frequency of parents' communication with teacher)	Student Questionnaire	√
	PropNoFSM	Proportion of students that are not eligible for free school meals	Teacher Questionnaire	√
	PropNoESL	Proportion of students that do not have English as their second language.	Teacher Questionnaire	X
	PropFluent	Proportion of students that are fluent in English.	Teacher Questionnaire	X
	PropNoRegSEN	Proportion of students that are not registered with special needs.	Teacher Questionnaire	√

Appendix E. Student questionnaire

(Howe et al., 2019)

Questionnaire for children

What is the name of your school?

When is your birthday?

How many schools have you been to before coming to this one? (Do not include nursery).

☐

Which language do you speak at home?	English Other

Tick the box that best describes how true each of these is for you

In the last year how often have you talked about each of these with your parents or guardians?

	Not at all	Once or twice	Three or more times
School activities or events that interest you			
Things you have learned in class			

In the last year have your parents or guardians done any of these things?

	Yes	No	I don't know
Come to a school meeting			
Spoken to your teacher			
Visited your class			
Come to a school event in which you participated			

How often do your parents or guardians do these things?

	Often	Sometimes	Hardly ever	Never
Check if you have done your homework				
Give you jobs to do in your house				
Limit your screen time				
Limit the amount of time you go to play with friends				

How often do you talk to these people about what you learn at school?

	Often	Sometimes	Hardly ever	Never
Your father or a male guardian				
Your mother or a female guardian				
Teachers				
Other adult relatives or friends				
Relatives or friends your own age				

References

Ackermann, E. (2001). Piaget's constructivism, Papert's constructionism: What's the difference. *Future of Learning Group Publication, 5*, 438.

Alexander, R. (2018). Developing dialogic teaching: Genesis, process, trial. *Research Papers in Education, 33*, 561–598.

Alexander, R. J. (2008). *Towards dialogic teaching. Rethinking classroom talk*. Cambridge: Dialogos.

Amodia Bidakowska, A. (2021). *Disciplinary dialogues: Examining the influence of subject on classroom dialogue, activity type and learning outcomes in English primary schools*. Doctoral dissertation: University of Cambridge.

Asterhan, C. S. C., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology, 99*, 626–639.

Aydeniz, M., & Ozdilek, Z. (2016). Assessing and enhancing pre-service science teachers' selfefficacy to teach science through argumentation: Challenges and possible solutions. *International Journal of Science and Mathematics Education, 14*, 1255–1273.

Bakhtin, M. (1986). *Speech Genres and other Late Essays*. Austin: University of Texas.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefit of teaching. *Journal of Educational Psychology, 72*, 593–604.

Barmby, P., Harries, T., Higgins, S., & Suggate, J. (2009). The array representation and primary children's understanding and reasoning in multiplication. *Educational Studies in Mathematics, 70*, 217–241.

Belenky, D. M., & Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review, 26*, 27–50.

Berns, R. G., & Erickson, P. M. (2001). *Contextual teaching and learning: Preparing students for the new economy*. Columbus: National Dissemination Center for Career and Technical Education.

Boyd, M., & Kong, Y. (2017). Reasoning words as linguistic features of exploratory talk: Classroom use and what it can tell us. *Discourse Processes, 54*, 62–81.

Boyd, M. P., Chiu, M. M., & Kong, Y. (2019). Signaling a language of possibility space: Management of a dialogic discourse modality through speculation and reasoning word usage. *Linguistics and Education, 50*, 25–35.

British Educational Research Association. (2011). *Ethical guidelines for educational research*. BERA.

British Psychological Society. (2009). Code of ethics and conduct. BPS.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema, & T. R. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19–32). Mahwah, NJ: Erlbaum.

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*, Article 101564.

Christodoulou, A., & Osborne, J. (2014). The science classroom as a site of epistemic talk: A case study of a teacher's attempts to teach science based on argument. *Journal of Research in Science Teaching, 51*, 1275–1300.

Collective, T.-S. E. D. A. (2023). *Toolkit for systematic educational dialogue analysis (T-SEDA): A resource for inquiry into practice* (Vol. 9). University of Cambridge. http://bit.ly/T-SEDA.

Cook, V., Warwick, P., Vrikki, M., Major, L., & Wegerif, R. (2019). Developing material-dialogic space in geography learning and teaching: Combining a dialogic pedagogy with the use of a microblogging tool. *Thinking Skills and Creativity, 31*, 217–231.

Craft, A., Cremin, T., Burnard, P., Dragovic, T., & Chappell, K. (2013). Possibility thinking: Culminative studies of an evidence-based concept driving creativity? *Education, 41*, 538–556.

Creswell, J. W., & Plano Clark. (2011). *Designing and conducting mixed methods research*. California: Sage Publications.

Danielson, C. (2013). *The framework for teaching. Evaluation instrument*. Princeton, NJ: The Danielson Group.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 287–312.

Economic and Social Research Council. (2012). *Framework for research ethics*. ESRC.

Erduran, S., Ardac, D., & Yakmaci-Guzel, B. (2006). Learning to teach argumentation: Case studies of pre-service secondary science teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 2*, 1–14.

Field, A. (2009). *Discovering statistics using SPSS*. London: Sage Publications.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review, 26*, 9–25.

Gainsburg, J. (2008). Real-world connections in secondary mathematics teaching. *Journal of Mathematics Teacher Education, 11*, 199–219.

GL Assessment. (2013). GL assessment. Retrieved from www.gl-assessment.co.uk/products/pupil-attitudes-to-self-and-school-pass/.

Hattie, J. A. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Hennessy, S., Howe, C., Mercer, N., & Vrikki, M. (2020). Coding classroom dialogue: Methodological considerations for researchers. *Learning, Culture and Social Interaction, 25*, Article 100404.

Hennessy, S., Rojas-Drummond, S., Higham, R., Torreblanca, O., Barrera, M. J., Marquez, A. M., García Carrión, R., Maine, F., & Ríos, R. M. (2016). Developing an analytic coding scheme for classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9*, 16–44.

Hennessy, S., Wishart, J., Whitelock, D., Deaney, R., Brawn, R., La Velle, L., McFarlane, A., Ruthann, K., & Winterbottom, M. (2007). Pedagogical approaches for technology-integrated science teaching. *Computers & Education, 48*, 137–152.

Howe, C. (2010). Peer dialogue and cognitive development. A two-way relationship? In K. Littleton, & C. Howe (Eds.), *Educational dialogues: Understanding and promoting productive interaction* (pp. 32–47). London: Routledge.

Howe, C. (2016a). *Adapting SEDA for our project*. Unpublished ESRC project design document.
Howe, C. (2016b). *Using Hattie (2009) to identify and assess confounds*. Unpublished ESRC project document.
Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education, 43*, 325–356.
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.
Howe, C., Ilie, S., Guardia, P., Hofmann, R., Mercer, N., & Riga, F. (2015). Principled improvement in science: Forces and proportional relations in early secondary-school teaching. *International Journal of Science Education, 37*, 162–184.
Howe, C., & Tolmie, A. (2003). Group work in primary school science: Discussion, consensus and guidance from experts. *International Journal of Educational Research, 39*, 51–72.
Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., Jessiman, E., & Donaldson, C. (2007). Group work in elementary science: Towards organisational principles for supporting pupil learning. *Learning and Instruction, 17*, 549–563.
Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*, 454–455.
Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. MET Project: Bill & Melinda Gates Foundation.
Kuhn, D. (2010). Teaching and learning science as argument. *Science Education, 94*, 810–824.
Kyriacou, C., & Issitt, J. (2008). *What characterizes effective teacher-pupil dialogue to promote conceptual understanding in mathematics key stages 2 and 3?* Social Science Research Unit: Institute of Education, University of London.
Larrain, A., Freire, P., & Howe, C. (2014). Science teaching and argumentation: One-sided versus dialectical argumentation in Chilean middle school science lessons. *International Journal of Science Education, 36*, 1017–1036.
Larrain, A., Howe, C., & Cerda, J. (2014). Argumentation in whole-class teaching and science learning. *Psykhe, 23*, 1–15.
Larrain, A., Howe, C., & Freire, P. (2018). 'More is not necessarily better': Curriculum materials support the impact of classroom argumentative dialogue in science teaching on content knowledge. *Research in Science & Technological Education, 36*, 282–301.
Lefstein, A. (2010). More helpful as problem than solution: Some implications of situating dialogue in classrooms. In K. Littleton, & C. Howe (Eds.), *Educational dialogues: Understanding and promoting productive interaction* (pp. 170–191). London: Routledge.
Lehesvuori, S., Hähkiöniemi, M., Ketonen, L., Lerkkanen, M. K., Pöysä, S., & Pakarinen, E. (2021). Reflections on dialogicity: Challenges and suggestions by mathematics student teachers. *Learning, Culture and Social Interaction, 31*, Article 100567.
Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex.
Leonard, S., Lamb, H., Howe, P., & Choudhoury, A. (2017). 'State of the nation' report of UK primary science education. Wellcome Trust Primary Science Campaign baseline research https://wellcome.ac.uk/sites/default/files/state-of-the-nation-report-of-uk-science-education.pdf.
Littleton, K., & Howe, C. (2010). *Educational dialogues: Understanding and promoting productive interaction*. Abingdon: Routledge.
Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. Abingdon: Routledge.
Ludvigsen, S., & Steier, R. (2019). Reflections and looking ahead for CSCL: Digital infrastructures, digital tools, and collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 14*(4), 415–423.
Maine, F. (2013). How children talk together to make meaning from texts: A dialogic perspective on reading comprehension strategies. *Literacy, 47*, 150–156.
McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching, 48*, 793–823.
Mercer, N. (1995). *The guided construction of knowledge*. Bristol: Multilingual Matters.
Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education, 40*, 430–445.
Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal, 30*, 359–377.
Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. London: Routledge.
Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education, 20*, 507–528.
Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education, 27*, 283–297.
Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Milton Keynes: Open University Press.
Muhonen, H., Pakarinen, E., Poikkeus, A. M., Lerkkanen, M. K., & Rasku-Puttonen, H. (2018). Quality of educational dialogue and association with students' academic performance. *Learning and Instruction, 55*, 67–79.
Muhonen, H., Rasku-Puttonen, H., Pakarinen, E., Poikkeus, A.-M., & Lerkkanen, M. K. (2016). Scaffolding through dialogic teaching in early school classrooms. *Teaching and Teacher Education, 55*, 143–154.
Nassaji, H., & Wells, G. (2000). What's the use of triadic dialogue? An investigation of teacher student interaction. *Applied Linguistics, 27*, 333–363.
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
Newton, P., Driver, R., & Osborne, J. F. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education, 21*, 553–576.
NFER, Centre for Assessment. (2016). *Technical information for NFER tests in reading and mathematics suite 2*. Slough: NFER.
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*, 135–198.
O'Connor, C., Michaels, S., & Chapin, S. (2015). Scaling down to explore the role of talk in learning: From district intervention to controlled classroom study. In L. B. Resnick, C. S. C. Asterhan, & Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 111–126). Washington, DC: American Educational Research Association.
O'Connor, M. C. (2001). Can any fraction be turned into a decimal? A case study of a mathematical group discussion. *Educational Studies in Mathematics, 46*, 143–185.
O'Donnell, A. M. (2006). The role of peers and group learning. In P. Alexander, & P. Winne (Eds.), *Handbook of educational psychology* (pp. 781–802). Mahwah: Lawrence Erlbaum.
Omland, M., Ludvigsen, S. R., & Rødnes, K. A. (2022). *The role of querying: Investigating subject-oriented meaning-making* (p. 100599). Culture and Social Interaction: Learning.
Osborne, J., & Chin, C. (2010). The role of discourse in learning science. In K. Littleton, & C. Howe (Eds.), *Educational dialogues: Understanding and promoting productive interaction* (pp. 88–102). Abingdon: Routledge.
Reznitskaya, A., & Gregory, M. (2013). Student thought and classroom language: Examining the mechanisms of change in dialogic teaching. *Educational Psychologist, 48*, 114–133.
Rojas-Drummond, S., Mazón, N., Fernández, M., & Wegerif, R. (2006). Explicit reasoning, creativity and co-construction in primary school children's collaborative activities. *Journal of Thinking Skills and Creativity, 1*, 84–94.
Schindler, A. K., Gröschner, A., & Seidel, T. (2015). Teaching science effectively: A case study on student verbal engagement in classroom dialogue. *Orbis Scolae, 9*, 9–34.
Schwarz, B. B., Hershkowitz, R., & Prusak, N. (2010). Argumentation and mathematics. In K. Littleton, & C. Howe (Eds.), *Educational dialogues: Understanding and promoting productive interaction* (pp. 103–127). Abingdon: Routledge.
Smith, F., Hardman, F., Wall, K., & Mroz, M. (2004). Interactive whole class teaching in the National Literacy and Numercy strategies. *British Educational Research Journal, 30*, 395–411.
Stylianides, G. J., & Stylianides, A. J. (2020). Posing new researchable questions as a dynamic process in educational research. *International Journal of Science and Mathematics Education, 28*, 583–598.
Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. California: Sage Publications.

Veenman, S., Denessen, E., Van Den Akker, A., & Van Der Rijt, J. (2005). Effects of a cooperative learning program on the elaborations of students during help seeking and help giving. *American Educational Research Journal, 42*, 115–151.
Von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching, 45*, 101–131.
Vrikki, M., Wheatley, L., Howe, C., Hennessy, S., & Mercer, N. (2019). Dialogic practices in primary school classrooms. *Language and Education, 33*, 85–100.
Vygotsky, L. S. (1978). *Mind in society*. Cambridge: Harvard University Press.
Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology, 79*, 1–28.
Webb, N. M., & Farivar, S. (1999). Developing productive group interaction in middle school mathematics. In A. M. O'Donnell, & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 117–149). Mahwah: Lawrence Erlbaum.
Webb, N. M., Franke, M. L., Ing, M., Turrou, A. C., Johnson, N. C., & Zimmerman, J. (2019). Teacher practices that promote productive dialogue and learning in mathematics classrooms. *International Journal of Educational Research, 97*, 176–186.
Wegerif, R. (2006). A dialogic understanding of the relationship between CSCL and teaching thinking skills. *International Journal of Computer-Supported Collaborative Learning, 1*, 143–157.
Wegerif, R. (2011). Towards a dialogic theory of how children learn to think. *Thinking Skills and Creativity, 6*, 179–190.
Wells, G. (1999). *Dialogic inquiry, toward a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.
Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. *Journal of the Learning Sciences, 15*, 379–428.
Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22*, 390–408.
Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*, 35–62.