



Developing dialogic teaching: genesis, process, trial

Robin Alexander

Wolfson College, University of Cambridge, Cambridge, UK

ABSTRACT

This paper considers the development and randomised control trial (RCT) of a dialogic teaching intervention designed to maximise the power of classroom talk to enhance students' engagement and learning. Building on the author's earlier work, the intervention's pedagogical strand instantiates dialogic teaching not as a single, circumscribed 'method' but as an interlocking set of permissive repertoires through which, steered by principles of procedure, teachers energise their own and their students' talk. The repertoires are directed both to teaching's improvement and to its larger epistemological, cultural and civic purposes. Its professional strand entailed teacher induction and training followed by a cyclic programme of planning, target-setting and review using mentoring and video/audio analysis. Supported by the UK Education Endowment Foundation it was piloted in London and trialled in three other UK cities with combined intervention/control cohorts of nearly 5000 year 5 (4th grade) students and 208 teachers. The independent evaluation calculated that after 20 weeks students in the intervention group were two months ahead of their control group peers in English, mathematics and science tests; while coded video data showed that the changes in both teacher and student talk were striking and in the direction intended. The RCT methodology affords limited explanatory purchase but insights are available from other studies. These, together with contingent questions and future possibilities, are discussed in the paper's conclusion.

ARTICLE HISTORY

Received 10 March 2018
Accepted 23 May 2018

KEYWORDS

Dialogic teaching; classroom talk; epistemology; pedagogy; randomised control trial

Introduction

This paper discusses the Cambridge Primary Review Trust / University of York Dialogic Teaching Project and the thinking that informed it. Funded 2014–2017 by the UK Education Endowment Foundation (EEF), the project piloted and implemented a programme designed to energise classroom talk and thereby enhance students' engagement, learning and attainment in contexts of social and educational disadvantage. In line with the EEF 'what works' funding criteria, the intervention was based on an existing approach – in this case a specific version of dialogic teaching (Alexander 2017a, 2017b) – for which there existed *prima facie* evidence of efficacy (Alexander 2003, 2005a, 2005b; Lefstein and Snell 2011), and it was subjected to randomised control trial (RCT) by an independent team.¹

The intervention had two strands, pedagogical and professional. Being contingent, both are described here. The ensuing account of the intervention's implementation and impact draws partly on reports from the externally led RCT, which focused chiefly on tested student learning outcomes (Jay et al. 2017); and partly on the project's in-house evaluation, which used interviews and coded video data to track the intervention's reception and progress, and its effect on the classroom talk that was the project's central concern (Alexander et al. 2017).²

Part 1 of this paper traces and describes the intervention's version of dialogic teaching and the professional development programme through which it was realised. Part 2 outlines the methodology and findings of the two evaluation exercises. Part 3 ventures conclusions, discussion and issues for further exploration.

Part 1 – Intervention

The intervention as pedagogy

Evidential basis and general character of the approach

There is no single and agreed definition of the term 'dialogic teaching'. It would be somewhat paradoxical if having intimated the liberality of dialogue this were not the case. Yet the various shades of meaning intersect with reasonable coherence. So in this paper the term will be used stipulatively to connote a pedagogy of the spoken word that is manifestly distinctive while being grounded in widely accepted evidence and in discourse and assumptions that have much in common.

The evidence has a number of strands – psycholinguistic, sociolinguistic, neuroscientific, philosophical, pedagogical – but in this context three are pre-eminent and should be briefly rehearsed. First, psychological research, increasingly supported by neuroscience, demonstrates the intimate and necessary relationship between language and thought, and the power of spoken language to enable, support and enhance children's cognitive development, especially during the early and primary years (for example, Britton 1969; Wood 1998; Tough 1979; Bruner 1983, 1996; Bruner and Haste 1987; Goswami 2015).

Second, classroom research testifies to the way that the recitation or IRE (initiation-response-evaluation) exchange structure, which centres on closed questions, recall answers and minimal feedback and in many schools remains the pedagogical default, resists change despite abundant evidence that it wastes much of talk's discursive, cognitive and educational potential (Barnes, Britten, and Rosen 1969, Sinclair and Coulthard 1975³; Barnes and Todd 1977; Mehan 1979; Nystrand et al. 1997; Galton et al. 1999; Alexander 2001, 2008; Cazden 2001; Hardman, Smith, and Wall 2003; Mortimer and Scott 2003; Smith et al. 2004; Galton 2008; Mehan and Cazden 2015; Resnick, Asterhan, and Clarke 2015).

Third, various remedies have been mooted. Several are exemplified in Mercer and Hodgkinson 2008; many more in Resnick, Asterhan, and Clarke 2015; while broad trends are identified by Lefstein and Snell 2014 and systematically compared by Kim and Wilkinson (2018). However, though sharing a commitment to elevating the profile and power of classroom talk, and though usually defined by their advocates as 'dialogic', as noted above they are far from identical, especially in respect of their scope. Some (e.g. Reznitskaya 2012; Reznitskaya and Gregory 2013) advocate a specific practice or method, while for others dialogue is situated less exclusively within a wider interactive spectrum. Some focus largely or exclusively on the talk of the teacher (e.g. Wragg and Brown 1993, 2001) or the student

(e.g. Norman 1992; Mercer 2000; Dawes, Mercer, and Wegerif 2004), while others, including the one under discussion here, aim to attend to both, arguing that although student talk must be our ultimate preoccupation because of its role in the shaping of thinking, learning and understanding, it is largely through the teacher's talk that the student's talk is facilitated, mediated, probed and extended – or not, as the case may be. Hence, the effort, to which all interested in dialogic pedagogy subscribe, to move beyond the monologic dominance of recitation/IRE and develop patterns of classroom interaction that open up students' speaking and listening, and hence their thinking, and which strive to distribute the ownership of talk more equitably.

In differentiating the various pedagogical approaches, Lefstein and Snell (2014) show how they vary not only in respect of strategy but also in the way they reflect contrasting notions of dialogue's nature and purposes, whether these be the perennial interplay of voices in culture and history (Bakhtin), the dialectic of argumentation and critique (Socrates), collaborative thinking as a route to acculturation as well as learning (Vygotsky), the nurturing of human relations (Buber) or human and social empowerment (Freire). Similarly, Alexander draws on his transnational and cross-cultural classroom research to show how classroom cultures, values and interactions are variously shaped by collective, communitarian and individualist emphases in accounts of social relations and by culturally located stances on human development, the nature and acquisition of knowledge and the act of teaching. Eschewing the popular dichotomising of teacher-centred/child-centred or transmission/discovery, he differentiates these as 'transmission', 'initiation', 'negotiation', 'facilitation' and 'acceleration' (Alexander 2001, 2008, 2009).

Given this diverse cultural and philosophical genealogy, it is inevitable that strategies for talk reform have different emphases. Here, again, the framework of Lefstein and Snell (2014) is helpful, though perhaps not sufficiently extensive. They identify four paradigms and typical proponents: dialogically organised instruction (Nystrand et al. 1997, 2003), exploratory talk (Mercer 2000; Mercer and Littleton 2007), accountable talk (Resnick, Michaels, and O'Connor 2010) and dialogic teaching (Alexander 2001, 2008, 2017a).

In turn, Alexander's take on dialogic teaching owes most to the foundational works of Vygotsky (1962, 1978), Bruner (1983, 1996) and Bakhtin (1981, 1986) while strategically it is closest to those of Nystrand et al. and Resnick, Michaels, and O'Connor (2010). Yet, it is also *sui generis*, for, as noted above, it devotes equal attention to the quality of teacher and student talk, and to the agency of others – fellow students as well as teachers – in the latter. It also rejects the view that there is one right way to maximise talk's quality and power (for example, through small group discussion, or 'interactive whole class teaching' as mandated in 1998 by the UK government's National Literacy Strategy), and instead advances the need for every teacher to develop a broad repertoire of talk-based pedagogical skills and strategies and to draw on these to expand and refine the talk repertoires and capacities of their students. Acknowledging the uniqueness of each classroom's personalities and circumstances it gives the teacher the responsibility for deciding how the repertoire should be applied. This responsibility is progressively shared with students, the development and autonomous deployment of whose own talk repertoires is the ultimate goal.

This commitment to repertoire combined with teacher and student agency is fundamental. It reaches back to Alexander's contribution to the UK government's 'three wise men' enquiry of 1991–1992 which made a similar case for repertoire-based teaching (Alexander, Rose, and Woodhood 1991), and to his objections to the either/or, them-and-us,

dichotomising tendency that has long characterised much of the wider educational and pedagogical discourse (Alexander 1984, 2008, chapter 4).

The approach is no less distinctive for treating talk not in isolation but as part of a generic model of teaching in which interaction takes its place alongside, and is contingent on, the invariants of *frame* (space, student organisation, time, curriculum, rules and routines), *form* (the lesson) and *act* (task, activity, interaction and judgement). That model was devised to make sense, as far as possible in a non-ethnocentric manner, of observational, video and interview data from classrooms in England, France, India, Russia and the United States (The model is described and applied in Alexander 2001; 265–528, and further developed in Alexander 2008, 45–50 and 180–183). Practical examples of its application in the present project might include the attention given to the relationship between the character and quality of talk and the teacher's handling of time and pace, student grouping and classroom layout, to the balance and relationship of oral and written activity, and to the mutually reinforcing and overlapping processes of dialogic interaction and assessment for learning (Black and Wiliam 1998; Black et al. 2003).

A dialogic teaching framework

As shown in Table 1, the dialogic teaching framework under discussion, for which space permits outline rather than detail, has four main components: justifications, principles, repertoires and indicators. The repertoires are the heart of the operation. They are guided and refreshed by the principles and indicators, while the justifications provide their springboard.

Table 1. Elements in a framework for dialogic teaching.

Justifications	Communicative – Social – Cultural – Political/civic – Psychological – Neuroscientific – Pedagogical
Principles	Collective – Reciprocal – Supportive – Cumulative – Purposeful
Repertoires	
<i>Interactive settings</i>	<i>Whole class – Group work (teacher-led) – Group work (student-led) – One-to-one (teacher–student) – One-to-one (student–student)</i>
<i>Everyday talk</i>	<i>Transactional – Expository – Interrogatory – Exploratory – Expressive – Evaluative</i>
<i>Learning talk</i>	<i>(a) Narrate – Explain – Speculate – Imagine – Explore – Analyse – Evaluate – Question – Justify – Discuss – Argue</i> <i>(b) Listen – Think about what we hear – Give others time to think – Respect others' views</i>
<i>Teaching talk</i>	<i>Rote – Recitation – Instruction – Exposition – Discussion – Dialogue</i>
<i>Questioning</i>	<i>Character – Response cue – Participation cue – Wait/thinking time – Feedback – Purpose – Structure</i>
<i>Extending</i>	<i>Time to think – Say more – Revoice – Rephrase/repeat – Evidence of reasoning – Challenge or counter-example – Agree/disagree – Add on – Explain what someone else means</i>
Indicators	

Justifications. Education is an ethical as well as instrumental endeavour, so teachers must consider why talk in general, and talk of the kind commended, are so important. Seven justifications are proposed, listed here as headings only but elsewhere enunciated more fully (Alexander 2017a, 9–14 and 33–34):

- Communicative
- Social
- Cultural
- Political/civic
- Psychological
- Neuroscientific
- Pedagogical

The first four justifications take up ethical positions, admittedly sustained by pragmatism. They argue that children need to be able to communicate, build relationships, participate in their culture, value collective identity and cohesion, and become engaged and active citizens. Language in all its forms is viewed as vital for each of these. The last three justifications derive from published evidence, of the kind referred to earlier, concerning the relationship between spoken language, synaptogenesis and cognitive development, and the character, possibilities and pitfalls of classroom talk as observational research has charted it.

Generous though this justificatory catalogue may be, it requires further comment on how the intended pedagogy relates to the broader educational aims it purports to serve. Between 2006 and 2010, the author led the Cambridge Primary Review, the UK's most comprehensive enquiry into the condition and future of primary education for half a century (Alexander 2010; Alexander et al. 2010).⁴ One of the review's ten metathemes was educational aims for the twenty-first century: what they currently are, and what they might be. After consulting widely in the UK and trawling other educational systems, the enquiry drafted, and after consultation confirmed, 12 aims for public education grouped in fours under the headings of 'The individual', 'Self, others and the wider world', and 'Learning, knowing and doing'.⁵ As will be inferred from footnote 5, several of these resonate with the notion of dialogic teaching. In relation to the citizenship aim, for example, Michaels et al. point out:

Dialogue and discussion have long been linked to theories of democratic education. From Socrates to Dewey and Habermas, educative dialogue has represented a forum for learners to develop understanding by listening, reflecting, proposing and incorporating alternative views. For many philosophers, learning through discussion has also represented the promise of education as a foundation for democracy. (Michaels, O'Connor, and Resnick 2008, 296)

The democratic claim for dialogic pedagogy has been furthered empirically by an extensive literature review on citizenship education (Deakin Crick et al. 2005), though lest this looks altogether too neat and unproblematic, Michaels et al. warn that with classroom discourse as with civic discourse 'the gap between the idealised and realised is daunting' (*ibid*).

One Cambridge Primary Review aim is more synoptic than the others, commanding attention to the idea that dialogue is much more than classroom talk and that education itself is dialogic⁶:

Enacting dialogue. To help children grasp that learning is an interactive process and that understanding builds through joint activity between teacher and pupil and among pupils in collaboration, and thereby to develop pupils' increasing sense of responsibility for what and how they learn. To help children recognise that knowledge is not only transmitted but also negotiated and re-created; and that each of us in the end makes our own sense out of the meeting of knowledge both personal and collective. *To advance a pedagogy in which dialogue is central: between self and others, between personal and collective knowledge, between present and past, between different ways of making sense.* (Alexander 2010, 399, my italics)

If one accepts the dialectical account of knowledge and its acquisition and growth intimated here,⁷ a pedagogy hegemonised by recitation/IRE is untenable not so much on the familiar grounds of efficiency – for by its own lights it can be very efficient – as because its account of efficiency is predicated on teaching as transmitting, learning as receiving and knowing as repeating. A dialogic pedagogy doesn't necessarily presuppose a dialogic epistemology, but a dialogic epistemology cannot realistically be fostered by other than a dialogic pedagogy.

Principles. The model's second element adumbrates five principles or tests of dialogic teaching.

- Collective (the classroom is a site of joint learning and enquiry)
- Reciprocal (participants listen to each other, share ideas and consider alternative viewpoints)
- Supportive (participants feel able to express ideas freely, without risk of embarrassment over 'wrong' answers, and they help each other to reach common understandings)
- Cumulative (participants build on their own and each other's contributions and chain them into coherent lines of thinking and understanding)
- Purposeful (classroom talk, though open and dialogic, is structured with specific learning goals in view).

These again subdivide (Alexander 2017a, 27–28). Collectivity, reciprocity and supportiveness characterise the classroom culture and pattern of relationships within which dialogue is most likely to prosper, its learning potential has the best chance of being realised, and students will be most at ease in venturing and discussing ideas. These three principles are consistent with the epistemological stance encapsulated in the aim from the Cambridge Primary Review cited above. But as the final principle reminds us, classroom discussion, though valuable in itself, is also a means to an educational end, and it must therefore square the circle of a Bakhtinian commitment to dialogue as unending and a pedagogical commitment to the student's understanding and mastery of specific ideas. Similarly, attention to the principle of cumulation, which underpins enquiry and knowledge growth in academic communities as well as classrooms, ensures that discussion is genuinely dialectical yet builds on what has gone before, advances understanding and is not merely circular.

Cumulation, we have found, is the most difficult of the principles to enact, because while collectivity, reciprocity and support relate to the *dynamics* of talk, cumulation attends to its *meaning* and, therefore, simultaneously tests teachers' mastery of the epistemological terrain being explored, their insight into students' understandings within that terrain, and their interactive skill in taking those understandings forward (Alexander 2017a, 49–53)⁸.

As an example of the definitional minefield that is dialogic teaching we might note that Mercer and Littleton characterise talk as 'cumulative' when 'speakers build positively *but uncritically* on what others have said' (Mercer and Littleton 2007, 51; my italics). While excluding critical interventions may meet our third criterion (supportiveness) it may also discourage argumentation and propel discussion towards premature or unfounded consensus, thereby foreclosing cumulation of the more exacting kind referred to in the previous paragraph. Within the dialogic teaching framework outlined here, cumulation is probably closer to 'accountable talk' as defined by Resnick, Michaels and O'Connor (Michaels, O'Connor, and Resnick 2008; Resnick, Michaels, and O'Connor 2010): 'Speakers make an effort to get their facts right and make explicit their evidence behind their claims or explanations. They challenge each other when evidence is lacking or unavailable.' (Michaels, O'Connor, and Resnick 2008, 283). Mercer and Littleton classify this kind of talk as 'exploratory' which suggests that their use of the term 'cumulative' deals more with the social relations of talk than its substance and trajectory.

The five principles are not confined to any one preferred pattern of organisation, and our interest in building a comprehensive pedagogical repertoire contrasts, say, with Mercer's initial preoccupation, following Barnes and Todd (1977, 1995), with the dynamics and

benefits of small group discussion (Mercer and Littleton 2007), which in the present model is just one of the several patterns of interactive organisation that dialogue can enrich, albeit an important one that should be fully exploited.

Repertoires. The framework includes six repertoires (Alexander 2017a, 37–40):

- (1) Interactive settings
- (2) Everyday talk
- (3) Learning talk
- (4) Teaching talk
- (5) Questioning
- (6) Extending

Repertoire 1 – interactive settings. Classrooms allow three organisational settings or modalities for student–teacher and student–student interaction: whole class, group and individual. Given that the actors include both teachers and students, these expand to five:

- Whole class teaching (teacher–student)
- Group work (teacher–student, teacher-led)
- Group work (student–student, student-led)
- One-to-one (teacher–student)
- One-to-one (student–student pairs)

Being forms of organisation rather than kinds of talk, these in effect mark one axis of a grid, while the various kinds of talk outlined in repertoires 2–6 below occupy the other axis. The resulting (virtual) grid immediately and vastly expands the possibilities captured by the framework, and reminds us that while group discussion, whole class teaching and paired talk offer distinct social, communicative and affective payoffs, the cognitive leverage they exert depends more on the character and quality of the talk being pursued than on their organisation as such, even though each organisational form entails opportunities and constraints that are different from the others.

Repertoire 2 – everyday talk. The other axis starts with six broad categories of talk to engender and sustain everyday interaction:

- Transactional
- Expository
- Interrogatory
- Exploratory
- Expressive
- Evaluative

This repertoire reminds us that whatever else schools do, they should equip children with the capacities to manage social encounters, tell and explain, ask different kinds of questions, explore ideas, articulate feelings and responses, and frame opinions and judgements. But in order to teach effectively, teachers themselves need to master, model and deploy this most basic range, remaining alive to the tendency of traditional teaching to omit the fourth and fifth above while restricting the scope of the others.

Repertoire 3 – learning talk. In classrooms, the kinds of everyday talk listed above can be expanded into eleven categories of student talk for learning:

- Narrate
- Explain
- Speculate
- Imagine
- Explore
- Analyse
- Evaluate
- Question
- Justify
- Discuss
- Argue

These are coupled with four conditions or capacities that students need to develop to allow such talk to happen and to take full advantage of its possibilities. They should:

- Listen
- Think about what they hear
- Give others time to think
- Respect alternative viewpoints

The last four are in the manner of norms or ground rules for discussion as proposed by Michaels, O'Connor, and Resnick (2008), Michaels and O'Connor (2015) and Mercer and Littleton (2007). The norms elaborate and facilitate the dialogic teaching principles of collectivity, reciprocity and supportiveness.

Repertoire 4 – teaching talk. It is axiomatic that teachers themselves need to command and preferably model the range of talk in repertoires 2 and 3. But while such repertoires are not unique to teachers, comparative classroom research has charted a narrower spectrum of talk strategies specific to teaching:

- Rote
- Recitation
- Instruction
- Exposition
- Discussion
- Dialogue

Although Philip Jackson was right to point out that unlike, say, 'doctors, lawyers, garage mechanics and astrophysicists' teachers lack an agreed technical vocabulary' (Jackson 1968, 143) – though if managerialist jargon and cliché count as technical vocabulary, the language of educational leadership, if not that of teaching, is changing fast – we know that in the classroom teachers do think and talk in professionally specific ways (Schön 1983), and recitation is perhaps the most uniquely teacherly kind of talk of all. The spectrum above arose from international classroom observation and video analysis (Alexander 2001), and serves as an empirical as well as conceptual corrective to the familiar opposition of 'transmission' and 'discovery' – telling children versus letting them find out for themselves. Currently, dare I

suggest, we might perhaps be wary of updating this, quasi-moralistically, to recitation (bad) versus dialogue (good). For as I note elsewhere:

There is a danger ... that we consign all but the last two of these forms of classroom talk to the despised archive of 'traditional' methods. In fact, exposition and recitation have an important role in teaching, for facts need to be imparted, information needs to be memorised, and explanations need to be provided, and even the deeply unfashionable rote has a place (memorising tables, rules, spellings and so on). However, the joint solving of problems through discussion, and the achievement of common understanding through dialogue, are undeniably more demanding of teacher skill than imparting information or testing recall through rote or recitation. (Alexander 2001, 526–527).

Dialogic teaching, therefore, encompasses the full range of teaching talk listed at Repertoire 4 but privileges the last two and especially dialogue in the more specific sense of interaction that, at best, extends the spectrum of student learning talk as in Repertoire 3. In contrast, though it does indeed have its place, the closed/recall questions that typically initiate recitation allow students to tell/narrate and, at a pinch, to explain, but not to speculate, imagine, explore, analyse, argue or ask questions of their own. Recitation is rarely other than monologic. Yet, as Nystrand et al. cogently conclude from their study of teaching in eighth and ninth grade literature classes:

The results of our study suggest that authentic questions, discussion, small-group work, and interaction, though important, do not categorically produce learning; indeed we observed many classes where this was not the case. We also found that recitation is not categorically ineffective; rather, its effectiveness depends on how teachers expand IRE sequences. The underlying epistemology of classroom interaction defines the bottom line for learning: what ultimately counts is *the extent to which instruction requires students to think, not just report someone else's thinking*. (Nystrand et al. 1997, 72, their italics).

Repertoire 5 – questioning

- Character: test, authentic
- Response cue: bidding (hands up to answer), nomination (question directed to specific student)
- Participation cue: rotation (short question and answer round the class), extension (longer exchanges confined to smaller number of students)
- Wait/thinking time: immediate, considered⁹
- Feedback: formative, evaluative
- Purpose: elicit, recall, develop, probe, manage
- Structure: closed, open, leading, narrow, discursive

Even allowing for progress since Flanders proposed his 'rule of two thirds', classrooms remain places where it is the teacher who asks most of the questions, so a dialogically informed questioning repertoire must allow for the possibility that students too will have questions to ask and must be encouraged and if necessary train them to do so. Repertoire 5, therefore, starts with the classic distinction of Nystrand et al. (1997) between 'test' and 'authentic' questions, which refers to the character and intention of questions that the teacher poses, and the latter's options for inviting and handling student responses, but ends with sub-repertoires of questioning purposes and structure that apply to *all* questions, regardless of who poses them.

Repertoire 6 – extending. This final repertoire, developed by Michaels and O’Connor (2012), was not in the dialogic teaching framework as initially ventured (Alexander 2004) but with the authors’ permission and following their own evaluation (Michaels and O’Connor 2015) it was incorporated into it for the present project, and to encouraging effect. ‘Extending’ proposes nine moves¹⁰ through which the teacher can help students to:

- Share, expand and clarify thinking
 - Time to think
 - Say more
 - Revoice
- Listen carefully to one another
 - Rephrase/repeat
- Deepen reasoning
 - Ask for evidence of reasoning
 - Challenge or counter-example
- Think with others
 - Agree/disagree and why
 - Add on
 - Explain what someone else means

This repertoire is the logical follow-on from questioning, and it offers as prompts utterances that are at once readily memorable and epistemologically positioned, and through which teachers can take student responses and contributions and build them into discussion chains, thereby modelling as well as advancing the principle of cumulation outlined above. However, practically suggestive though these moves are, Michaels and O’Connor echo Nystrand et al. (quoted above) in warning that ‘the simple deployment of talk moves does not ensure coherence in classroom discussion or robust student learning.’ (Michaels and O’Connor 2015, 358).

Indicators. The framework is completed by a list of 61 indicators that specify in practical terms how dialogic teaching looks and sounds (Alexander 2017a, 40–44). Space does not allow these to be listed here in full. Suffice it to say that they cover (i) the contexts within which dialogic teaching is placed and the classroom conditions that optimally support it (ii) the properties of the talk itself. The list is preceded by a warning:

What follows is intended to serve a heuristic purpose, not to be translated into a checklist to which teachers are required to conform. If that were to happen, its dialogic intention would be defeated. (Alexander 2017a, 41)

If a 61-item list, heuristic or otherwise, appears daunting or unmanageable, the framework ends with a more succinct summation of the kind of talk we are hoping for:

- *interactions* which encourage students to think, and to think in different ways
- *questions* which invite more than simple recall
- *answers* which are justified, followed up and built upon rather than merely received
- *feedback* which, as well as evaluating, leads thinking forward
- *contributions* which are extended rather than fragmented or prematurely closed

- *exchanges* which chain together into coherent and deepening lines of enquiry
- *discussion and argumentation* which probe and challenge rather than unquestioningly accept
- *scaffolding* which provides appropriate linguistic and/or conceptual tools to bridge the gap between present and intended understanding
- *professional mastery of subject matter* which is of the depth necessary to liberate classroom talk from the safe and conventional
- *time, space, organisation and relationships* which are so disposed and orchestrated as to make all this possible.

The ultimate test of genuinely dialogic teaching as defined here is captured in two quotations, the first of which we have already encountered:

What ultimately counts is the extent to which instruction requires students to think, not just report someone else's thinking. (Nystrand et al. 1997, 72).

If an answer does not give rise to a new question from itself, it falls out of the dialogue. (Bakhtin 1986, 168).

Here, Nystrand et al. remind us that supposedly 'dialogic' patterns of talk are not intrinsically productive, and that while classroom talk is in part directed to communicative facility and effectiveness, if its impact is not primarily *cognitive* then the prospects for learning – and indeed the value of what is communicated – are greatly diminished. Complementing that warning, and moving from the cognitive power of exchanges to their component moves, Bakhtin's sense of dialogue as an unending process or quest argues a shift in the centre of discursive gravity from what the teacher asks, instructs or tells – the main focus of traditional classroom observation instruments, perversely even those used by some committed dialogists (Alexander 2015, 433¹¹) – to what the pupil says and, especially, what the teacher *does* with what the pupil says.

Prior evidence of the effectiveness of the approach

So: dialogic teaching is here expanded or reduced to specific *justifications, principles, repertoires* and *indicators* in pursuit of the aim of socially and cognitively empowering talk. Although it is correct for the EEF project evaluation team to assert, as they do, that until this project there had been no randomised control trial of its version of dialogic teaching in operation (Jay et al. 2017), internal evaluations of its precursors in London and North Yorkshire, using different methods, reported broadly positive outcomes (Alexander 2003, 2005a, 2005b). Subsequently, Lefstein and Snell provided an external, ethnographically forensic perspective on the TTDI (London) project, and investigated the resulting changes in classroom talk and the dilemmas these provoked in one of that project's schools (Lefstein 2010; Lefstein and Snell 2011, 2014). Three years on from the initial intervention they found clear evidence of a well-embedded signature pedagogy traceable to TTDI (Lefstein and Snell 2014, 31), and confirmed Alexander's (2005b) findings on the specific ways that both teacher and pupil talk had changed:

... more use of 'questions which probed and/or encouraged analysis and speculation'; greater use of 'paired talk to prepare for whole class discussion'; a 'more flexible mix of different kinds of talk – recitation, exposition, discussion, dialogue'; and 'an increase in pupil contributions of an expository, explanatory, justificatory or speculative kind.' These changes were apparent in our 2008–9 observations. (Lefstein and Snell 2014, 36).

Once this version of dialogic teaching is located within the extended family of talk reform approaches with which it has most in common, we find further evidence that dialogue makes a difference. Thus, Hattie's synthesis of 800 meta-analyses relating to student attainment shows that the biggest effect sizes available by the mid 2000s related to teaching strategies in which the quality of talk is paramount: reciprocal teaching, feedback and student self-verbalisation, for example (Hattie 2009). Extending the nexus, Black, Harlen and others find essential affinities between dialogic teaching and assessment for learning (Black et al. 2003; Harlen 2014). Meanwhile, Galton compares the pedagogical talk of mainstream teachers and arts practitioners. For the latter, expressive and evaluative talk are as prominent as the other kinds of 'everyday talk' – transactional, expository, interrogatory, evaluative – listed in repertoire 2 above, and one only has to listen for a few moments to actors or chamber musicians rehearsing to understand how cumulation depends on collectivity and reciprocity, and how all three are essential to the melding of acquired artistic skill, creative impulse and divergent viewpoints for the furtherance of a creative activity's purpose. But what is particularly striking about Galton's observations and interviews is how students felt about the talk encouraged by the visiting arts practitioners. Contrasting it with fear of boredom and making mistakes in mathematics lessons, they found these arts sessions engaging and stimulating, and the talk far less likely to be imbued with negative comment or dominated by procedural niceties and time-watching. Meanwhile Galton observed patterns of interaction that could have come straight from a dialogic teaching manual:

Compared to teachers, creative practitioners ... gave pupils more time to think when planning and designing activities ... extended questioning sequences so that classroom discourse was dialogic rather than ... the more usual 'cued elicitations' ... offered more precise feedback ... tended to extend rather than change pupils' initial ideas ... built appropriate scaffolding into the task instead of using teacher dominated approaches such as *guided discovery* ... were more consistent in their management of learning and behaviour. (Galton 2008, x)

In the context of an intervention such as the one under discussion, where efficacy in the official evaluation is judged solely by student scores in standardised tests, the motivational power of this kind of talk can all too easily be missed. It should not be. Engagement, after all, is a prerequisite for learning, and engagement in contexts of social disadvantage is not always easily won. Nor should we puritanically sniff at what Galton's young interviewees called the 'fun' of creative activities, for why shouldn't learning be enjoyable?

In 2011, many workers in the field of classroom talk reform met in Pittsburgh USA under the auspices of the American Educational Research Association (AERA). The resulting research compendium opened by summarising the evidence as follows:

Students who had experienced this kind of structured dialogic teaching *performed better* on standardised tests (i.e. tests that the investigators did not control) than similar students who did not have discussion experience. The data also showed that some students *retained* their learned knowledge for two or three years. More surprising, in some cases students *even transferred* their academic advantage to a different domain (e.g. from science instruction to an English literature exam). (Resnick, Asterhan, and Clarke 2015, 1, authors' italics)

Later, having reviewed the conference papers Lauren Resnick confirms the editors' initial premise that dialogic teaching 'can produce learning gains that go well beyond the topics actually discussed' and adds that it is able to generate not only this extent of learning transfer but also 'a more general ability to learn, an ability that we often attribute to intelligence' (Resnick 2015, 441).

This seemingly bold claim, that talk not only advances learning as it is defined for educational purposes but also ‘socialises’ the intelligence that is held to condition such learning (hence, the AERA book’s telling title ‘Socialising Intelligence ...’) in fact elides with current thinking about the malleability of human intelligence, the capacity of schooling to raise IQ scores (Berliner and Biddle 1995, 50), and evidence that in modern societies measured intelligence appears to be growing (Flynn 1987; and Neisser 1998, cited in Resnick 2015, 443–444).

The next step is to identify what it is about dialogic teaching that makes the difference. Resnick suggests three possible explanations: (i) through dialogue students learn specific, powerful and transferable skills; (ii) by being treated as thinkers and reasoners they ‘develop a more positive view of their own intellectual learning competence, which leads them to engage more actively and successfully in future learning activities’; (iii) they are socialised into ‘a culture of argumentation ... [which] privileges standards of reasoning over ‘correct’ forms of expression, providing students with a safe space to hone their ideas.’ (Resnick 2015, 444). I shall revisit these hypotheses later.

The intervention as professional development

We turn now to the professional development programme through which dialogic teaching as outlined above was to be fostered.

Stages and schools

The intervention had three stages: pilot (2014–2015), trial (2015–2016), and follow-up (2016–2017).

The pilot took place in 10 of the London primary schools¹² that had been involved in the earlier project reported in Alexander 2005b and revisited by Lefstein and Snell (2014). It was hoped that returning to these schools would facilitate a fruitful conversation with the teachers involved about the benefits and challenges of the chosen strategies.

The trial took place in 78 schools in the cities of Birmingham, Bradford and Leeds. The target figure was 80; 78 agreed to participate; being in federated pairs four schools were treated as two, so for analytical purposes the number was 76. By the start of the trial this had reduced to 72, but – questionably – although the pre-intervention withdrawals played no part in the programme they were included in the RCT’s data analysis on the basis of ‘intention to treat’.

The project’s third stage enabled the in-house evaluation to be completed and the project team to honour its commitment to those schools which participated in the project hoping for developmental and pedagogical benefits but found themselves in the control group, a situation that often prompts post-randomisation dropout. To reduce this risk we offered these schools the opportunity to participate in a repeat of the programme as trialled, though without the same intensity of monitoring and support or, of course, the tests.

The Education Endowment Foundation, which commissioned and funded the project, uses UK government funding to select and trial promising initiatives for closing the attainment gap between disadvantaged children and the rest. Accordingly, all schools in both pilot and trial met the standard poverty criterion of a high proportion of children (at least 20 per cent) eligible for free school meals (FSM). Further, English was the additional language (EAL) rather than the mother tongue of about half the students involved. Participating

schools were also required to be large enough for there to be at least two parallel classes in each year. The students were all in their penultimate year of primary education, that is aged 9–10 or Y5 (year 5 or US 4th grade).

Strategies

The intervention's professional development programme combined seven strategies:

- Induction and training
- Mentoring
- Video/audio
- Guided planning, target-setting and review
- Whole school involvement
- Materials and professional study
- Monitoring and support from the project team.

Strategy 1: induction and training. In July 2015, following randomisation by the evaluation team, the mentors, Y5 teachers and headteachers in each intervention school were invited to attend full-day induction sessions in Birmingham (for Birmingham schools) or Leeds (for schools in Leeds and the neighbouring city of Bradford). Mornings were devoted to introducing dialogic teaching, afternoons to explaining the professional development programme. In September 2015, following the summer holiday/recess and immediately before the start of the scheduled programme, mentors returned for a full-day training and simulation session. Between these two sessions participants were asked to read and familiarise themselves with the project print material of which all were given copies at the induction, and to ensure that they knew how to operate the digital video camera and audio voice recorder with which, on the same occasion, each school was provided.

At the end of each of the intervention's two phases, that is in December 2015 and May 2016, participants met again in Birmingham and Leeds to share experiences, review progress and hear about the development team's in-house evaluation.

Strategy 2: mentoring. The mentors – one in each school – were expected to be experienced teachers though not necessarily members of the school's senior management team. Indeed, a hierarchical or inspectorial view of mentoring was strongly discouraged and schools were asked instead to foster a relationship of peers embarking on a shared journey in which professional learning is mutual and discussion is open, advisory and non-judgemental. Such a relationship is particularly important in the arena of classroom talk, which is at the heart of every teacher's professional activity yet also raises questions which are as much personal as professional and need to be handled with sensitivity to the feelings of those involved.

In relation to each of the programme's fortnightly cycles, described at 'phases and cycles' below, mentors worked with their teacher mentees to foreground within the planned lessons those aspects of teacher and/or student talk on which the handbook (Alexander 2017c) invited them to focus. At the end of each cycle they jointly reviewed video examples of the resulting practice in accordance with protocols provided in a linked review booklet (Alexander 2017d), noting strengths and areas for further development. Typically, this meeting then moved from review to planning for the next cycle.

It was for participants to choose whether the mentoring sessions should be one-to-one or joint – that is, with all of a school's project teachers coming together to share experiences and view and discuss each other's video extracts. Most opted for the latter arrangement and testified to its value.

Strategy 3: video/audio. Video and audio are not only ideal for capturing classroom interaction as both sound and behaviour, for talk is signalled by body language and gesture as well as by what we say and hear. They are also powerful tools for professional self-evaluation and development. In this project, video was used during the induction and training, but its principal purpose was to provide the material upon which teachers and mentors would jointly work. Recordings fixed teachers' entering pedagogy as baselines for tracking and assessing their progress; and when subjected to close analysis they enabled teacher and mentor to identify aspects of classroom talk on which it might be beneficial and in some instances necessary to work. Video fulfilled the secondary purpose of enabling the development team to assess the intervention's progress and impact, adding matched recordings from control schools.

As one who had used video and audio recording since the 1980s both for studying talk and supporting professional development (Alexander 1988, 2001, 2003, 2005a, 2005b; Alexander and Willcocks 1995), I had noticed teachers' tendency when viewing classroom video clips to concentrate more frequently and critically on the actors' observable behaviours than on the words uttered or meanings exchanged. In order to alert teachers themselves to this viewer bias as well as to counter it, the training sessions worked in turn on three versions of the same lesson extract. First an audio recording directed attention to the form and meaning of talk, and to these alone. Then a transcript allowed more detailed study in which specific speech exchanges, moves or acts could be revisited as many times as necessary. Finally a video clip enabled participants to observe the interplay of talk's linguistic and paralinguistic features and place the talk heard, read and now viewed in its full pedagogical context.

Incidentally, though these days digitalised video is *de rigueur*, the power of audio alone to concentrate attention on the *language* of interaction should be emphasised, and of course it is less intrusive. Re-reading James Britton's pioneering discussion of lesson transcripts (Britton 1969) gives one a sense of the investigative wonderment unleashed by the invention of reel-to-reel audiotape.

[For the technically minded, each school was given one Panasonic HC-W570EB-K Full HD Camcorder with Twin Camera, one Olympus VN-732PC 4 Gb Digital Voice recorder, one Hama Star 61 Tripod, 1 Transcend 64 Gb Premium SDXC Class 10 Memory card, a camera case and batteries. The induction day included a video training session and participants were able to access online a specially prepared training DVD].

Strategy 4: guided planning, target-setting and review. Each cycle in the post-training programme began with planning and target-setting and ended with review, and the video/audio recordings contributed to one or both of these. What was being planned for each cycle was not a deviation from the intended curriculum but a sharper and more self-conscious focus on the part within it that oral pedagogy might play. This required teachers to audit and map the talk in their classrooms, single out those aspects on which, within their planned lessons and the required focus of each of the intervention's cycles, they should work, and

then do so in a systematic way. Baseline sessions in September initiated this process, but development came from repeating the planning/review sequence rather than waiting until the end of the programme to assess progress. Hence, the 11 planning/review/refocusing cycles discussed below.

Strategy 5: whole school involvement. Individual development and innovation are most successful when they are supported by the school's leadership and embedded in its everyday professional discourse. Although the trial was confined to teachers and students in Year 5, schools were encouraged to take ownership of both the pedagogy and the development strategy and to explore their application across the school.

Strategy 6: materials and professional study. In this project it was axiomatic that teachers learn from examining the practice of both themselves and others, and that effective professional development requires understanding of the ideas and evidence on which the objectives of the development are based. To this end, all participants were provided with a suite of print materials and online access to these and video material. The print materials included:

- A 68-page handbook (Alexander 2017c) which sets out the project's aims and processes and details the intervention programme in full and week by week, supporting this with brief accounts of dialogic teaching and mentoring, and exemplificatory transcripts of lesson extracts.
- A booklet containing planning/review forms for each cycle, with appropriate prompts (Alexander 2017d).
- Alexander's *Towards Dialogic Teaching: rethinking classroom talk*, this being the work on which the intervention was principally based.
- The *Talk Science Primer* of Michaels and O'Connor (2012) which lists and provides the rationale for the nine 'extending' moves listed in Repertoire 6 above.
- A laminated sheet entitled *Dialogic Teaching Repertoires* which reduces dialogic teaching, for easy and daily reference, to its barest essentials.

To these was added, to deepen mentors' understanding and sensitise them to the uses of video, Lefstein and Snell's book of case studies with its linked website with lesson video clips (Lefstein and Snell 2014). The project's own on-line resources included the handbook and planning/review forms together with two DVDs created by the project team: *Dialogic Teaching* and *Video Recording in Classrooms*. All these materials were distributed at the July 2015 induction sessions in the expectation that they would be studied before the programme started in September.

Strategy 7: monitoring and support. During each of the programme's two phases every intervention school was visited at least once by a member or members of the project team. The visits entailed meetings with the Y5 teachers, their mentors and, where possible, school heads. Progress was reviewed, video clips were discussed, planning/review booklets were examined, and problems needing resolution were identified and, ideally, resolved. The problems were typically of a kind that teachers themselves were happy to articulate, but in some cases, for example where the relationship between mentor and teachers was not working or the planning/review sessions were being skimped, it was necessary for the

project team to intervene. If resolution was not possible in one visit, the project's school liaison officer would return to the school and then maintain contact by phone and/or email for as long as was necessary.

Alongside monitoring and trouble-shooting, the visits fulfilled the more formal purpose of acquiring interview data for the project's in-house evaluation. This aspect is dealt with more fully later in this paper.

Phases and cycles

The programme comprised eleven planning/review/refocussing cycles. Each lasted for two weeks except the one-week opening and closing cycles. The programme was spread across two school terms or phases. Phase 1 (cycles 1–6) ran from 21 September to 18 December 2015, Phase 2 (cycles 7–11) from 4 January to 18 March 2016. The entire programme occupied 20 weeks.

In Phase 1, entitled *Expanding Repertoires*, teachers and mentors:

- Made video/audio lesson recordings to use as baselines for future development and comparison, and scheduled dates for subsequent recording and mentoring sessions (cycle 1).
- Discussed and agreed with students the conditions and norms for talk on which the success of this short but intensive programme would partly depend (cycle 2).
- Mapped and began to refine talk repertoires for whole class teaching, focusing first on the teacher's questions, instructions and explanations (cycle 3), then on moves to extend the contributions of the student (cycle 4).
- Shifted the focus to repertoires for small group and one-to-one discussion, both teacher-led and student-led (cycle 5), consolidating the norms and generic talk repertoires from the previous cycles.
- With the building blocks in place, and mindful that successful learning and teaching depend on the student's engagement, identified and worked on those kinds of talk which are most likely to secure this engagement and make the talk truly inclusive (cycle 6).
- Pulled together material from the programme so far in preparation for the mentors' plenary (cycle 6).

In Phase 2, entitled *Advancing Dialogue*, teachers and mentors:

- Initiated and implemented an intensive six-week programme of teaching, in the National Curriculum core subjects English, mathematics and science, which consciously applied and sustained the full range of talk repertoires opened up during Phase 1 (cycles 7, 8 and 9)
- Explored dialogic teaching across the wider curriculum by focusing in addition on National Curriculum non-core foundation subjects (cycle 10)
- Pulled together evidence from phases 1 and 2 for the final plenary for all project participants (cycle 11).

Directed and responsive foci

The piloted version of this programme had been relatively flexible, presenting in some detail the properties of dialogic teaching to be aimed for while leaving teachers and mentors free

to devise their own routes to these. This produced variation between teachers, as to both focus and quality, which was greater than could be accommodated by the methodology of the randomised control trial, which required a high degree of implementation fidelity. The loose framework of the pilot was therefore replaced by the intensive programme, outlined above, in which the focus and tasks for each cycle were precisely specified. At the same time, there needed to be room for variation in each teacher's circumstances, capacities and needs, so a distinction was made between the programme's 'directed' and 'responsive' foci. To quote from the project handbook:

The *directed focus* is what we ask all schools to follow during the cycle in question in order to ensure consistency and progression. Within each cycle lasting a fortnight (all but two of them), planning/target setting should be done early in the first week and review/refocussing at the end of the second. This will ensure that each such cycle includes at least 6-7 days for teaching to the cycle's aims and targets. The *responsive focus* is an aspect or aspects of classroom talk to which individual teachers and mentors decide they would like to attend in order to extend or add to the directed focus. This will reflect the unique circumstances and needs of each teacher and his/her class. As a reminder of the need for the responsive focus and the discretion it offers, a blank 'responsive focus' column is included in the planning/review forms. (Alexander 2017c, 24)

Allowing for a responsive focus within each cycle alongside what the project itself directed seemed an appropriate compromise between the fidelity demanded by the trial and the flexibility necessary in all teaching. However, the shift from the latitude that teachers were given in the pilot to the closely prescribed framework of the trial was of such an order of magnitude that the team feared that teachers might find it excessive. It was with some relief that when we raised this at the plenary session at the end of phase 1, mentors affirmed that they and their teaching colleagues not only found the framework helpful rather than overly prescriptive but, given the intervention's complexity and brevity, anything less would have left some of them floundering. But they also welcomed the deviation that the responsive focus allowed.

Planning and review protocols

For each cycle mentors were asked to refer to and complete with their mentees a planning/review form either in longhand in a special booklet or electronically on forms downloadable from the project website (Alexander 2017d). The forms reminded teachers and mentors of each cycle's directed focus and provided prompts for planning and review. The planning prompts included reference to relevant sections and pages in the print materials. The planning and review prompts appeared in the first column while the second and third columns, headed 'Directed focus – plan/review' and 'Responsive focus – issues to take forward', were mostly kept clear for the mentor's comments.

Part 2 – Evaluation

Evaluation design

*Randomised control trial*¹³

The trial organised independently by a separate team at another university (Sheffield Hallam) used a three-level clustered RCT design (pupils within classes with schools), with randomisation at school level and the classes divided equally into intervention and control groups. From the target number of 80 schools, as noted earlier, 78 were recruited (counted

as 76 because two pairs of schools were federated) and were included in the trial on the basis of intention to treat, though the actual number participating in the intervention, after pre-intervention withdrawals, was 72. To be included in the trial, schools needed to have at least two parallel Y5 classes (US 4th grade) and at least 25 per cent of their students eligible for free school midday meals (FSM). The actual average FSM proportion, overall, was 35 per cent.

The evaluation team used as outcome measures GL Assessment Progress Tests in English, mathematics and science.¹⁴ Students were randomised at classroom level to participate in one of these so that at the assessment point, in late May 2016, one third of each class took each test. The evaluation team argued that this reduced the testing burden on pupils and teachers without significantly reducing the statistical power of the analysis.

Tests were scored by GL Assessment, the company that published the tests. The scoring was blind and for each measure the raw, unstandardised score was used in the analysis. The primary analysis of the intervention's impact, based on intention to treat (38 intervention group schools) rather than actual participation in the intervention (31 schools), was a multilevel regression model of each outcome measure, using as covariates (i) membership of the intervention group, (ii) KS1 point score¹⁵ and (iii) FSM eligibility.

The primary analysis included subgroup analyses for FSM students to assess whether the intervention had differential impact on these students, who being economically the most disadvantaged are the main target group for all EEF projects and trials.

Controversially, but in accordance with EEF RCT practice, there were no pre-tests. EEF argues that suitably rigorous sampling and randomisation enables intervention effect sizes to be reliably calculated from outcome measures alone.

School, student and teacher numbers at each stage of the trial process are shown in Table 2. The school and student attrition recorded above allowed EEF to give it a security rating of 3 out of 5. EEF deemed this sufficient for concluding that attainment gains made by intervention group students over their control group peers were due to the intervention rather than other factors.

Table 2. School, student and teacher numbers in the dialogic teaching trial.

Recruitment	Schools approached: 80 Schools agreeing to participate: 78 Counted in the RCT as: 76 (two pairs of schools were federated) Pre-intervention withdrawals: 6 schools Actual participation: 72 schools
Allocation, intention to treat	78 schools counted as 76 Intervention group: 38 schools, 2492 students, 118 teachers ^a Control group: 38 schools, 2466 students, 90 teachers
Test data collected	Intervention group: 31 schools, 2097 students ^b Control group: 38 schools, 2466 students
Test data analysed	Intervention group: 31 schools, 1832 students ^c Control group: 31 schools, 2080 students

^a80 class teachers, 38 mentors.

^bIn addition to the 6 pre-intervention school withdrawals, one school failed to return test data. 395 students were therefore lost to the follow-up.

^cComplete datasets were unavailable for 265 students in the intervention group and 386 in the control group either because of student absence or opt-out or because tested students could not be matched with available data on KS1 attainment or FSM eligibility.

External process evaluation

Alongside the trial the Sheffield team organised a qualitative implementation and process evaluation. This included a postal survey intended to include all intervention group teachers, mentors and headteachers but which yielded a very low return rate, together with telephone interviews with six teachers, eight mentors and three headteachers, and classroom observation and interviews in three schools.¹⁶

In-house formative evaluation

The in-house evaluation by the York-based development team, which was entirely separate from the trial team's process evaluation, had two strands: an interview programme undertaken in intervention schools only; and a comparative analysis of videotaped lessons from both intervention and control schools.

Interviews. Two sets of interviews were conducted with teachers, mentors and headteachers in each intervention school. The first took place during the programme's phase 1 (autumn 2015), the second towards the end of phase 2 (spring 2016). The focus in each case was participants' adherence to and divergence from the specified programme, challenges encountered during its implementation, and its perceived impact on teaching, learning, student engagement and classroom talk. The interviewers, who were all members of the York team, worked to an agreed schedule of questions. Answers were recorded on a proforma for later analysis using the NVivo software for qualitative data. To the interview programme was added scrutiny of the cycle-specific planning/review forms completed by the mentors for each of the teachers they were supporting, to allow deeper insights into fidelity, feasibility and utility.

*Video data*¹⁷. In order to assess the pedagogic impact of the intervention it was necessary to videotape lessons in sub-samples of both the intervention and control groups, and to do so twice so as to track development and progress over time. Video recordings of English, mathematics and science lessons were made (i) early in phase one (week beginning 21 September 2015) to provide a baseline, and (ii) towards the end of phase 2 (fortnight beginning 22 February 2016).

15 teachers from the intervention group and 11 from the control group agreed to be video-recorded. The intervention group teachers were self-selected in response to a request for volunteers at the July 2015 induction sessions. Self-selection was the only realistic possibility: to impose selection on top of the other demands of the project could have been counterproductive for retention. The control group teachers were selected on the basis of school-school matching.

Each teacher was recorded twice, in phase 1 and again in phase 2, yielding a theoretical total of 156 lessons (2 English, 2 mathematics and 2 science in each case). In fact, because not all of the designated teachers taught science, the total number of lessons recorded was 134 (67 in each phase). The resulting recordings were subjected to both quantitative and qualitative analysis.

For the quantitative analysis, some of the key verbal indicators of typical classroom talk, both traditional and dialogic¹⁸, became the basis for a coding system that was piloted in the London schools before being finalised and applied to the trial stage video data. Coders were trained and checked to maximise coding consistency. The coding system was uploaded into

the Noldus Observer XT 12.5 software in order to generate quantitative data from the coded exchanges. These were then statistically analysed using SPSS. The analysis was undertaken twice for the purpose of cross-validation, first internally at the University of York, then externally by Kirkdale Geometrics.¹⁹

The quantitative analysis undertaken to date covers:

Student and teacher talk (aspects of repertoires 3, 4 and 5)

- Change in the ratio of teacher talk to student talk over time, and intervention/control comparisons.
- Change in the ratio of recitation to discussion/dialogue over time, and intervention/control comparisons.

Teacher talk (aspects of repertoire 6)

- Intervention/control differences in teacher extending moves.
- Development in teacher extending moves over time (from phase 1 to phase 2) and intervention/control comparisons.
- Teacher extending move differences between English, mathematics and science.

Student talk (aspects of repertoires 3, 4 and 5)

- Intervention/control differences in the ratio of brief to extended student contributions.
- Change in the ratio of brief to extended student contributions over time (from phase 1 to phase 2) and intervention/control comparisons.
- Brief/extended ratio differences between English, mathematics and science.
- Frequency of sub-types of extended student contributions, derived from the project's categories of learning talk (repertoire 3).

Evaluation outcomes

Although the headline findings of the RCT may be of greater public interest – and after the release of its report they did indeed provoke press attention²⁰ – chronology and logic require that the findings from the in-house evaluation, which tracked the intervention's progress towards the point at which student attainment was tested, be presented first.

In-house evaluation: interview data²¹

Programme impact. Overall, participating teachers claimed direct positive gains from the programme for classroom talk, student engagement and student learning, and for their own professional understanding and skill. Specifically:

- Norms for student talk were fairly quickly established and embedded.
- Teachers learned to extend their basic talk repertoires and their skill in using them.
- Specific teaching strategies such as questioning, discussion and feedback became more systematic and effective.
- Exchanges were lengthened, sustained and deepened.
- Students' preparedness to listen to each other improved.
- Interaction became more inclusive, with fewer students isolated, silent or reluctant to participate, while previously dominant students became less inclined or able to monopolise the talk and teachers' attention.

- With an increased emphasis on a supportive, reciprocal talk culture, students gained in confidence and became more patient and better attuned to each others' situations and keen to provide mutual support in both talking and learning.
- There were also subject-specific gains. In English, teachers reported improved student vocabulary, better discussion, and evidence of transfer of verbal gains from oral to written work. In mathematics, students became more adept at explaining the reasoning behind their solutions and thus providing teachers with a firm basis for both feedback and further extension. In science, the democratisation of questioning that is an aspiration of dialogic teaching fed into a more genuinely scientific stance in students' investigations and discussions.

Programme implementation. Overall, checks on mentors' entries in the planning/review forms confirmed interview claims that the programme was implemented with a reasonable degree of fidelity and that it was both useful and feasible. Specifically:

- By the end of the intervention, every school had completed all eleven cycles of the programme.
- 57 per cent of schools reported that they had followed the programme as specified in the handbook while 43 per cent had made modifications, though still within the specified framework.
- The main challenges faced were: insufficient time for teachers and mentors to plan and review (33 per cent of schools in both phases 1 and 2); national curriculum and assessment changes; staff changes (including among those immediately involved); pupil changes (high student turnover affected many project schools); and unanticipated events such as staff illness, especially in phase 2.
- However, most of the challenges that teachers faced were intrinsic to school life rather than generated by the project. No innovation would have escaped them.
- For the critical role of mentor, which requires time for preparation and follow up as well as face-to-face meetings, time pressures were less of an issue for those mentors who held senior positions allowing administrative release.²²

The more limited process evaluation undertaken by the external team in conjunction with the RCT arrived at similar conclusions to those above, so it is not necessary to list them separately. However, one frequently voiced opinion from the external team's survey and interviews should be noted here. When interviewed before the results of the trial were announced, teachers believed that the intervention was too short to achieve a discernible impact on student learning – even though they also asserted that its positive effect on students' engagement and patterns of talk, prerequisites for the hoped-for learning gains, was speedily apparent – and that a period of at least a full school year would have been more effective. In the event, teachers' pessimism about learning outcomes was misplaced, for the intervention's impact on student test scores was no less discernible than on student engagement and talk, as we shall see below. But the point is well made, and the project team itself had already registered it with EEF, the funding body whose evaluation paradigm had been responsible for the intervention's brevity. The follow-up project is likely to extend the intervention from 20 weeks to a full school year.

*In-house evaluation: video data*²³

Comparison of coded talk acts and exchanges in intervention and control classrooms showed significant differences emerging between the two groups over the two terms of the intervention. These differences were striking in both teacher and student talk. For example (but bearing in mind that at the time of writing the video analysis is incomplete):

Closed and open teacher questions. In all three core subjects, the ratio of closed to open teacher questions was fairly evenly balanced in phase 1 but by phase 2 intervention teachers were making greater use of open questions than their control group peers. The argument here is that while closed or ‘what?’ questions require largely pre-ordained responses dependent on recall or at best instant calculation, open and ideally authentic questions launched by ‘how?’ or ‘why?’ or ‘what if?’ encourage reasoning, speculation and more active cognitive and indeed social engagement – provided, of course, that the teacher allows appropriate wait/thinking time (Alexander et al. 2017, Figure 1 and Tables A and B).

Teacher extending moves. Intervention teachers were trained to deploy a variety of moves to probe, extend and follow up student contributions (repertoire 6) on the principle that these would both increase students’ interest and engagement and enhance their cognitive gains. Differences between the two groups in respect of these were most marked in mathematics and science, where by phase 2 the intervention teachers were making significantly greater use of wait/thinking time, revoicing, rephrasing, seeking evidence of reasoning, challenging, requesting justification and so on (Alexander et al. 2017, Figures 2a/2b and Tables A and B).

Balance of recitation and discussion/dialogue. In English and mathematics, comparable ratios of recitation to discussion and dialogue in the intervention and control groups were transformed into significant differences by phase 2, with intervention teachers making much greater use of discussion and dialogue. In this matter, science was again somewhat different in that in phase 1 the intervention group was already making greater use of discussion and dialogue than the control group. This lead was sustained into phase 2 and increased as the intervention progressed (Alexander et al. 2017, Figure 4 and Table C).

Balance of brief and extended student contributions. In English and mathematics, the ratio of brief to extended student contributions in phase 1 was the same in intervention and control classrooms. By phase 2, there were statistically significant differences between the groups in respect of an increase in extended student contributions and a decrease in brief contributions. In science, the intervention group started the programme with a higher ratio of extended to brief student contributions than the control group (Given that this happened after the induction and training it may suggest that the programme’s messages in this regard were more readily implemented in science than the other two subjects, or even that primary science teaching is more instinctively dialogic). This difference was sustained into phase 2 (Alexander et al. 2017, Figure 3 and Tables A and B).

The repertoire of student talk. As emphasised earlier, the present version of dialogic teaching attends as closely to the talk of the teacher as to that of the student, because it is through the teacher’s talk that the student’s talk is either confined within the tightly controlled boundaries of recitation or encouraged through discussion and dialogue to

enlarge its discursive and semantic repertoire and hence its cognitive power. Hence the focus above on the balance of closed and open questions, recitation and dialogue, and on brief and extended student contributions. For while dialogic teaching as conceived here accepts the need in certain circumstances for closed questions, recitation and brief student contributions, it also affirms that unless the quantity and quality of student talk are extended well beyond these traditional patterns of exchange into a more extensive interactive repertoire, the full communicative and cognitive potential of classroom talk will remain largely unrealised. In the end, therefore, it is the student's talk that matters most, and it is to the teacher's agency in securing the enhancement of student talk that dialogic teaching is directed.

To judge student talk merely by the length of utterances, as in the brief/extended analysis referred to above, is useful only as a preliminary or general indicator of quality. What matters is the form of student talk that opportunities for its temporal extension allow, for extended talk may be – in terms of the most demanding of the five criteria of successful dialogic teaching – cumulative, or it may be merely circular, and this is a particular risk in classroom discussion (cf. the observation of Michaels and O'Connor (2015, 358) that 'the simple deployment of talk moves does not ensure coherence in classroom discussions or robust student learning').

Here, the 11 categories of learning talk in repertoire 3 (narrate, explain, speculate, imagine, explore, analyse, evaluate, question, justify, discuss, argue) provided the necessary analytical indicators. These were modified for coding purposes as 12 sub-types of extended student contributions which also included student responses to some of the key extending moves in repertoire 6. The modified coding categories for student learning talk were: expand/add, connect, explain/analyse, rephrase, narrate, evaluate, argue, justify, speculate, challenge, imagine, shift position. These were applied to video transcript samples from both the intervention and the control groups in weeks 1 and 18–19.

The differences by the latter stage of the intervention were striking. By then, intervention group students had become markedly more expansive in their contributions and exhibited higher levels of explanation, analysis, argumentation, challenge and justification. Their talk, then, was without doubt more dialogic than that of their control group peers. Though there were between-subject differences, the overall pattern of intervention/control contrast obtained across all three subjects tested for the RCT (Alexander et al. 2017, Table D).

Randomised control trial

Thus, the in-house evaluation undertaken by the project team demonstrated that the professional development programme, though brief, produced changes in teachers' thinking and practice which led to the widening and intensification of the students' own talk repertoires, and the Sheffield team's external process evaluation confirmed these trends, though without the benefit of the video data which would have enabled it to compare talk in the intervention and control groups and to do so systematically.

To those who believe in the value of dialogue as an educational end in itself, as considered earlier in this paper, this finding could have been sufficient, but EEF exists to trial and disseminate interventions that reduce the attainment gap among disadvantaged children, so for them the test scores mattered above all else. The results of these as presented by the independent evaluation team are shown in Table 3.

Table 3. Summary of impact of dialogic teaching intervention on the RCT's primary outcomes.

Subject	Group	Effect size	Months' progress after 20 weeks
Mathematics	Treatment vs. control	0.09	1
	Treatment vs. control (FSM only)	0.16	2
Science	Treatment vs. control	0.12	2
	Treatment vs. control (FSM only)	0.11	2
English	Treatment vs. control	0.15	2
	Treatment vs. control (FSM only)	0.12	2

Jay et al. (2017, 5)

The evaluation's overall conclusions, which added findings from the process evaluation to the effect sizes, recalculated by EEF as months of progress, were as follows (here quoted verbatim from the Sheffield team's report):

- (1) Children in Dialogic Teaching Schools made two additional months' progress in English and science, and one additional month's progress in mathematics, compared to children in control schools, on average.
- (2) Children eligible for free school meals (FSM) made two additional months' progress in English, science and mathematics compared to FSM children in control schools.
- (3) The intervention was highly regarded by headteachers, mentors and teachers, who thought that the Dialogic Teaching approach had positive effects on pupil confidence and engagement.
- (4) The majority of participating teachers felt that it would take longer than two terms to fully embed a Dialogic Teaching approach in their classrooms. It could, therefore, be valuable to test the impact of the intervention over a longer period.
- (5) The intervention requires teachers to change classroom talk across the curriculum, supported by training, handbooks, video and regular review meetings with mentors. Future research could aim to differentiate the effects of these different elements. (Jay et al. 2017, 4)

The cost of the intervention was calculated as GBP52 per student per year, which made this intervention one of the cheapest trialled by EEF. About its main findings and conclusions the EEF evaluation report also observed:

- At 20 weeks / two terms the intervention was too short to achieve its maximum impact, and a longer intervention could well have produced even larger differences in attainment between the two groups.
- Given that it can take two or three years for a complex intervention to be fully implemented the RCT was undertaken too soon.
- The findings may have been affected by unobserved differences between schools that withdrew from the trial and those that remained.
- The intervention was unusual among those trialled for EEF in that it sought to improve attainment across the curriculum, and succeeded in doing so. Most EEF projects deal with one subject only. (Adapted from Jay et al. 2017, 44–46)

And added this significant admission:

Limitations to the methodological approach taken here mean that there may be positive effects of the intervention that could not be detected, or that effects observed may be underestimated. (Jay et al. 2017, 44)

Converting effect sizes to months of progress in order to give teachers and policy makers a more meaningful basis for their decisions is of course far from straightforward. EEF explains its conversion procedure, assumptions and caveats in a technical appendix (Higgins, Kokatsaki, and Coe 2012) which remains on the Foundation's website even though I understand that EEF no longer accepts its rigid low/medium/high banding of effect sizes because these have no regard to an intervention's character or – crucially – its length (Additional progress of two months after 20 weeks is clearly more noteworthy than it would be after 60 weeks). EEF's decision to headline and give considerable media prominence to the two-month gain from our intervention, signalling confidence in its significance, is discussed further below.

Part 3 – Discussion

Intervention and outcomes

To summarise. The intervention had two strands, pedagogical and developmental. Grounded in an established approach to dialogic teaching, steered by print materials, in-school mentoring and video/audio analysis, and supported by externally provided training and monitoring, it used a 20-week cyclic programme of planning, target-setting and review to encourage teachers to expand their and their students' repertoires of classroom talk in the direction of dialogue and argumentation. The required changes were argued on evidential and ethical grounds as a necessary basis for increasing engagement, improving learning and enriching education among all children, but especially those from the poorest families, where educational underachievement is at its most marked and intractable.

The version of dialogic teaching underpinning the intervention, while having obvious affinities with several others and incorporating among its six repertoires one developed elsewhere,²⁴ is in three respects distinct.

First, while privileging dialogue narrowly defined, it embeds it within a larger set of talk repertoires which – controversially perhaps, though in my view necessarily – retains for certain circumstances two kinds of talk with which dialogue is usually opposed. The approach, therefore, requires a *professional* dialogue about the resulting interactive options to inform the *pedagogical* dialogue with and between students. This contrasts with approaches that pin their hopes and advocacy on teachers adopting set patterns of talk, a stance that perhaps risks appearing to commend dialogue for others but not for itself.

Second, the approach is one element in a larger pedagogical framework encompassing the handling of lesson structure, space, student organisation, content, time, routines and rules, learning tasks and activities, and assessment (Alexander 2001, 323–335). The character of classroom talk relates to and is contingent upon all of these, so they need as far as possible to be in harmony.²⁵

Third, dialogic teaching as defined and operationalised here is not merely a technical matter. True, it tries to make talk more collective, reciprocal, supportive, cumulative and purposeful and hence more inclusive and efficient, and it draws on psychological, neuroscientific and pedagogical evidence to make its case in this regard. But it does so with larger educational aims in view that foreground dialogue as cultural and civic imperatives, and that propose a dialogic stance on the nature and growth of human knowledge. This is why, though citing Bakhtin has these days become a fashionable claim to educational

provenance, it seems legitimate in this instance to do so, for dialogic teaching as defined here is an epistemology and a habit of mind, not just a way of upgrading talk, and it 'seeks to grasp human behaviour through the use humans make of language.' (Holquist 2002, 15).

Turning to the findings, these can be more succinctly summarised. Interviews conducted in parallel by the development and evaluation teams found teachers highly supportive of the programme's aims and strategies though mindful of its challenges. The development team's analysis of video-recorded lesson episodes from both intervention and control group classrooms showed how talk in the intervention classrooms begin to shift in the intended direction early on, and continued to do so, with considerable divergence in patterns of both teacher and student talk evident by week 19. Similarly, and we believe consequently, after the 20-week programme the randomised control trial found that students in the intervention group were two months ahead of their control group peers in standardised tests of English, mathematics and science, despite the fact that problems such as school attrition and the required brevity of the intervention are thought to have caused its impact to be understated.

Limitations of the randomised control trial

It has been objected that RCTs are not well suited to an arena as complex, idiosyncratic and ephemeral as teaching, and that their claim to represent the 'gold standard' in educational and social research is overstated, misguided and perhaps even imperialist (Berliner 2002; Prideaux 2002; Norman 2003; Sullivan 2011; Ginsburg and Smith 2016; Pogrow 2017). This is not the place to assess such claims, but it is right briefly to note reservations about the RCT paradigm as it was applied in this particular case.²⁶

Firstly, the RCT used checked but nevertheless subjective Y2 teacher assessments as measures of prior attainment, instead of a proper pre-test of the target groups immediately before the Y5 trial. Further undermining the reliability of this procedure, the rapid student turnover that characterises inner-city schools with high proportions of disadvantaged/FSM/EAL students schools meant that the cohort tested in 2016 was appreciably and perhaps unsustainably different from that tested in 2013.

Secondly, the RCT failed to deploy appropriate procedures to address the acknowledged problem of missing data at the school and classroom levels.

Thirdly, the RCT relied on standardised tests in English, mathematics and science as sole measures of the programme's educational outcomes, to the exclusion of measures of student engagement and spoken language which were actually agreed at the outset and would have allowed the programme's impact to have been evaluated in a manner more consistent with its aims. Fully consistent and valid measures would also most certainly have yielded a bigger effect size.

Fourthly, the decision to calculate effect sizes separately by subject, especially in view of the fact that only one third of the students took each test, further weakened the power of the analysis. A combined multivariate analysis would have reduced error and made the effect calculation more secure.

Fifthly, the report failed to comment on the EAL dimension, despite the fact that English was the second language of half of the students and that the intervention group had a larger proportion of EAL pupils than the control group (53 per cent compared with 47 per cent – Jay et al. 2017, 24). Talk reform is doubly challenging in the EAL context, and the weighting

of this measure of disadvantage against the intervention group may have further reduced compromised effect size.

Sixthly, the report also said little or nothing about other obvious differences within the student population such as gender, ethnicity, cultural background and special needs. So, for example, combining several of these variables we note that a significant proportion of the students were girls of Muslim Asian heritage, and while the RCT may have been methodologically unsuited to tracking the qualitative impact of the intervention on sub-groups such as this, the external process evaluation arguably could and should have done. Similarly, 16 per cent of the students in each group (intervention and control) were classified as having special educational needs (SEN) but their progress was not separately investigated. The only sub-group to be separately analysed was FSM pupils. These accounted for 35 per cent of each group and, as noted above, they did better than the intervention group as a whole in the mathematics test. But we are not told why.

Finally, although the report acknowledged that 'limitations to the methodological approach taken here mean that there may be positive effects of the intervention that could not be detected, or that effects observed may be underestimated' (Jay et al. 2017, 44), nothing was done to reduce the risk or address the problem. And since the effects are quantified it would have been helpful if the extent of their possible underestimation could also have been quantified.

And yet, a two-month student attainment advantage from a 4.5 month intervention might be rated pretty impressive; and if in the estimation of the independent evaluation team and EEF – neither of which had an interest in talking up the outcomes – the balance of the RCT's unreliability leans towards underestimation of the effect size rather than exaggeration, then we might as well accept the finding. Further, although we in the development team objected to aspects of the RCT as conceived, conducted and reported, we also embraced the advantage, in terms of its public credibility, of escaping the stricture that RCTs are too frequently compromised by association with the intervention's developer (Ginsburg and Smith 2016, ii). As EEF points out, all its evaluators are entirely independent of the teams developing the interventions they evaluate, and 'are appointed through a competitive tendering process and reviewed for any academic conflict of interest' (Nevill 2016, 2).

Reaching for explanation

The evaluation report (Jay et al. 2017) recommends that in a future project elements of the professional development programme be disaggregated and implemented separately so that their effects can be individually tested. Though we certainly need to know what in this intervention made such a difference in so short a time, this recommendation seems to reflect both misunderstanding of the intervention's essentially holistic character and the extent to which education RCTs have been colonised by the dosage mindset of the drug trial industry from which they seem to have been imported, as if training = 1 dose, training + mentoring = 2 doses, training + mentoring + video = 3 doses, training + mentoring + handbook = 4 doses, and so on. We wonder what would constitute a dialogic teaching overdose.

This brief diversion into educational pharmacology aside, the external evaluation report offered little by way of explanation or diagnosis from the quantitative and qualitative evidence it assembled. Lacking the essential and agreed measure of spoken language it was unable to track the intervention through to the transformation of students' talk that was its

object, and it was left to the development team to plug this critical gap with its own video data. Nor, crucially, could the RCT differentiate the relative impact of the intervention's pedagogical and developmental strands. Referring merely to 'the intervention', this RCT, for all its statistical wizardry, was a remarkably blunt instrument.

However, since the publication of the evaluation report in July 2017, EEF has indicated its interest in considering a further project in a much larger number of schools, and at the time of writing is in discussion with this paper's author and others about how the evaluation of such a project might achieve greater diagnostic and explanatory power than was afforded by the RCT to which we refer here. It is also hoped that a follow-up project can be funded to enable the intervention to last for a full school year, at least.

Frustratingly, in matters of cause and effect and the differential response to a talk-rich intervention from the various student sub-groups mentioned, we can draw little from the official evaluation and must fall back on prior experience and extrapolation from other sources. In doing so we should remember that both the pedagogical and professional strands of this project were designed holistically, and that to attempt to weigh the relative merits of, say, repertoires 5 (questioning) and 6 (extending) is as fruitless as trying to assess the differential impact on effect sizes of the training and mentoring elements of the intervention's professional strand. This is because the repertoires track a logical sequence from everyday talk via options for organising talk to teaching talk and its constituents, and then on again to learning talk. In this case, for example, there can be no extending without prior questioning, and extending itself is in part interrogatory. Each repertoire relates to the others. All are essential.

Further, the very comprehensiveness of the approach as conceived is one of its strengths in practice, because in bearing simultaneously on talk's diverse aspects and actors, it is more likely to generate an interactive culture that is pervasively dialogic than if we were try to transform, say, questioning, extending or feedback alone. And a teacher who is relatively less skilled in handling one repertoire may have compensating skill in another.

The approach is comprehensive in another sense: it aims to have application and purchase across the curriculum. That is why the intervention required teachers to work in the contrasting domains of English, mathematics, science and one other subject of their choice; and it is why, in the EEF trial, learning outcomes were assessed in three curriculum domains when in most EEF trials they are confined to one. In as far as a positive effect was found in all three tested subjects, the case appears to be upheld, as EEF noted in its press release when the evaluation report was published:

The consistent results across subjects suggest that the approach may improve children's overall thinking and learning skills rather than their subject knowledge alone. This is backed up by evidence summarised in the Sutton Trust/EEF Teaching and Learning Toolkit that advises that metacognition approaches – strategies that encourage pupils to plan, monitor and evaluate their learning – are a particularly effective way of improving results.²⁷

Back, then, to the question of cause and effect. Reviewing the various experimental studies brought together in the 2015 AERA collection referred to earlier, Resnick finds convincing and replicated evidence for successful cross-curriculum transfer which supports our own experience, EEF's deduction that dialogic teaching is a generic pedagogy rather than a subject-specific one, and its findings from other EEF projects about the impact of metacognition-directed interventions on learning outcomes.

Resnick then looks more closely at how transfer might work, drawing on the work of Kuhn and Zillmer (2015), Koedinger and Wiese (2015) and Adey and Shayer (1993). For instance:

Adey and Shayer described an English assessment task that asked students to listen to a segment of dialogue, choose one of the characters, describe that character's views, and state how they differed from those of another character in the story. Koedinger and Wiese argue that the process of identifying a character's views is comparable to variable extraction in a science context, if we think of a 'view' as a feature that differs one character from another. Variable extraction, moreover, is a skill of argumentation, because claims are often statements of relationships among variables. Thus, through science discussions, students may have learned a specific skill – variable extraction – that helped them on their English exams. (Resnick 2015, 444–445)

Or, in order not to imply a hierarchy of domain value, we might add *vice versa*. Indeed, in the present project we found argumentation to be a prominent feature of English lessons, especially in students' discussions of the meaning of stories and poems.

From the transfer of specific skills between subjects, Resnick moves to a second explanation for dialogue's success, which she calls 'I can learn'. Once again drawing on studies in the AERA collection, she concludes that dialogic teaching

... can change students' perceptions of themselves as learners, especially when the discussion highlights reasoning and gives students opportunities to explain their ideas. The examples of dialogic teaching ... may work by actively treating students as thinkers and reasoners, thereby modifying the ways in which they engage with content. (Resnick 2015, 446)

Finally, Resnick proposes that dialogue yields a 'culture of argumentation':

Students engage in a process of argumentation that has the potential to go beyond any individual student's power of reasoning. The students challenge one another, call for evidence, change their minds and restate their claims, just as adults do in virtually every discipline of knowledge in the world outside of school. (Resnick 2015, 446)

She adds that what matters is the argumentation as such rather than adherence to the 'rules of argument observed, or at least claimed, by logicians, and that 'the focus is on reasoning and knowledge rather than its forms of expression' (Resnick 2015, 447).

The synergy between these three explanations and key aspects of the present project's dialogic teaching repertoire will be evident. Transfer was already presumed in the project's rationale and design, the RCT confirmed it and the funding agency highlighted it. Resnick's 'specific skills' explanation gives added point to the scope and diversity of our 'learning talk' repertoire. This includes, as we have seen, forms of talk such as explaining, speculating, imagining, analysing, exploring, evaluating, justifying, questioning, discussing and arguing that are equally essential to literary, historical and scientific discussion and enquiry, while the comparison of teachers and creative practitioners by Galton (2008), referred to earlier, shows how these forms are no less fundamental to artistic activity. Similarly, Resnick's 'I can learn' explanation underlines the importance of dialogic teaching's principle of supportiveness, while her 'culture of argumentation' endorses the principles of collectivity and reciprocity.

One further explanation suggests itself, allowing us to dig deeper. Resnick properly concentrates on the transferability of dialogic habits of *student* talking and thinking across curriculum domains. But the 'culture of argumentation' embraces *teachers* no less than students, and 'I can learn' is most likely to convince the students when their teacher believes 'I can teach' (dialogically). After all, it takes two, at least, to dialogue. One reason why recitation persists is that it enables the teacher to retain control of the trajectories of both lesson content

and student behaviour. Exposure to the consequences of publicly and perhaps incorrectly answering ‘test’ questions (Nystrand et al. 1997) is highly risky for students, and some teachers prefer to keep things that way. Hence the well-documented student counter-culture of classroom risk-avoidance (Doyle 1983; Pollard 1985; Galton 2008) and the tactics that students adopt in response to their teachers’ view of ‘communicative competence’ – by, for example, bidding to answer questions in a way that ‘balances the risks of not being noticed against the risks of being ignored as too enthusiastic’ (Edwards 1992, 235).

In contrast, dialogic teaching is predicated on ceding to students a degree of control of both content and behaviour, and it therefore transfers at least some of the risk of public exposure back to the teacher, and not all teachers are happy with either scenario. Moreover, the dialogic teaching principles of collectivity, reciprocity and supportiveness aim in different ways to minimise students’ sense of risk and their fear of its consequences, because only then will they talk as freely as true dialogue requires; while through the various teacher talk repertoires the teacher scaffolds exchanges that ‘guide, prompt, reduce choice and expedite ‘handover’ of concepts and principles’ (Bruner 1978, 2006; Alexander 2001, 528). Similarly, Galton’s highly suggestive descriptions of creative practitioners working with children show – perhaps as much because they are not teachers as because they work in the creative domain – how they instinctively allow wait/thinking time and share control of exchanges: ‘Creative practitioners seem more comfortable with silence ... [and frequently] reverse roles so that the pupil and not the adult asks the questions’ (Galton 2008, 38).

In pursuit of all the conditions and outcomes of dialogically induced learning, the teacher talk repertoires listed earlier – interactive settings, teaching talk, questioning, extending – are no less essential than the range of learning talk to which they are directed, for however far dialogic teaching may democratise pedagogy, empower students and respect their thinking, these things happen only if the teacher makes them happen. The continuing dominance of recitation must constantly remind us of that.

Resnick extrapolates one more important finding from the AERA collection: that most of the studies achieved positive results, in terms of retention and transfer, on the basis of relatively short but intensive spells of dialogue. That, too chimes with the findings of the present study, where an intervention of just 20 weeks accelerated attainment gains by two months, and it prompts Resnick to suggest that teachers might devote one or two lessons a week to ‘well-planned and carefully guided discussions’ (Resnick 2015, 449).

If we have to start somewhere, that may make sense, but I would not wish it to be concluded, and I am sure it was not intended, that 90 minutes a week will be sufficient, or that a brief weekly burst of dialogic teaching can sit happily alongside the monologic teaching that still dominates many students’ experiences of school. For dialogic teaching, at least as conceived here, is a total pedagogy. It certainly allows for ‘traditional’ forms of teaching talk like rote, recitation and exposition, but within rather than in opposition to a broader repertoire in which the overall centre of gravity, across all teaching and the whole curriculum, is shifted decisively towards discussion, argumentation and dialogue. But our account is also underpinned by, and by these means steers students towards, a distinctively dialogic account of knowledge, culture, civic engagement, and education itself.

That prompts a final thought. Within the fast-growing literature on dialogue in teaching and learning a fault line can be detected. On one side dialogic teaching is viewed essentially as technique, as a tool of effective teaching of anything and towards any end. That, I think, is how it has been perceived and evaluated within the ‘what works’ paradigm of EEF, and

against the backdrop of the UK's rising tide of child poverty and inequality anything that closes the attainment gap is of course to be welcomed. For when we combine the headline finding of the project discussed here with Resnick's review of other experimental studies and Nystrand's finding that 'a single authentic question or a single student question significantly increased the measured probability of a subsequent dialogic spell' (Nystrand et al. 2003), we find that the evidence points unerringly to dialogue's sheer efficiency, even (dare I say) in small doses.

But on the other side, and offering a richer and even more persuasive case, we have dialogue as a pedagogy that is not only effective in these terms but is also a fitting way to enact the larger aims for which, at best, education claims to stand. Since these two arguments for dialogic teaching are not mutually exclusive, there is every reason to resist the reductionist corollary of 'what works': that what works in teaching is all that matters and as long as our most disadvantaged children can be helped to improve their literacy and numeracy 'outcomes' the rest of their education can take care of itself. In England, we saw this impoverished and misguided view translated into policy when, in 1998, the government 'disapplied' the national curriculum requirement to teach a broad curriculum encompassing the arts and humanities so that armed with officially prescribed literacy and numeracy strategies (DfEE 1998, 1999) teachers could concentrate without distraction on raising standards in the so-called basics. That policy disenfranchised even as claimed to empower; and illustrating by its denial the potency of transfer it also proved counterproductive for student attainment.²⁸ Let this not be the fate of dialogic teaching.

Notes

1. Details at <https://educationendowmentfoundation.org.uk/our-work/projects/dialogic-teaching/>
2. The development project was based at the University of York and the evaluation at Sheffield Hallam University.
3. Sinclair and Coulthard named the typical three-part recitation exchange IRF (initiation-response-feedback). Following Mehan (1979) and Cazden (2001), most researchers in the field now use IRE. The third move may or may not provide explicit feedback, but it is rarely neutral in intent or interpretation, so 'evaluation' is more exact. In any event, the 'feedback' in much classic recitation may pass judgement but provide little or no useful information.
4. The review's final report is Alexander 2010. For further information and access to the 31 interim reports and 40 briefing papers, see <http://cprtrust.org.uk/cpr/>. The review was succeeded by the Cambridge Primary Review Trust http://cprtrust.org.uk/about_cpr/. This took forward many of the Review's recommendations, commissioned follow-up research and initiated the present project. The aims for education are set out in full at http://cprtrust.org.uk/about_cpr/aims/.
5. **The Individual:** Well-being; Engagement; Empowerment; Autonomy.
Self, Others and the Wider World: Respect and Reciprocity; Interdependence and Sustainability; Local, National and Global Citizenship; Culture and Community.
Learning, Knowing and Doing: Exploring, Knowing, Understanding and Making Sense; Fostering Skill; Exciting the Imagination; Enacting Dialogue.
6. I first advanced this theme in a public lecture in Hong Kong entitled 'Education as Dialogue' (Alexander 2006). It was revised as 'Pedagogy for a runaway world' in Alexander 2008, 121–153).
7. Interestingly, though my approach to dialogic teaching has affinities with Lauren Resnick's idea of 'accountable talk', the paper by Michaels, O'Connor, and Resnick (2008) seems to tie its

important principle of ‘accountability to knowledge’ to a somewhat more canonical account of knowledge than the one adopted here.

8. For further discussion of the challenge of cumulation, see Alexander 2008, 114–119.
9. Though ‘wait time’ is the term famously coined by Rowe (1974, 1986) and since then has been used by many others, teachers in this project found ‘thinking time’ to be more useful since it eliminates any vestiges of Beckettian doubt about what or whom the questioner is waiting for.
10. Michaels and O’Connor call them ‘talk moves’ but in relation to the comprehensive ambition of our dialogic teaching framework this label risks excluding talk whose agents, character and purposes go beyond what the authors list. I therefore prefer the more precise term ‘extending’, since in relation to students’ contributions and hence their thinking that is what the Michaels and O’Connor moves aim to do.
11. Just as IRE is dominated by the teacher’s utterances, so traditional interaction analysis coding instruments tend to explicate the first move in the IRE sequence in greater detail than the other two. The landmark Flanders FIAC instrument from which so many later observation instruments derive (Flanders 1960) had seven kinds of teacher talk but only two for the student and a pretty basic and content-free twosome at that: ‘Student talk – response’ and ‘Student talk – initiation’. We might be unconcerned about this on the grounds that it accurately reflected the way teaching was in the 1960s – if it invariably was, which I doubt – were it not for the fact, even in the context of the analysis of dialogue, that even today ‘we still make teacher talk, and especially teachers’ questions, the centre of observational and analytical gravity, providing many more categories of teacher talk than of student talk and therefore allowing ourselves a far less nuanced study of the latter.’ (Alexander 2015, 433).
12. In England, the primary phase caters for children aged 4–11.
13. The description in this section draws on the account provided in the evaluation report of Jay et al. (2017).
14. Test details are at <https://educationendowmentfoundation.org.uk/our-work/projects/dialogic-teaching/>.
15. KS1 point score: students’ scores in the tests at the end of Key Stage 1 (Y2, US 1st grade). These were used not to establish a baseline but to assess how far the intervention’s impact depended on prior attainment.
16. A full account of the independent process evaluation, and copies of relevant protocols, appear in Jay et al. 2017.
17. This section draws on Jan Hardman’s contribution to Alexander et al. (2017).
18. The dialogic teaching indicators were modified from Alexander (2017a, 40–44).
19. For coding frames and details of the analytical procedures and software used, see Alexander et al. (2017).
20. See links to media stories at http://cprtrust.org.uk/about_cprrt/media/media-coverage/.
21. This section draws on the contributions of Taha Rajab and Mark Longmore to analysis of the interview data and to its reporting in Alexander et al. (2017).
22. Alexander (2017a, 5–6).
23. This section again draws on Jan Hardman’s contribution to Alexander et al. (2017). The figures and tables referred to may be viewed at <http://www.robinalexander.org.uk/wp-content/uploads/2017/07/Alexander-et-al-EEF-in-house-interim-report-final-170714.pdf>.
24. As explained earlier, Michaels and O’Connor (2012).
25. One of the best-known examples of a lack of congruence between pedagogical form and intention is Maurice Galton’s finding, persisting when his initial ORACLE project was repeated 20 years later, that in English primary classrooms students were almost always observed seated *in* groups but rarely working *as* groups (Galton, Simon, and Croll 1980; Galton et al. 1999).
26. I am grateful to Frank Hardman and Harvey Goldstein for their technical assessments of this project’s RCT.
27. <https://educationendowmentfoundation.org.uk/news/eef-publishes-four-new-independent-evaluations/>, 7 July 2017. Accessed May 2018.
28. This episode and the attendant evidence are presented in Alexander (2010, 240–243, 2011).

Acknowledgements

The initiative discussed here was supported 2014–2017 by the Education Endowment Foundation (EEF) as a joint project of Cambridge Primary Review Trust and the University of York. Thanks are due to EEF for their generous grant, to project colleagues Frank Hardman (my co-director), Jan Hardman, Taha Rajab, David Reedy and Mark Longmore, and to the students, teachers, mentors and heads in the project's schools in London, Birmingham, Bradford and Leeds.

Sarah Michaels and Cathy O'Connor kindly allowed their talk moves to be incorporated into the project's version of my dialogic teaching framework. I thank them and the many others from whose excursions into pedagogy in general and classroom talk in particular I have gained so much over the years, especially Douglas Barnes, Courtney Cazden, Tony Edwards, Maurice Galton, Adam Lefstein, Neil Mercer, Martin Nystrand, Lauren Resnick, and the late Jerome Bruner and Brian Simon.

This paper was initially presented in the symposium 'Professional Development in Dialogic Teaching: commonalities and constraints' at the 2017 biennial EARLI conference in Tampere, Finland. I am grateful to my symposium colleagues Alexander Grösschner, Jonathan Osborne, Alina Reznitskaya and Ian Wilkinson for giving me the opportunity to share and discuss the EEF project, the ideas that informed it and the issues it has raised.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Robin Alexander is Fellow of Wolfson College at the University of Cambridge, Professor of Education Emeritus at the University of Warwick and Honorary Professor of Education at the University of York. www.robinalexander.org.uk

References

- Adey, P., and M. Shayer. 1993. "An Exploration of Long-Term Far-Transfer Effects following an Extended Intervention Program in the High School Science Curriculum." *Cognition and Instruction* 11 (1): 1–29.
- Alexander, R. J. 1984. *Primary Teaching*. London: Cassell.
- Alexander, R. J. 1988. "Garden or Jungle? Teacher Development and Informal Primary Education." In *Informal Primary Education Today: Essays and Studies*, edited by W. A. L. Blyth, 148–188. London: Falmer Press.
- Alexander, R. J. 2001. *Culture and Pedagogy: International Comparisons in Primary Education*. Oxford: Blackwell.
- Alexander, R. J. 2003. *Talk for Learning: The First Year*. Northallerton: North Yorkshire County Council. http://www.robinalexander.org.uk/docs/NYorks_EVAL_REP_03.pdf.
- Alexander, R. J. 2004. *Towards Dialogic Teaching: Rethinking Classroom Talk*. 1st ed. York: Dialogos.
- Alexander, R. J. 2005a. *Talk for Learning: The Second Year*. Northallerton: North Yorkshire County Council. http://www.robinalexander.org.uk/docs/TLP_Eval_Report_04.pdf.
- Alexander, R. J. 2005b. *Teaching through Dialogue: The First Year*. London: Barking and Dagenham Council. <http://www.robinalexander.org.uk/bardagreport05.pdf>.
- Alexander, R. J. 2006. *Education as Dialogue: Moral and Pedagogical Choices for a Runaway World*. Hong Kong and York: HKIEd and Dialogos.
- Alexander, R. J. 2008. *Essays on Pedagogy*. London: Routledge.
- Alexander, R. J. 2009. "Towards a Comparative Pedagogy." In *International Handbook of Comparative Education*, edited by R. Cowen and A. M. Kazamias, 922–941. New York, NY: Springer.

- Alexander, R. J., ed. 2010. *Children, Their World, Their Education: Final Report and Recommendations of the Cambridge Primary Review*. London: Routledge.
- Alexander, R. J. 2011. "Evidence, Rhetoric and Collateral Damage: The Problematic Pursuit of "World Class" Standards." *Cambridge Journal of Education* 41 (3): 265–286.
- Alexander, R. J. 2015. "Dialogic Pedagogy at Scale: Oblique Perspectives." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 429–440. Washington, DC: AERA.
- Alexander, R. J. 2017a. *Towards Dialogic Teaching: Rethinking Classroom Talk*. 5th ed. York: Dialogos.
- Alexander, R. J. 2017b. *Dialogic Teaching and the Study of Classroom Talk: A Developmental Bibliography*. <http://www.robinaalexander.org.uk/wp-content/uploads/2017/01/Alexander-dialogic-teaching-bibliography.pdf>.
- Alexander, R. J., ed. 2017c. *The CPRT/IEE Dialogic Teaching Project, Trial Stage, 2015–16: Handbook for Schools*. York: Cambridge Primary Review Trust and University of York.
- Alexander, R. J., ed. 2017d. *The CPRT/IEE Dialogic Teaching Project, Trial Stage, 2015–16: Planning and Review Forms for Mentors*. York: Cambridge Primary Review Trust and University of York.
- Alexander, R. J., C. Doddington, J. M. Gray, L. Hargeaves, and R. Kerschner, eds. 2010. *The Cambridge Primary Review Research Surveys*. London: Routledge.
- Alexander, R. J., F. Hardman, J. Hardman, T. Rajab, and M. Longmore. 2017. *Changing Talk, Changing Thinking: Interim Report from the In-House Evaluation of the CPRT/UoY Dialogic Teaching Project*. York: University of York. <http://www.robinaalexander.org.uk/wp-content/uploads/2017/07/Alexander-et-al-EEF-in-house-interim-report-final-170714.pdf>.
- Alexander, R. J., J. Rose, and C. Woodhood. 1991. *Curriculum Organisation and Classroom Practice in Primary Schools*. London: DES.
- Alexander, R. J., and J. Willcocks. 1995. "Task, Time and Talk." In *Versions of Primary Education*, by R. J. Alexander with contributions from J. Willcocks, K. Kinder and N. Nelson, 103–219. London: Routledge.
- Bakhtin, M. M. 1981. *The Dialogic Imagination*. Austin, TX: University of Texas.
- Bakhtin, M. M. 1986. *Speech Genres and Other Late Essays*. Austin, TX: University of Texas.
- Barnes, D., J. Britten, and H. Rosen. 1969. *Language, the Learner and the School*. Harmondsworth: Penguin.
- Barnes, D., and F. Todd. 1977. *Communication and Learning in Small Groups*. London: Routledge and Kegan Paul.
- Barnes, D., and F. Todd. 1995. *Communication and Learning Revisited: Making Meaning through Talk*. Portsmouth NH: Heinemann.
- Berliner, D. C. 2002. "Educational Research: The Hardest Science of All." *Educational Researcher* 31: 18–20.
- Berliner, D. C., and B. J. Biddle. 1995. *The Manufactured Crisis: Myths, Fraud and the Attack on America's Public Schools*. Cambridge MA: Perseus Books.
- Black, P., C. Harrison, C. Lee, B. Marshall, and D. Wiliam. 2003. *Assessment for Learning: Putting It into Practice*. Maidenhead: Open University Press.
- Black, P., and D. Wiliam. 1998. *Inside the Black Box*. Slough: NFER-Nelson.
- Britton, J. 1969. "Talking to Learn." In *Language, the Learner and the School*, by D. Barnes, J. Britton and H. Rosen, 81–115. Harmondsworth: Penguin Books.
- Bruner, J. S. 1978. "The Role of Dialogue in Language Acquisition." In *The Child's Conception of Language*, edited by A. Sinclair, R. Jarvella and W. Levelt. NY: Springer.
- Bruner, J. S. 1983. *Child's Talk: Learning to Use Language*. Oxford: OUP.
- Bruner, J. S. 1996. *The Culture of Education*. Cambridge, MA: Harvard University Press.
- Bruner, J. S. 2006. *In Search of Pedagogy*. Vol. 1. London: Routledge.
- Bruner, J. S., and H. E. Haste. 1987. *Making Sense: The Child's Construction of the World*. London: Routledge.
- Cazden, C. B. 2001. *Classroom Discourse: The Language of Teaching and Learning*. Portsmouth NH: Heinemann.
- Dawes, L., N. Mercer, and R. Wegerif. 2004. *Thinking Together: A Programme of Activities for Developing Speaking and Listening*. Birmingham: Imaginative Minds.

- Deakin Crick, R., M. Taylor, S. Ritchie, E. Samuel, and K. Durant. 2005. *A Systematic Review of the Impact of Citizenship Education on Student Learning and Achievement*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Department for Education and Employment (DfEE). 1998. *The National Literacy Strategy: Framework for Teaching*. London: DfEE.
- Department for Education and Employment (DfEE). 1999. *The National Numeracy Strategy: Framework for Teaching Mathematics from Reception to Year 6*. London: DfEE.
- Doyle, W. 1983. "Academic Work." *Review of Educational Research* 53: 159–199.
- Edwards, A. D. 1992. "Teacher Talk and Pupil Competence." In *Thinking Voices: The Work of the National Oracy Project*, edited by K. Norman. London: Hodder.
- Flanders, N. A. 1960. *Interaction Analysis in the Classroom: A Manual for Observers*. Ann Arbor: University of Michigan.
- Flynn, J. R. 1987. "Massive IQ Gains in 14 Nations: What IQ Tests Really Measure." *Psychological Bulletin* 101 (2): 17–91.
- Galton, M. 2008. *Creative Practitioners in Schools and Classrooms*. Cambridge: University of Cambridge Faculty of Education.
- Galton, M. J., L. Hargreaves, C. Comber, D. Wall, and A. Pell. 1999. *Inside the Primary Classroom: 20 Years on*. London: Routledge.
- Galton, M., B. Simon, and P. Croll. 1980. *Inside the Primary Classroom*. London: Routledge.
- Ginsburg, A., and M. S. Smith. 2016. *Do Randomized Control Trials Meet the 'Gold Standard'?*. Washington, DC: American Enterprise Institute.
- Goswami, U. 2015. *Children's Cognitive Development and Learning*. CPRT Research Survey 3. York: Cambridge Primary Review Trust.
- Hardman, F., F. Smith, and K. Wall. 2003. "Interactive Whole Class Teaching" in the National Literacy Strategy." *Cambridge Journal of Education* 33 (2): 197–215.
- Harlen, W. 2014. *Assessment, Standards and Quality of Learning in Primary Education*. CPRT Research Survey 2. York: Cambridge Primary Review Trust.
- Hattie, J. 2009. *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*. London: Routledge.
- Higgins, S., D. Kokatsaki, and R. Coe. 2012. *The Teaching and Learning Toolkit: Technical Appendices*. London: Education Endowment Foundation/Sutton Trust. Accessed May 2018. [https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Technical_Appendices_\(July_2012\).pdf](https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Technical_Appendices_(July_2012).pdf)
- Holquist, M. 2002. *Dialogism: Bakhtin and His World*. London: Routledge.
- Jackson, P. W. 1968. *Life in Classrooms*. New York: Holt, Rinehart and Winston.
- Jay, T., R. Taylor, N. Moore, C. Burnett, G. Merchant, P. Thomas, B. Willis, and A. Stevens. 2017. *Dialogic Teaching: Evaluation Report and Executive Summary*. London: Education Endowment Foundation with Sheffield Hallam University. https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Dialogic_Teaching_Evaluation_Report.pdf.
- Kim, M.-Y. and I. A. G. Wilkinson. 2018. *What is Dialogic Teaching? Constructing, Deconstructing and Reconstructing a Pedagogy of Classroom Talk*. Unpublished manuscript. Columbus, OH: Department of Teaching and Learning, Ohio State University.
- Koedinger, K. R., and E. S. Wiese. 2015. "Accounting for Socializing Intelligence with the Knowledge-Learning-Instruction Framework." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 275–288. Washington, DC: AERA.
- Kuhn, D., and N. Zillmer. 2015. "Developing Norms of Discourse." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 77–86. Washington, DC: AERA.
- Lefstein, A. 2010. "More Helpful as Problem than Solution: Some Implications of Situating Dialogue in Classrooms." In *Educational Dialogues: Understanding and Promoting Productive Interaction*, edited by K. Littleton and C. Howe, 170–191. London: Taylor and Francis.
- Lefstein, A., and J. Snell. 2011. "Classroom Discourse: The Promise and Complexity of Dialogic Practice." In *Applied Linguistics and Primary School Teaching*, edited by S. Ellis and E. McCartney, 165–185. Cambridge: Cambridge University Press.

- Lefstein, A., and J. Snell. 2014. *Better than Best Practice: Developing Teaching and Learning through Dialogue*. London: Routledge.
- Mehan, H. 1979. *Learning Lessons: Social Organization in the Classroom*. Cambridge MA: Harvard University Press.
- Mehan, H., and C. B. Cazden. 2015. "The Study of Classroom Discourse: Early History and Current Developments." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 13–34. Washington, DC: AERA.
- Mercer, N. 2000. *Words and Minds: How We Use Language to Think Together*. London: Routledge.
- Mercer, N., and S. Hodgkinson, eds. 2008. *Exploring Talk in School*. London: Sage.
- Mercer, N., and K. Littleton. 2007. *Dialogue and the Development of Children's Thinking: A Sociocultural Approach*. London: Routledge.
- Michaels, S., and C. O'Connor. 2012. *Talk Science Primer*. Cambridge MA: TERC.
- Michaels, S., and C. O'Connor. 2015. "Conceptualizing Talk Moves as Tools: Professional Development Approaches for Academically Productive Discussions." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 347–361. Washington, DC: AERA.
- Michaels, S., C. O'Connor, and L. B. Resnick. 2008. "Deliberative Discourse Idealized and Realized: Accountable Talk in the Classroom and Civic Life." *Studies in Philosophy and Education* 27 (4): 283–297.
- Mortimer, E. F., and P. H. Scott. 2003. *Meaning Making in Secondary Science Classrooms*. Buckingham: Open University Press.
- Neisser, U. E. 1998. *The Rising Curve: Long-Term Gains in IQ and Related Measures*. Washington, DC: American Psychological Association.
- Nevill, C. 2016. *Do EEF Trials Meet the New "Gold Standard"?*. London: Education Endowment Foundation.
- Norman, K. 1992. *Thinking Voices: The Work of the National Oracy Project*. London: Hodder and Stoughton.
- Norman, G. 2003. "RCT = Results Confounded and Trivial: The Perils of Grand Educational Experiments." *Medical Education* 37: 582–584.
- Nystrand, M., A. Gamoran, R. Kachur, and C. Prendergast. 1997. *Opening Dialogue: Understanding the Dynamics of Language and Learning in the English Classroom*. New York, NY: Teachers College Press.
- Nystrand, M., L. L. Wu, A. Gamoran, S. Zeiser, and D. A. Long. 2003. "Questions in Time: Investigating the Structure and Dynamics of Unfolding Classroom Discourse." *Discourse Processes*. 35 (2): 135–198.
- Pogrow, S. 2017. "The Failure of the US Education Research Establishment to Identify Effective Practices: Beware Effective Practices Policies." *Education Policy Analysis Archives* 25 (5): 2–19.
- Pollard, A. 1985. *The Social World of the Primary School*. London: Cassell.
- Prideaux, D. 2002. "Researching the Outcomes of Educational Interventions: A Matter of Design." *BMJ* 324: 126–127.
- Resnick, L. B. 2015. "Talking to Learn: The Promise and Challenge of Dialogic Teaching." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 441–450. Washington, DC: AERA.
- Resnick, L. B., C. S. C. Asterhan, and S. N. Clarke. 2015. "Talk, Learning and Teaching." In *Socializing Intelligence through Academic Talk and Dialogue*, edited by L. B. Resnick, C. S. C. Asterhan and S. N. Clarke, 1–12. Washington, DC: AERA.
- Resnick, L., S. Michaels, and C. O'Connor. 2010. "How Well Structured Talk Builds the Mind." In *From Genes to Context: New Discoveries about Learning from Educational Research and Their Applications*, edited by R. Sternberg and D. Preiss, 163–194. New York: Springer.
- Reznitskaya, A. 2012. "Dialogic Teaching: Rethinking Language Use during Literature Discussions." *Reading Teacher* 65 (7): 446–456.
- Reznitskaya, A., and M. Gregory. 2013. "Student Thought and Classroom Language: Examining the Mechanisms of Change in Dialogic Teaching." *Educational Psychologist* 48 (2): 114–133.

- Rowe, M. B. 1986. "Wait Time: Slowing down May Be a Way of Speeding up." *Journal of Teacher Education* 37 (1): 43–50.
- Rowe, M. B. 1974. "Wait-Time and Rewards as Instructional Variables, Their Influence on Language, Logic, and Fate Control: Part One-Wait-Time." *Journal of Research in Science Teaching* 11 (2): 81–94.
- Schön, D. A. 1983. *The Reflective Practitioner: How Professionals Think in Action*. London: Temple Smith.
- Sinclair, J. Mc H, and R. M. Coulthard. 1975. *Towards an Analysis of Discourse*. Oxford: Oxford University Press.
- Smith, F, F. Hardman, K. Wall, and M. Mroz. 2004. "Interactive Whole Class Teaching in the National Literacy and Numeracy Strategies." *British Educational Research Journal* 30 (3): 403–419.
- Sullivan, G. M. 2011. "Getting off the "Gold Standard": Randomized Control Trials and Educational Research." *Journal of Graduate Medical Education* 3 (3): 285–289.
- Tough, J. 1979. *Talk for Teaching and Learning*. London: Ward Lock Educational.
- Vygotsky, L. S. 1962. *Thought and Language*. Cambridge, MA: MIT Press.
- Vygotsky, L. S. 1978. *Mind in Society*. Cambridge, MA: Harvard University Press.
- Wood, D. 1998. *How Children Think and Learn*. Oxford: Blackwell.
- Wragg, E. C., and G. Brown. 1993. *Explaining*. London: Routledge.
- Wragg, E. C., and G. Brown. 2001. *Questioning*. London: Routledge.