International Journal of Educational Research xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

International Journal of Educational Research

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

The role of dialog in philosophy for children

K.J. Topping [a,*], S. Trickey [b]

[a] University of Dundee, Scotland, UK[1]
[b] American University, Washington DC, USA

ARTICLE INFO

Article history:
Received 17 May 2012
Accepted 8 January 2013
Available online xxx

Keywords:
Philosophical enquiry
Primary
Elementary
Interaction
Dialog
Questioning
Reasoning

ABSTRACT

Philosophy for Children (P4C) stems from the work of Mathew Lipman. It involves teacher modeling of exploratory talk in relation to a complex stimulus, then has students discuss together in pairs, larger groups, and the whole class to achieve consensus. The effects of P4C on quantity and quality of interactive dialog on 180 10-year-old children in primary (elementary) mainstream classes in Scotland were investigated. Video recordings of teacher-led and pupil-pupil classroom dialog before, and seven months into, participation in the program were analyzed. Changes in intervention classes included: increased use of open-ended questions by the teacher, increased participation of pupils in classroom dialog, and improved pupil reasoning in justification of opinions. There were no changes in comparison classes. Increases were found in the amount of teacher-student dialog and the amount of student-student dialog, with decreases in the amount of teacher-whole-class talk. This study found gains in cognition, which were sustained from elementary school into high school even when no further P4C was done. Socio-emotional aspects of learning also improved. Investigation of process found changes in teacher behavior and child behavior. Implications for future research, policy and practice are outlined, particularly for enhancing consistency and expanding the program.

© 2013 Elsevier Ltd. All rights reserved.

The purpose of this paper is to explore the Philosophy for Children method and investigate the relationship between program effectiveness and any changes in teacher and student dialog in implementation with "normal" class sizes.

No other studies have done this, although a number of examples of successful interventions in thinking skills have been reported (see reviews by Adey & Shayer, 1994; Andrews, Torgerson, Low, & McGuinn, 2009; McGuinness, 1999; Moseley et al., 2004; Reznitskaya et al., 2009; Sternberg & Bhana, 1996; Thinking Skills Review Group, 2004). Nussbaum (2011) offered a theoretical review of alternative frameworks of argumentation that could be useful for both research and practice. However, many of these cognitive interventions were intensive, long-lasting and costly, with implications for cost-effectiveness, sustainability and replicability. Some involved high teacher/pupil ratios (e.g. Feuerstein, Rand, Hoffman, & Miller, 1980). Relatively few offered evidence of implementation integrity.

Implementing effective dialogic collaborative learning requires changes in the verbal and non-verbal behavior of both teacher and pupils (Mercer, 2000; Topping & Ehly, 1998). This chapter explores what variability in implementation in terms of process changes in teacher and pupil behavior were associated with cognitive gains. Other research reporting outcomes such as cognitive gains is briefly addressed. The thinking skills intervention was an adaptation of Lipman's (1981) "Philosophy for Children" (P4C). This emphasizes the development of critical reasoning skills through questioning and dialog

* Corresponding author at: School of Education, University of Dundee, Nethergate, Dundee DD1 4HN, Scotland, UK. Tel.: +44 01382 381400; fax: +44 01382 381511.
E-mail address: k.j.topping@dundee.ac.uk (K.J. Topping).
[1] http://www.dundee.ac.uk/eswce/people/kjtopping.htm.

0883-0355/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijer.2013.01.002

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

between pupil, teacher and peers. However, changing teacher and pupil verbal behavior is neither straightforward nor easy, as the literature testifies.

1. Context from the literature

Systematic investigation of classroom dialog began in the 1960s in the UK, when studies first noted that teachers rather than children did most of the talking and that there was a lack of open questioning (Alexander, 2004). Subsequent studies (Galton & Simon, 1980; Kumpulainen & Wray, 2002; Mercer, 2000; Wegerif, Mercer, & Dawes, 1999) evidenced alarming findings: that teacher-led discussions mostly only offered students the opportunity to make the briefest of responses, closed questions predominated, teacher talk which challenged pupils to think for themselves was scarce, many classroom questions had a low level of cognitive demand, and autonomous pupil discussion and problem-solving were rare. However, Boyd and Rubin (2006) found that despite the picture of teachers controlling talk patterns and of students producing minimal amounts of mainly procedural talk or recitation-type talk, authenticity of teacher questions was not necessarily a condition for triggering socially engaged and structurally elaborated student utterances. Teacher questions that elicited extended student talk were contingent on previous student utterances, rather being open-ended.

Teacher asking of higher order questions is however widely associated with higher student achievement (Redfield & Rousseau, 1981; Samson, Strykowski, Weinstein, & Walberg, 1987). Nonetheless, over an average school year, a pupil in elementary or secondary school in the US was found to ask only one question per month (Dillon, 1988, 1990). Dillon observed that in 1 h, teachers asked 84 questions, whilst all the pupils combined only asked two. Most teacher/student interaction consisted of rapid teacher questions requiring factual answers, allowing little room for pupil questions or meaningful discussion, and primarily testing memory. Dillon suggested that inhibiting factors were the teacher's need to "cover the curriculum" and a concern that increasing student autonomy would be a threat to discipline.

English, Hargreaves, and Hislam (2002) reported that only ten per cent of children's oral contributions during the 'literacy hour' in primary classrooms in England consisted of more than three words. Uninterrupted interactions of more than 25 s between the teacher and a child/group of children had declined following the introduction of mandatory highly structured whole class activities. Indeed, Dull and Morrow (2008) found that while dialogic questioning can and does occur in classrooms, students in heterogeneous tracks and lower-income schools have fewer experiences with dialogic questioning. Elsewhere, even 'expert' literacy teachers were found not to use effective interactive behaviors such as open questioning consistently across teaching contexts (Topping & Ferguson, 2005). Additionally, these expert teachers were not aware of this inconsistency, or of effective behaviors they did not use at all.

When oral collaborative activities *are* well structured, they have been found to maintain children's on-task behavior more consistently than solitary written tasks (Alexander, 1995). Cohen and Manion (1980) found that open-ended questions were important to allow depth of discussion, clear up misunderstandings, test the limits of the respondent's knowledge, encourage co-operation and establish rapport, facilitate a truer assessment of the respondent's beliefs, and yield unanticipated and unexpected answers which might suggest hitherto un-thought of relationships or hypotheses. Webb's (1989) review of 19 studies in grades 2–11 found strong associations between student achievement and group interactions, particularly when the interactions involved *giving* high-level explanations or elaborations. Receiving explanations influenced achievement only with active application thereof on the part of the recipient (Webb, Troper, & Fall, 1995).

Dialog has been found important not only in subjects like reading (Fitch & Hulgin, 2008), but also in subjects like mathematics and science (Braund, 2009; Dawes, 2004; Oliveira, 2010). Pupil propositions and explanations have been found crucial for cognitive gains in collaborative learning in science (Christie, Tolmie, Thurston, Howe, & Topping, 2009; Howe et al., 2007; Thurston, Christie, Howe, Tolmie, & Topping, 2008). A study of peer tutoring in mathematics showed significant increases in use of math words and strategic dialog (Topping, Campbell, Douglas, & Smith, 2003). The question of embedding and sustainability is important. However, Jennings and Mills (2009) showed that it was possible to sustain a discourse of enquiry over a five-year period in an elementary school.

More recently, technology has been used to facilitate the development of dialog. Thus Mercer, Hennessy, and Warwick (2010) used interactive whiteboards as a tool for encouraging and supporting classroom dialog. Others have investigated the used of an Evidence-Based Dialog Mapping software tool (Okada & Shum, 2008). Groenke and Paulus (2008) successfully paired pre-service English teachers in online chat rooms with local middle school students to talk about young adult literature.

So does the problem lie with the teacher, not with the pupils? There is evidence that even when teachers seek to enhance interactivity in the classroom, pupils might not respond. Pupils might think that asking questions could suggest inability to understand, and be reluctant to ask questions because of possible negative reactions from their teachers or peers – so they protect their self worth. Timson (1997) found that student questioning increased when teacher questioning reduced *and* the threat to student self worth was lessened. Teacher/pupil interaction became more communicative when questions were chosen for discussion that perplexed pupils and teacher alike.

Evidence that there was a lack of 'good talk' in classrooms led Dawes, Mercer, and Wegerif (2004) to suggest that there was often a false assumption that children knew *how* to discuss. Shared ground rules facilitated development, not only in terms of explicit reasoning but also in verbal creativity (Mercer, Dawes, Wegerif, & Sams, 2004; Wegerif, 2005). Similarly, Meloth and Deering (1999) complained that the cooperative learning paradigm had not adequately examined students' verbal discourse in naturalistic classroom settings, nor adequately examined the kinds of task-specific talk that might be beneficial to learning.

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

However, cycles of dialog have been found in some peer learning contexts, albeit often with post-school learners (Person & Graesser, 1999). In some cases these were a significant improvement on the simple sequences commonly observed, and corresponded to the desirable Question – Response – Feedback – Collaboration – Evaluation model. The present study suggests a modification of this model for a philosophical enquiry intervention in schools, to a cycle of: Teacher Stimulus – Teacher Questioning – Pupil Response – Pupil Proposition – Pupil Agree/Disagree Judgment – Pupil Explanation/Rationale for Judgment – Pupil Explanation/Rationale for Own Position. This new model clearly implies the need for a considerable shift from teacher talk to pupil talk, as well as a need for teacher questioning to be open-ended if it is to lead to the ensuing elements in the cycle and raise the quality as well as quantity of pupil elaborations and active response to those elaborations.

However, creating such change in classrooms is not easy. For example, Hargreaves and Moyles (2002) aimed to promote 'deeper' forms of interactive teaching in their Study of Interactive Teaching project. The intervention included video feedback and opportunities for reflective teacher discussion of changing practice. Teachers succeeded in increasing their ratio of questions to statements. However, despite the intervention, pupil responses were rarely 'extended' and learning remained heavily teacher dominated. The quality of teacher questioning as well as its volume is important.

Clearly, intervention intentions are not the same as actual delivery. The present study explored the Philosophy for Children program with regard to teacher and pupil verbal behaviors in relation to implementation quality and outcomes. It investigated whether a structured and scaffolded intervention program could sustain implementation quality and impact upon teacher and pupil interactive verbal behavior in the classroom, at low resource cost delivered into normal size classes in a range of primary schools. It also sought to test the new theoretical model of dialogic development. In both respects the study addresses gaps in the literature.

2. Intervention method

The Philosophy for Children program (P4C) (Lipman, 1981, 2003; Lipman & Sharp, 1978; Lipman, Sharp, & Oscanyon, 1980) can be infused into a range of subject domains (as discussed in Fisher, 1999a), although it initially tends to be used with children as a separate activity (as in this study). P4C incorporates features that Adey (2001) has suggested are key for promoting cognitive skills and educational attainment, including the thoughtful dialog described as necessary for promoting rich learning environments, and the heightened awareness of thinking and learning processes important for metacognitive development (Carnell & Lodge, 2002; Fisher, 2007). A key element is the emphasis on developing a community approach to 'enquiry' in the classroom. The process is characterized by use of open-ended Socratic questioning by the teacher, which challenges children to think more independently, promotes reciprocal teacher-pupil and pupil-pupil dialog, and encourages children to ask questions of each other of a similar nature.

Trickey and Topping (2004) offered a systematic review of controlled outcome studies of the P4C method in elementary and high schools. Ten studies met the stringent criteria for inclusion, measuring outcomes by norm-referenced tests of reading, reasoning, cognitive ability and other curriculum-related abilities, by measures of self-esteem and child behavior, and by child and teacher questionnaires. All studies showed some positive outcomes. The mean effect size was 0.43 with low variance, indicating a consistent moderate positive effect for P4C on a wide range of outcome measures.

2.1. Thinking through philosophy program

However, Lipman's own writings on P4C tend to describe the process in somewhat general terms, potentially leaving teachers new to the method unclear about how to operationalize the principles and risking considerable variation across different implementations (making the consistency of the meta-analytic findings all the more remarkable). Haynes (2002) helpfully operationalized the process of philosophical enquiry in nine pragmatic steps: (1) getting started – e.g. agreeing rules of interaction, beginning with scripted relaxation exercise, aiming to create an alert but relaxed state; (2) sharing a stimulus to prompt enquiry; (3) pause for thought; (4) questioning – the pupils think of interesting or puzzling questions; (5) connections – making links between the questions; (6) choosing a question to begin an enquiry; (7) building on each other's ideas – during which the teacher has to strike a balance between encouraging the children to follow on from each other's ideas and allowing related lines of enquiry to open up; (8) recording discussion – graphic mapping; and (9) closure and review – summarizing, encouraging children to reflect on the discussion, reflecting on the process itself, how their thinking might have progressed during it, whether minds were changed.

The intervention in this study included all these elements, but used a more contemporary process – 'Thinking through Philosophy' (Cleghorn, 2002). This included specific structured lesson plans, designed to support relatively novice teachers and minimize variation in implementation quality. It also included a greater emphasis on linking each session with the previous session (curriculum progression and developmental continuity being under-emphasized in previous work) and explicitly incorporating pair work developing into group work (step-wise building of transferable collaborative group work skills being under-emphasized in previous work). Additionally, the components of effective peer dialog were more precisely specified, to enable teachers to more effectively formatively assess and intervene in dialog as it occurred. Finally, "take-away" thinking tasks were specified, to encourage generalization out of the P4C session and indeed beyond school, and indeed maintenance of cognitive gains (again under-emphasized in previous work).

Additional elements specified in Cleghorn's (2002) program were: (1) linking with the previous week-reinforcing memory of what took place the previous week and providing an opportunity to bring forward any new related thinking; (2)

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

pair/group work – an opportunity to check the children's initial understanding of any story or other thinking stimulus; (3) scaffolded dialog – the teacher encouraging pupils to: communicate their views in response to an agreed subject of 'enquiry', support their views with reasons, listen respectfully to views being expressed, indicate whether they agree or disagree with those views, provide alternative viewpoints, and gradually develop a process of dialog that helps the class construct a deeper understanding (or better solution) than would be possible individually; (4) thought for the week – highlighting a practical idea to provide 'homework' for the rest of the week, to help relate that idea to real situations. Thus the structure for each intervention session was held constant, while teachers remained free to exercise their own choice of stimulus, mainly from but not confined to the stimulus bank supplied, as deemed relevant to class age, needs and developmental progression.

In-service professional development support was led by Cleghorn, aided by two senior teachers who were already experienced in leading classroom enquiry. During the period of the evaluation, the combined time allocation of these three teachers to the initiative amounted to 0.2 full time equivalent. Intervention class teachers from the last two primary years (pupils leave primary school at the age of 12 in the region studied) received 10–12 h of professional development during the first year of the initiative. Pre-service, this included a full day input plus observation of an expert classroom practitioner and debriefing with that practitioner. Subsequently, each term intervention teachers attended a 2-h after-school group session to share progress and talk through issues arising. The teachers thus formed their own "community of enquiry". Further guidance was available from the support teachers on a 'call-out' basis. As this research involved evaluating a natural curricular development which was in half the schools in the authority, permissions were sought at head teacher and class teacher level, but not at child or parent level.

2.2. Related research

The present chapter focuses upon interaction and dialog, but the overall project investigated other aspects of functioning which are briefly mentioned here. Firstly, cognitive gains were made by students participating in P4C. Primary school pupils aged 10 at pre-test engaged in collaborative enquiry 1 h per week over 16 months. Randomly selected intervention children ($n = 105$, 4 classes/schools) showed significant pre-post gains on the Cognitive Abilities Test, while controls following a normal curriculum ($n = 72$, 3 classes/2 schools) did not. Significant gains for intervention children were evident in verbal and in non-verbal and quantitative aspects of reasoning, consistent across intervention schools. Both boys and girls made significant gains. The highest quartile of pre-test ability showed the smallest gains (Topping & Trickey, 2007a).

There was a follow-up study of the same pupils, at long-term two-year follow-up, after the participants had transferred to secondary (high) school without experiencing further philosophical enquiry in the interim. Sample attrition was greater in the control than in the experimental group; 96 experimental and 52 control subjects were available. The significant pre-post cognitive ability gains in the experimental group in primary school were maintained toward the end of their second year of secondary school. Higher achieving pupils were somewhat advantaged in sustaining these gains. The control group showed an insignificant but persistent deterioration in scores from pre-test to post-test to follow-up (Topping & Trickey, 2007b).

Concern with academic gain is all very well, but there was also interest in other effects. Two measures were used to investigate the socio-emotional effects on children aged 11 at pre-test in five experimental and three control primary school mainstream classes. In a pre-post design, experimental pupils used collaborative enquiry for 1 h each week over a seven month period, while control pupils followed a normal curriculum. On a test of self-esteem as a learner, experimental pupils ($n = 119$) gained significantly while controls ($n = 52$) did not. In the former there was evidence of significant reduction in dependency and anxiety, and of greater self-confidence. Girls tended to gain more than boys. These results were fairly consistent across schools and classes (Trickey & Topping, 2006).

Finally, the views and perceptions of the participants were investigated – the subjective evaluative responses of student ($n = 77$), class teacher ($n = 16$) and head teacher ($n = 8$) participants. The 11-year-old students from three randomly selected elementary school classes had used collaborative enquiry for 1 h weekly over seven months. Student responses indicated a majority very much enjoyed it (process more than content), although positivity varied between schools. Widening participation, cognitive gains and improvement in listening and concentration were commonly reported. Half the students reported gains in "emotional intelligence", particularly relationships, social behavior and empathy, self-confidence, and self-regulation of emotion. Two thirds of students reported generalization of effects outside the enquiry sessions. Student perceptions were largely confirmed by the teachers (Trickey & Topping, 2007).

3. Research aims

The present study aimed to:

1 Investigate whether a structured and scaffolded intervention program would impact desired teacher and pupil verbal behavior in the classroom, despite low resource cost and delivery into normal size classes in diverse primary schools.
2 Investigate whether program implementation was associated with (i) increased use of open-ended questions by the teacher; (ii) increased participation of pupils in classroom discussion, and (iii) increased rational underpinning for pupil judgments.

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

3 Explore variation or consistency across schools.

As cognitive gains had already been found, it was expected that at least some of the process behaviors the literature suggested were implicated in these gains would be found. Which, and how many, and how consistently, were all more difficult to predict. Changes in quality of interaction as well as changes in quantity of interaction were expected.

4. Method

4.1. Research design

The research design was a 2×2 pre-post intervention/comparison design, comparing pre-program to during-program verbal behavioral change in intervention group teachers and pupils with comparison group change. It was quasi-experimental as sampling was not totally random and not all school/class effects could be controlled. Both groups were observed under the same conditions on both occasions.

4.2. Context and sample

The intervention was located in a small ethnically homogeneous local education authority (school district) of mixed socio-economic status but including pockets of severe disadvantage. The authority refused to accept totally random sampling for this study. All class teachers from the last two years of primary school were invited to a meeting where attendees were invited to participate in the first phase of the intervention. All 19 schools expressed interest, but existing commitments to school development plans left eight schools that were free to engage with the first phase. The eight initial intervention schools were thus to some degree self-selected, but not on the basis of highest motivation or best organization. Each intervention school had two intervention classes.

The schools, teachers and pupils in this study were drawn from the schools participating in the intervention study. Four comparison schools (who had of course also expressed interest in participating in the intervention) were selected as comparable in terms of balance of pupil ability, size of school and social disadvantage factors. The pre-test mean standardized cognitive ability score was 99 for the intervention pupils (standard deviation 12.1) and 101 for the comparison pupils (s.d. 11.1), both normally distributed. Class size averaged 25 pupils with no difference between experimental and control. The mean roll (total number of pupils) for the intervention schools was 328, for comparison schools 365. The average number of children receiving free school meals (as an indicator of socio-economic disadvantage) was 52 in intervention schools and 61 in the comparison schools. The comparison children received classroom experiences and the class teachers continuing professional development experiences "normal" for the school district during the same time as the intervention, which included some emphasis upon pupil participation in formative assessment – also involving teacher-pupil dialog. However, class time on tasks involving dialog was specifically manipulated only in the intervention schools. All comparison classes subsequently became involved in the second phase of implementation. Obviously, not all possible relevant confounding variables were directly controlled or controllable, and effects attributable to unmeasured teacher variability and pupil variability were expected to add layers of "noise" to results in both intervention and comparison groups.

To make workload and curriculum intrusion acceptable, different types of measurement were randomly shared across the eight intervention and four comparison schools. Four intervention and two comparison classes from six different schools were randomly selected for the interaction analysis aspect of the study from the pools of intervention and comparison schools and classes.

4.3. Measures

Video recordings were made of classroom discussions of the same somewhat ambiguous Greek fable (Aesop's "Mercury and the Axe") in October in the six classes, prior to initiation of the program in the four intervention classes. The teacher read out the story and then explored its meaning with the class through collaborative enquiry. A number of "prompt questions" (from Fisher, 1999b, p. 45) were provided to structure common elements in the subsequent small group enquiry:Key question: What is truth?

1. Do you think this is a true story? Why?
2. What do we mean when we say something is true?
3. What do we call something that is not true? What does 'false' mean?
4. What is a lie?
5. What do we call a story which is not true? What is fiction/a fable/a fairy tale?
6. Which man in the story was honest? What does 'honest' mean?
7. Which man in the story was a liar? What does 'liar' mean?
8. Is it better to tell the truth or to tell lies? Why?

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

9. Have you ever told a lie? Can you say when or why?
10. Is it every right to tell a lie? Is it ever wrong to tell the truth?

All six teachers repeated the same task with the same instructions in May the following year, and all six class discussions were again video-recorded.

The video-recorded classroom discussion was scored using a structured observation schedule, which focused on five primary dependent variables:

1. Occurrence of teacher asking an open-ended question (including follow up questions)
2. Ratio of the amount of time pupils were talking to the amount of time their teacher was talking,
3. Number of discrete pupil utterances,
4. Occurrence of pupil agreeing or disagreeing with a view expressed by another pupil and providing a reason for this,
5. Occurrence of pupil supporting their own view/opinion with a reason.

A sixth secondary dependent variable (mean length of pupil utterance) was later derived.

This combination of variables were intended to capture changes in the balance of *quantity* of interaction as distributed across the actors, and also changes in the *quality* of interactive verbal behavior in teachers and pupils. They were chosen on the combined grounds of theoretical and practical importance (in relation to concerns expressed in the literature) and psychometric stability in terms of inter-rater reliability when observed (other possible variables found to lack the latter were discarded during piloting).

The average length of discussion was 30 min, but analysis of a sample of video recordings for the whole of the discussion yielded results no different from the first 10 min, and it was considered that the beginning of discussions offered a more valid comparison than the totality of discussions of very different lengths. Consequently, for all variables except the pupil/teacher talk variable, every instance of the specified verbal behavior occurring was tallied over the first ten continuous minutes of the discussion. For the pupil/teacher talk variable, observations were made every 30 s during the first 10 min. Twenty observations were therefore made per discussion. A percentage was then calculated for the number of times pupils were talking as a proportion of total observed talk. Two observers received training to enable them to use the schedule. The observers independently scored a random sample of video recordings and these scores were compared. All the components in all the variables had a high level of inter-observer agreement, minimally 85% (range 85–97%).

4.4. Analysis

Given the sampling constraints, non-parametric statistical analysis was considered appropriate. Pre-post comparisons were made within each group using the Wilcoxon Test, which is of course conservative compared to parametric tests.

5. Results

The quantitative results from the six pre-test and six post-test recordings for each school (intervention or comparison) are given separately in Table 1 to explore variation between schools. Overall, intervention schools as a group showed statistically significant pre-post gains on teacher open-ended questioning, proportion of pupil talk, pupil reasoned response to another's view, and pupil reasoned support of own view. Regarding number of discrete pupil utterances, half the intervention schools increased and half decreased, only one school showing a large increase. However, this variable is best interpreted in relation to total teacher/pupil talk. Comparison schools showed no significant gains, and their erratic pattern suggested this was not merely attributable to smaller total numbers in this group.

On average, intervention classes increased their ratio of pupil talk from 41% to 66%. Considerable variation was evident in the intervention schools. Intervention schools C and D appeared to be the biggest improvers (not necessarily from the lowest baseline), while school A showed more modest gains. School B, however, showed only slight gain coupled with slight decline.

From the pupil/teacher talk variable, the total number of pupil utterances showed little change, but the proportion of pupil talk to total classroom talk increased substantially, so pupil utterances must have tended to become longer. Of course, longer utterances are not necessarily better utterances, so quality variables must be considered in parallel. However, while the intervention classes increased the duration of their utterances (amount of time talking divided by number of utterances) on average by 58% (see Table 2), the comparison classes showed a decrease.

5.1. Qualitative impressions

In order to illuminate the process and quality of some of the thinking articulated by some of the intervention pupils, two very brief sample sections of transcript are given below:

Stuart (Discussing happiness): "Happiness is inside you. It is sort of chained to you so it can never leave you."
Teacher: "Do we use happiness up? If I am happy too much might I run out of it when I'm older?"

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

Table 1
Pre-post scores of intervention and comparison schools and group significance.

Variable	Exp school A PRE	Exp school A POST	Exp school B PRE	Exp school B POST	Exp school C PRE	Exp school C POST	Exp school D PRE	Exp school D POST	Exp group significance (p)	Comp school E PRE	Comp school E POST	Comp school F PRE	Comp school F POST	Comp group significance (p)
Occurrence of teacher asking an open-ended question	1	6	5	0	0	6	0	7	0.03	0	6	0	1	0.32
Percentage of time pupils speak as proportion of total talk	40	45	45	45	40	85	40	70	0.05	50	40	45	40	1.00
Number of discrete pupil utterances	34	41	33	20	31	27	21	22		34	62	40	44	
Occurrence of pupil agreeing/disagreeing with another pupil and providing a reason for this	0	1	0	1	0	7	0	3	0.05	0	0	0	0	0.18
Occurrence of pupil supporting their own opinion with a reason	2	6	0	4	6	9	6	12	0.03	9	6	7	3	0.16

Figures are tallies of observations of the specified behaviors during the 10 min observation period analyzed.

Holly: "No, it's like a car going to the petrol station to get more petrol. When something makes you happy it sort of tops it up again."

Teacher: "But where is happiness?"

Holly: "Happiness is not to be found without – it can only be found within". (This was the exact language used by Holly)

Hannah: "Things are not real happiness – they are just like a comfort blanket over you that makes you think you are happy. It's the same *you* underneath. Money won't make you happy."

Teacher: "Can you say more about that?"

Hannah: "Happiness is better thought of as a contentment inside."

Sam: "Happiness is like a spark in a fire – it's always there but it needs something extra to make it come to life."

* * * * * * * *

Teacher: "What does it mean to be human?"

Ashleigh: "You know you are human because everyone is equal."

Gemma: "Yes, everyone is equal, but different in their own special way."

Kelly: "Our feelings and thoughts are special to us."

Nick: "You are your own world."

Teacher: "What do you mean by that?"

Table 2
Duration of individual utterances before and after P4C in minutes.

	Average duration of pre-test utterance	Average duration of post-test utterance
Intervention school A	0.23	0.23
Intervention school B	0.27	0.45
Intervention school C	0.38	0.64
Intervention school D	0.38	0.63
Comparison school E	0.24	0.12
Comparison school F	0.23	0.18

Mean pre-test duration of utterance for intervention schools = 0.29.
Mean post-test duration of utterance for intervention schools = 0.48.
Both comparison schools declined in duration of utterance.

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

Nick: "Well, I think if I'm a nice person, the world will be nice. If I am feeling miserable the world looks bad and then I feel even worse! And so it goes on."

Laurie: "Our senses tell us about the world, but our feelings are what we are. That's why we are different from the animals."

6. Discussion

This study reports good news – that changes in teacher and pupil verbal behavior widely seen as important for the enhancement of cognitive skills can be obtained at modest intervention cost. However, these welcome results should not be over-interpreted, over-generalized, or accepted uncritically.

6.1. Interpretation

Teachers generally increased their use of open-ended questioning, but this was a little erratic, one or two teachers showing atypical results. On average the percentage of time that pupils contributed to discussion increased from 41% to 66%. This is a striking change in the traditional balance of pupil talk to teacher talk. It also implies that the children felt sufficiently confident to contribute in this way. The individual contributions each child made were not separately quantified. However, several teachers subsequently commented that there were always a number of children who did not participate well in the whole-class discussion (not necessarily the same children in different sessions). However this was helped by activities involving 'thinking in pairs' and 'thinking in fours' – suggesting the way the program was structured had an additive effect. The teachers considered that the thinking of children who did not actively participate in any whole-class discussion still benefited from listening to the contributions of their peers.

The results suggest that changes in pupil participation did not reflect larger numbers of brief comments being made by the children, but longer and more elaborated pupil comments. Longer pupil responses are of course not necessarily the same as better quality responses. However, when all six variables encompassing both quality and quantity of interaction are taken into account, teachers were talking less and asking more open questions, while children were talking more and demonstrating more elaborate thinking in contributions to discussion. Beyer (1985) defined critical thinking as 'the ability to assess the authenticity, accuracy and/or worth of knowledge claims and arguments'. Supporting an expressed view with reasoning would seem an important step in the process of developing critical thinking.

The study thus adds to the literature by indicating that, and how, a more structured and scaffolded intervention program can secure adequate implementation quality and impact upon desired teacher and pupil verbal behavior in the classroom, even at low resource cost in normal size classes in a range of primary schools. It also adds by supporting the new theoretical model, which speaks to the verbal behavioral features of dialogic collaborative learning in thinking skills holding promise from both research and practitioner perspectives.

6.2. Methodological issues

Obviously this study had imperfections. Sampling was random within the pool of schools, but members of that pool were self-selected (although there was no evidence of bias in the sample, and pure random sampling might have affected external validity). The findings may not be generalizable outside of the school district undertaking this project, and certainly not out of elementary school. The study relied on measurement of a limited number of variables, but these were chosen as suitable for the purpose and of established reliability. However, further triangulation with additional measures might have been advantageous had it been possible. The possibility of Hawthorne Effect must be considered, in terms of teacher response to the novelty of the intervention. However, the intervention period was rather long for Hawthorne Effects to be sustained. Experimenter effects from researchers recording classroom interaction on video must also be considered, although these pupils were very familiar with such recording occurring. While pupils' rationale for their own opinions and those of their peers were more frequently observed after the intervention, the depth and logical security of their rationale was not assessed in this study, not least because establishing adequate inter-assessor reliability on such judgments is extremely difficult.

6.3. Implications for future research, practice and policy

This study suggests several lines of possible future investigation. Longitudinal studies are needed to explore the sustainability of gains across time and contexts, especially over the transition from primary to secondary education, and from there to post-school. This is particularly relevant at a time when concerns have been expressed about a drop in the performance of pupils following their transition from primary to secondary education (e.g. Galton, Gray, & Rudduck, 1999). It would be useful to investigate the nature of teacher effects and exactly how differences between the verbal and non-verbal behaviors of teachers influence outcomes in dialog and thinking. Further research on the relationship between outcomes of philosophical enquiry and socio-economic factors, initial cognitive ability and gender factors is also desirable. More detailed investigation is necessary to explore whether there are differences in the ways that boys and girls respond to collaborative classroom processes.

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

From a policy and practice perspective, clarifying the verbal behavior changes involved in securing positive outcomes has implications for future intervention implementation integrity and quality improvement, replicability and sustainability. The new theoretical model is likely to be helpful in this. There was some evidence of individual 'teacher effects' in this study. Some teachers appeared more able to assimilate the concept of collaborative enquiry and put this into practice. If teachers operating collaborative enquiry are to capitalize on the gains apparently available, there are implications for their initial teacher training and their continuing professional development.

7. Conclusion

The present study found that:

1 The structured and scaffolded intervention program used did impact desired teacher and pupil verbal behavior in the classroom, despite low resource cost and delivery into normal size classes in diverse primary schools.
2 Program implementation was associated with (i) increased use of open-ended questions by the teacher; (ii) increased participation of pupils in classroom discussion, and (iii) increased rational underpinning for pupil judgments, of both the position of others and their own position. Intervention schools showed statistically significant pre-post gains on teacher open-ended questioning, proportion of pupil/teacher talk, pupil reasoned response to another's view, and pupil reasoned support of own view. Intervention classes on average increased their ratio of pupil/teacher talk from 41% to 66%. The number of pupil utterances did not increase, although the duration did. Comparison schools showed no significant gains.
3 Considerable variation was found across schools, indicating a need for further work on securing more consistent process gains. This might involve further refining the structuring and scaffolding in the program described.

Collaborative philosophical enquiry enhanced reciprocal communicative interaction in the classroom between teacher and children and between children and children, in terms of both quantity and quality. These interactions seem likely to be implicated in the summative cognitive gains also found in this project.

A systematic review by Baumfield, Butterworth, and Edwards (2005) concluded that thinking skills interventions have the potential to support and encourage teachers to develop pedagogy that enables students to achieve greater understanding, engagement and higher achievement, while also providing teachers with a vehicle for formative assessment and greater insight into pupils' learning. However, this process requires close partnerships and sustained involvement of teachers working together within and across schools. The present study echoes this finding, adding that the quantity and quality of dialog in the classroom seems key to harvesting these multiple benefits. Furthermore, fine-grain scaffolding of pupil/teacher dialog in communities of enquiry in diverse classrooms should be paralleled with scaffolding effective dialog in teacher communities of enquiry within and between schools.

References

Adey, P. (2001). Cognitive acceleration: Thinking as intelligence. *Teaching Thinking, 5*, 38–41.
Adey, P., & Shayer, M. (1994). *Really raising standards: Cognitive intervention and academic achievement*. London: Routledge.
Alexander, R. (1995). *Versions of primary education*. London: Routledge.
Alexander, R. (2004). *Towards dialogic teaching*. Cambridge, England: Dialogos.
Andrews, R., Torgerson, C., Low, G., & McGuinn, N. (2009). Teaching argument writing to 7- to 14-year-olds: An international review of the evidence of successful practice. *Cambridge Journal of Education, 39*(3), 291–310.
Baumfield, V. M., Butterworth, M., & Edwards, G. (2005). *The impact of the implementation of thinking skills programmes and approaches on teachers*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Beyer, B. K. (1985). Critical thinking: What is it? *Social Education, 49*, 270–276.
Boyd, M., & Rubin, D. (2006). How contingent questioning promotes extended student talk: A function of display questions. *Journal of Literacy Research, 38*(2), 141–169.
Braund, M. (2009). Talk in primary science: A method to promote productive and contextualised group discourse. *Education 3-13, 37*(4), 385–397.
Carnell, E., & Lodge, C. (2002). *Supporting effective learning*. London: Paul Chapman Publishing.
Christie, D., Tolmie, A., Thurston, A., Howe, C., & Topping, K. (2009). Supporting group work in Scottish primary classrooms: Improving the quality of collaborative dialogue. *Cambridge Journal of Education, 39*(1), 141–156.
Cleghorn, P. (2002). *Thinking through philosophy*. Blackburn, England: Educational Printing Services.
Cohen, L., & Manion, L. (1980). *Research methods in education*. London: Routledge.
Dawes, L. (2004). Talk and learning in classroom science. *International Journal of Science Education, 26*(6), 677–695.
Dawes, L., Mercer, N., & Wegerif, R. (2004). *Thinking together* (2nd edition). Birmingham: Imaginative Minds Ltd.
Dillon, J. T. (1988). *Questioning and teaching*. London: Routledge.
Dillon, J. T. (1990). *The practice of questioning*. London: Routledge.
Dull, L. J., & Morrow, S. E. (2008). Is dialogic questioning possible in social studies classrooms? *Theory and Research in Social Education, 36*(4), 391–412.
English, E., Hargreaves, L., & Hislam, J. (2002). Pedagogical dilemmas in the National Literacy Strategy: Primary teachers' perceptions, reflections and classroom behaviour. *Cambridge Journal of Education, 32*(1), 9–26.
Feuerstein, R., Rand, Y., Hoffman, M., & Miller, M. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability*. Baltimore, MD: University Park Press.
Fisher, R. (1999a). *Teaching thinking: Philosophical enquiry in the classroom*. London: Cassell.
Fisher, R. (1999b). *First stories for thinking*. Oxford: Nash Pollock.
Fisher, R. (2007). Dialogic teaching: Developing thinking and metacognition through philosophical discussion. *Early Child Development and Care, 177*(6-7), 615–631.
Fitch, E. F., & Hulgin, K. M. (2008). Achieving inclusion through CLAD: Collaborative learning assessment through dialogue. *International Journal of Inclusive Education, 12*(4), 423–439.
Galton, M., & Simon, B. (1980). *Inside the primary classroom*. London: Routledge.

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002

Galton, M., Gray, J., & Rudduck, J. (1999). *The impact of school transitions and transfers on pupil progress and attainment*. London: Department for Education and Employment.
Groenke, S. L., & Paulus, T. (2008). The role of teacher questioning in promoting dialogic literary inquiry in computer-mediated communication. *Journal of Research on Technology in Education, 40*(2), 141–164.
Hargreaves, L., & Moyles, J. (2002). How do school teachers define and implement 'interactive teaching' in the National Literacy Strategy in England? *Paper presented at the American Research Association Annual Meeting.*
Haynes, J. (2002). *Children as philosophers*. London: Routledge Falmer.
Howe, C., Tolmie, A., Thurston, A., Topping, K. J., Christie, D., Livingston, K., et al. (2007). Group work in elementary science: Organisational principles for classroom teaching. *Learning & Instruction, 17*(5), 549–563.
Jennings, L., & Mills, H. (2009). Constructing a discourse of inquiry: Findings from a five-year ethnography at one elementary school. *Teachers College Record, 111*(7), 1583–1618.
Kumpulainen, K., & Wray, D. (2002). *Classroom interaction and social learning: From theory to practice*. London: Routledge Falmer.
Lipman, M. (1981). Philosophy for children. In Costa, A. L. (Ed.). *Developing minds: Programs for teaching thinking* (vol. 2, pp. 35–38).Alexandria, VA: Association for Supervision and Curricular Development.
Lipman, M. (2003). *Thinking in education* (2nd edition). Cambridge & New York: Cambridge University Press.
Lipman, M., & Sharp, A. M. (1978). *Growing up with philosophy*. Philadelphia, PA: Temple University Press.
Lipman, M., Sharp, A. M., & Oscanyon, F. (1980). *Philosophy in the classroom*. Philadelphia, PA: Temple University Press.
McGuinness, C. (1999). *From thinking skills to thinking classrooms: A review and evaluation of approaches for developing pupil's thinking*. Research Report RR115. London: Department for Education and Employment.
Meloth, M. S., & Deering, P. D. (1999). The role of the teacher in promoting cognitive processing during collaborative learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 235–255). Mahwah, NJ: Lawrence Erlbaum.
Mercer, N. (2000). *Word and minds: How we use language to think together*. London: Routledge.
Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3), 359–377.
Mercer, N., Hennessy, S., & Warwick, P. (2010). Using interactive whiteboards to orchestrate classroom dialogue. *Technology, Pedagogy and Education, 19*(2), 195–209.
Moseley, D., Baumfield, V., Higgins, S., Lin, M., Miller, J., Newton, D., et al. (2004). *Thinking skill frameworks for post-16 learners: An evaluation*. Guildford: Learning and Skills Development Agency.
Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: Alternative frameworks for argumentation research in education. *Educational Psychologist, 46*(2), 84–106.
Okada, A., & Shum, S. B. (2008). Evidence-based dialogue maps as a research tool to investigate the quality of school pupils' scientific argumentation. *International Journal of Research & Method in Education, 31*(3), 291–315.
Oliveira, A. W. (2010). Developing elementary teachers' understanding of the discourse structure of inquiry-based science classrooms. *International Journal of Science and Mathematics Education, 8*(2), 247–269.
Person, N. K., & Graesser, A. G. (1999). Evolution of discourse during cross-age tutoring. In A. O'Donnell & A. King (Eds.), *Cognitive perspectives in peer learning* (pp. 69–86). Mahwah, NJ: Lawrence Erlbaum.
Redfield, D. L., & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research, 51*, 237–245.
Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R. C., et al. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education, 39*(1), 29–48.
Samson, G. E., Strykowski, B., Weinstein, T., & Walberg, H. J. (1987). The effects of teacher questioning levels on student achievement. *Journal of Educational Research, 80*, 290–295.
Sternberg, R., & Bhana, K. (1996). Synthesis of research on the effectiveness of intellectual skills programs: Snake oil remedies or miracle cures? *Educational Leadership, 44*(2), 60–67.
Thinking Skills Review Group. (2004). *Thinking skills approaches to effective teaching and learning: What is the evidence for impact on learners?* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Thurston, A., Christie, D., Howe, C. J., Tolmie, A., & Topping, K. J. (2008). Effects of continuing professional development on group work practices in Scottish primary schools. *Journal of In-service Education, 34*(3), 263–282.
Timson, J. (1997). *The role and potential of the teacher as a facilitator of classroom discussion*. Masters dissertation in education, University of Lancaster.
Trickey, S., & Topping, K. J. (2004). Philosophy for children: A systematic review. *Research Papers in Education, 19*(3), 365–380.
Topping, K. J., & Ehly, S. (Eds.). (1998). *Peer-assisted learning*. Mahwah, NJ/London, UK: Lawrence Erlbaum.
Topping, K. J., & Ferguson, N. (2005). Effective literacy teaching behaviours. *Journal of Research in Reading, 28*(2), 125–143.
Topping, K. J., & Trickey, S. (2007a). Collaborative philosophical enquiry for school children: Cognitive effects at 10-12 years. *British Journal of Educational Psychology, 77*, 271–288.
Topping, K. J., & Trickey, S. (2007b). Collaborative philosophical enquiry for school children: Cognitive gains at two-year follow-up. *British Journal of Educational Psychology, 77*(4), 787–796.
Topping, K. J., Campbell, J., Douglas, W., & Smith, A. (2003). Cross-age peer tutoring in mathematics with seven- and 11-year-olds: Influence on mathematical vocabulary, strategic dialogue and self-concept. *Educational Research, 45*(3), 287–308.
Trickey, S., & Topping, K. J. (2006). Collaborative philosophical enquiry for school children: Socio-emotional effects at 11-12 years. *School Psychology International, 27*(5), 599–614.
Trickey, S., & Topping, K. J. (2007). Collaborative philosophical enquiry for school children: Participant evaluation at 11 years. *Thinking: The Journal of Philosophy for Children, 18*(3), 23–34.
Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21–40.
Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87*, 406–423.
Wegerif, R. (2005). Reason and creativity in classroom dialogues. *Language and Education, 19*(3), 223–238.
Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction, 9*(5), 493–516.

K.J. Topping is professor of Educational and Social Research in the School of Education at the University of Dundee, Scotland.

S. Trickey is assistant professor at the American University, Washington. He was formerly a senior educational psychologist with Clackmannanshire Council, Scotland.

Please cite this article in press as: K.J. Topping, S. Trickey, The role of dialog in philosophy for children, International Journal of Educational Research (2013), http://dx.doi.org/10.1016/j.ijer.2013.01.002