Abstract

This study explores the theory that individual reasoning ability, as measured using standard reasoning tests, has part of its origin in dialogue with others. In the study 64 eight and nine year old children were taught the use of 'Exploratory talk', a type of talk in which joint reasoning is made explicit. The relationship between the talk of the children and the solving of Raven's test problems was followed using discourse analysis of groups working together. The findings of the study support four claims: that use of exploratory talk can improve group reasoning, that exploratory talk can be taught, that the teaching of exploratory talk can successfully transfer between educational contexts and that individual results on a standard non-verbal reasoning test significantly improved as a result of the intervention teaching exploratory talk. Our results offer support to the hypothesis that experience of social reasoning can improve scores on measures of individual reasoning. The stronger hypothesis that general cognitive development is a product of induction into social reasoning is not supported.
The growing influence of the socio-cultural paradigm has led to several studies of how children and others learn to think in particular ways through being inducted into social practices and ways of using language (see, for example, Rogoff, 1990; Resnick, Levine & Teasley, 1991; Mercer, 1995; Wertsch, 1991). These studies of specific cognitive development raise a question about the possibility of studying cognitive development in general. If learning to think is always induction into a specific social context, then how are we now to understand the findings of non-culturally based research on cognitive development? Gauvain, writing from a socio-cultural perspective, claims that this question poses an important challenge to researchers in the field of learning and instruction:

Although the extent to which cultural considerations can be incorporated into traditional views of development is unclear, it would be foolhardy to dismiss a long tradition of careful research rooted in such traditions simply because it did not take into account cultural influences on development. Thus a difficult task for the future is the reconciliation of findings from non-culturally based research with culturally based investigations of the same and related phenomena (Gauvain, 1995, p. 42)

In this paper we present a study that, in its theoretical basis, its methods and also its findings, goes some way towards constructing a bridge between a socio-cultural perspective on learning to think and the non-culturally based tradition of research on cognitive development that Gauvain refers to.

The study is based on the argument that a type of language use which (following Barnes & Todd, 1977) we call 'exploratory talk' embodies the kind of reasoning which is valued in a range of 'educated' cultural activities. In the study children were taught to use this kind of talk and the effects of this experience on their joint activity and their ability to solve the problems of a standardised psychological test of non-verbal reasoning (in groups and as individuals) were assessed. The test used - the Raven's Progressive Matrices - is commonly held to test 'the ability to reason and solve problems involving new information' (Carpenter, Just & Shell, 1990, p.404), while also correlating with measures of academic attainment. Analysis of discourse enabled us to relate changes in the way that children talked together while jointly solving the problems of one version of the Raven's test to changes in their individual scores on a second version of the same test. This methodology, combining qualitative discourse analysis with quantitative measurement and controlled experiment, enabled us to explore connections between cultural practice, social interaction and individual development.

**Theoretical framework**
The empirical study was designed to explore a model of individual cognitive development that combines a dialogical description of reasoning with a version of Vygotsky's account of individual development.

Resnick, Salmon, Zeitz, Wathen, & Holowchak. (1993) argue that whereas traditional psychology has described reasoning in terms of logical rules or other formalisms, various trends and arguments in contemporary cognitive psychology point to the need to see reason as a form of social practice. Elsewhere we have argued, influenced by Habermas's concept of communicative rationality (Habermas, 1990, p. 89; White, 1988; Habermas, 1995) and the arguments of other philosophers and communication theorists (for example Rorty, 1991, p. 39; Burbules & Rice, 1992), that to describe reason as a social practice requires a description in terms of inter-personal orientations and associated ground rules (Wegerif & Mercer, 1997a). By labelling our description of reason 'dialogical' we mean that it is not a model of reason drawn from the outside, after the event, as if reason was a closed and finished system, but it is a description of those rules and orientations which inform a type of dialogue from within, specifically those rules and orientations which serve to maintain a free and open encounter between different perspectives and ideas (see Rommetveit, 1992, for further explication of the significance of the move from 'monological' to 'dialogical' accounts of cognition).

Vygotsky's model of individual development, which has had a strong influence on the current socio-cultural perspective, stresses that 'all that is internal in the higher mental functions was at one time external' (Vygotsky, 1991, p36). Vygotsky's categories of internal and external have proved problematic for the contemporary socio-cultural perspective. Nonetheless Vygotsky's claim that an individual's ability to perform cognitive tasks when acting alone presupposes and stems from a prior socialisation process is still a basic tenet of the socio-cultural perspective. Some authors, following Leont'ev (1981), refer to the movement of development which Vygotsky called 'internalisation' as a process of the personal appropriation of cultural capital that results from a period of 'guided participation' or 'cognitive apprenticeship' (Rogoff, 1990; Newman, Griffin & Cole, 1989; Rojas-Drummond, Hernandez, Velez & Villagran, 1998).

The dialogical account of reasoning outlined above implies that reasoning is embedded in a social practice. Neo-Vygotskian accounts of individual cognitive development focus on induction into specific social practices. Combining these two views, a dialogical view of reasoning and a neo-Vygotskian view of development, would lead to the conclusion that learning to reason is essentially induction into a social practice. This model may be an oversimplification but as it is an extrapolation of trends in contemporary theory it is worth articulating and evaluating. In the strong form presented here it is a possible socio-cultural alternative to models of general cognitive development which are central to the non-cultural tradition of developmental psychology. Some of the conceptual issues...
Exploratory talk and reasoning

In order to apply and test the theoretical framework we have outlined, the notion of reasoning as a social process has to be specified rather more precisely, in terms of actual situated social practice, than has been done by philosophers such as Habermas and Rorty. To do this we have used the concept of 'exploratory talk' whose origin lies in empirical studies of classroom discourse.

In an article in Learning and Instruction, Mercer (1996a) used observational research in British primary schools to typify three kinds of talk, which he also described as representing different 'social modes of thinking'.

1. The first way of talking is Disputational talk, which is characterised by disagreement and individualised decision making. There are few attempts to pool resources, or to offer constructive criticism of suggestions. [...] Disputational talk also has some characteristic discourse features. notably short exchanges consisting of assertions and counter-assertions.

2. Next there is Cumulative talk, in which speakers build positively but uncritically on what the other has said. Partners use talk to construct a "common knowledge" by accumulation. Cumulative discourse is characterised by repetitions, confirmations and elaborations. [...]

3. Exploratory talk occurs when partners engage critically but constructively with each other's ideas [...] Statements and suggestions are offered for joint consideration. These may be challenged and counter-challenged, but challenges are justified and alternative hypotheses are offered. Compared with the other two types, in exploratory talk knowledge is made more publicly accountable and reasoning is more visible in the talk. Progress then emerges from the eventual joint agreement reached.

(Mercer, 1996a, pp. 368-369)

These types of talk are discussed and illustrated in detail elsewhere (Mercer, 1995, 1996; Wegerif & Mercer, 1997a). Of the three types we claim that exploratory talk is the closest to reasoning as a social practice.

The more detailed elaboration of exploratory talk which follows stems from three influences: conceptual considerations, particular those raised by Habermas (see Wegerif, 1996), the literature on research on effective collaborative learning (see review in Mercer 1995, p. 90-95) and our experience in classrooms working closely with teachers (see Dawes, 1997). Out of this combination of sources the following pragmatic ground rules for exploratory talk are provisionally proposed:

1. all relevant information is shared
2. the group seeks to reach agreement
3. the group takes responsibility for decisions
reasons are expected
challenges are acceptable
alternatives are discussed before a decision is taken
all in the group are encouraged to speak by other group members

The first three rules in the list are ground rules that are shared with cumulative talk, rules that serve to bind the group, share information together and construct knowledge together through seeking agreement. Rules four and five focus on the explicit reasoning that characterises exploratory talk as opposed to other types of talk. The role of challenges is important in distinguishing between cumulative, disputational and exploratory orientations. In exploratory talk challenges stimulate joint reasoning, in cumulative talk they are experienced as disruptive and often lead to a loss of cooperation and a switch into disputational talk. In disputational talk participants may still offer apparent arguments but are in fact focusing on 'winning' rather than on understanding or solving a problem together.

Ground rule six, that alternatives are discussed, reflects the findings of research on collaborative problem solving, particularly that of Kruger (1993) which has found that groups which do best are those which consider alternatives before deciding. In contradistinction to some researchers (e.g. Howe, 1992) we argue that this generation of alternative views does not necessarily imply different initial conception of the problem by the participants in collaboration but can itself be generated by the ground-rules of the talk. Finally rule seven was a product of empirical experience working with groups of children. We found that offering the abstract right to participate, found for example, in Habermas’s characterisation of the ideal speech situation (1991, p. 87) was not sufficient. In practice children needed to be actively encouraged to speak and to put forward views by their peers. These ground rules again emphasise our focus on the generative power of the interaction as opposed to an emphasis on the prior dispositions and views of the participants.

'Exploratory talk', in which reasoning is made visible and publicly accountable through the discussion of alternatives, offers us an empirically grounded version of what Habermas calls 'communicative rationality'. Engaging in this type of talk actively constructs participants as 'reasonable', (BenHabib, 1992) that is both giving reasons for claims and being open and responsive to the reasoning of others. While communicative rationality of this sort is historically and culturally situated it is nonetheless valued across a wide range of contemporary contexts. The 'exploratory talk' found and promoted in the classroom situation is a version of a type of language use given central importance in contemporary cultural activities such as science, law, government and the negotiation of business.

Aims of the study
The study reported had two broad aims, one to do with issues of practical pedagogy the other to do with theory. The first aim was to find out if the type of talk we are calling ‘exploratory talk’ can be taught and if this teaching can be successfully transferred from the teacher who originated the programme in her classroom to other teachers in different schools. The second aim was to explore the theory that performance on individual non-verbal reasoning tests is connected to prior participation in social reasoning as embodied in ‘exploratory talk’. This theory generates the testable hypothesis that the effective coaching of exploratory talk will increase children's individual results on established tests of reasoning. In order to test this hypothesis it is not enough to show a connection between coaching exploratory talk and individual test results, it is also necessary to show, in so far as this is possible, firstly that the ground rules of exploratory talk were instrumental in helping to solve reasoning test problems when children worked in groups and secondly that the coaching was effective in leading to the production of more exploratory talk.

The main hypotheses explored in this study are:

1. that the ground rules of ‘exploratory talk help groups to solve problems working together on group reasoning exercises;

2. that the incidence of talk showing the features of exploratory talk can be increased by the use of specially-designed teacher-led and peer-group activities;

3. that activities designed to teach exploratory talk which have been developed and implemented in one school by one dedicated teacher can transfer effectively to further schools and further teachers;

4. that increase in the use of exploratory talk in group exercises leads to increased scores by individual pupils working alone on reasoning tests.

In addition the study as a whole explored whether coaching in exploratory improved collaborative learning within the normal curriculum. This aspect of the study is not the focus of this paper and is commented on more fully in Wegerif, Mercer and Dawes, 1998.

Methods

Phases of the study

The empirical study had two phases. In the first phase an intervention programme to coach exploratory talk was developed and implemented in one base school. In the second phase this intervention programme was implemented in classes in two further schools.

Experimental design
Each run of the intervention programme in a school was treated as a separate field experiment. For each intervention:

- each target class was matched with a control class of the same age group in another local state school;
- target and control classes were divided into groups of 3, and in those groups the children attempted to solve the Raven’s Standard Progressive Matrices (SPM - see illustration in figure 1). They were given this test before the intervention and then again after the intervention had been completed in the target school;
- members of both target and control classes were given Raven’s Coloured Progressive Matrices (CPM), which is made up of similar non-verbal problems, as individuals both before and after the intervention programme;
- one focal group of three children in each class was video-taped while doing the group reasoning tests. (Other video-recordings of classroom events were also made, but the analysis of that data will not be discussed in this paper.)

Participants

60 children, aged 9 and 10 years old, in three local state middle schools composed the target classes that took part in the TRAC implementation programme. As mentioned above, each target class in each school was matched with a control class of the same age in another local state middle school. This produced three control classes of 64 children in all, who had no part of the intervention programme but who were observed and tested in the same ways as the target classes. (Because of absences on the days of testing, a few children initially included as subjects in both control and target classes dropped out of the study.) For the purpose of the tests and many of the exercises in the intervention programme the children were required to work together in mixed gender groups of three (plus one or two groups of two if numbers did not divide by three). These groups were organised by the class teacher, so as to include a range of ability in each. There were in total 23 target groups and 25 control groups. (Because of absences, three groups in target classes and four in control classes varied slightly in membership between the beginning and end of the study.)

The intervention programme

The TRAC intervention programme will be described only briefly here. A teachers’ guide is being prepared (Dawes, Mercer and Wegerif, in preparation) and more detail about the content of the programme is provided elsewhere by its main originator (Dawes, 1995, 1997). The programme
consists of a series of nine lessons. Each lesson is designed to last for about one hour and focuses on one or more of the ground rules of exploratory talk which were outlined earlier. The first few lessons deal with skills such as listening, sharing information and co-operating, while later lessons encourage children to make critical arguments for and against different cases. The children are given opportunities to practice discussing alternative ideas, giving and asking for reasons and ensuring that all members of the group are invited to contribute. Some computer-based group activities are included, using specially-designed software.

The use of the ground rules was taught to the children through explicit modelling by the teacher, coaching their use in whole group and small group discussions and giving opportunities for their use by the children working in small groups without the teacher. The explicit modelling phase involved the teacher at the front of the class illustrating the ways in which she wanted the children to talk together. Asking ‘why?’, using ‘because’ to give reasons for statements, asking other children what they think, reaching agreement before making a final decision. Different teachers modelled the language of exploratory talk differently. Some of these differences were picked up through analysing transcripts of the talk of the children. A key lesson in the programme involved eliciting the ground rules for the children in their own words. This was the third lesson after the children had all had some practice in collaborative activities. In a guided discussion the teacher drew from the class the kind of rules that they think should be used in group work. The list that resulted was then put on the wall in large letters. In the rest of the programme and in other lessons where collaborative learning was used the teacher or children can then refer to the rules on the wall and say ‘remember our rules’. Although each of the three classes studied produced a different set of ground rules they were all similar to the ground rules for exploratory talk which we listed earlier. Here is an example of the ground rules for one of the classes in the study:

**Class 5D’s Ground rules for talk**

1. Discuss things together. That means
   * ask everyone for their opinion
   * ask for reasons why
   * listen to people.
2. Be prepared to change your mind.
3. Think before you speak.
4. Respect other people’s ideas- don’t just use your own.
5. Share all the ideas and information you have.
6. Make sure the group agrees after talking.
Discourse analysis

In each of the four target classes, one 'focal group' of children was video-recorded as they worked on the Raven's SPM test, before and after the intervention. All the groups were arranged by the class teacher to be both mixed ability and mixed gender. The teachers were asked to select a focal group that was typical of the other groups in the class. The results of these groups on the group Raven's test were never the best results in the class but were above average for two of the three classes. From our observation of their talk together we hypothesise that the extra attention of having a camera focused upon them may have encouraged them to treat the task more carefully then they might have done without the camera.

The video-tapes of these focal groups were used for qualitative discourse analysis of the kind developed and described by the present authors and associates (Edwards & Mercer, 1987; Mercer & Fisher, 1993; Mercer, 1996b; Wegerif & Mercer 1997b). In addition transcripts of the talk of the children were made and a computerised text analysis concordancer was used to search for pre- and post-intervention differences. This method (described in more detail in Wegerif and Mercer, 1997b) involves integrating qualitative analysis of the full transcript with the abstraction of 'key words in context' in order to generalise significant features and compare different transcripts. Although essentially qualitative this approach also facilitates linking qualitative evaluation to quantitative descriptions of texts. For example detailed qualitative analysis of the deliberations of all the groups revealed that the word 'think', as in 'I think', was being used to put forward reasons. This word was then included in a computerised search for key features in the talk of the focal groups that indicated exploratory talk and so could be used to provide a quantitative measure for comparison with other transcripts.

Raven's Progressive Matrices

Raven's Progressive Matrices (RPM) consist of graphical puzzles of a type illustrated in figure 1. They are widely used in education and psychology as a test of 'non-verbal' reasoning. Raven's tests are particularly appropriate for exploring the link between language practices and the non-culturally based tradition of research in cognitive development, because they correlate well with other similar tests of reasoning and with measures of academic achievement (Raven, Court & Raven, 1995, Richardson, 1992, p. 129). Carpenter, Just & Shell write:

The centrality of the Raven test indicates not only that it is a good measure of intelligence, but also that a theory of the processing in the Raven test should account for a good deal of the reasoning in the other tests (Carpenter, Just & Shell, 1990, p. 428)
Richardson (1992) has argued that RPM do not measure abstract mental processes but rather the ability to read a particular kind of representation. He has demonstrated that if the same logical problems as those found in Raven's tests are presented in a way that makes more 'human sense', using pictures of cars and teddy bears for example instead of abstract shapes, children respond very differently with a different distribution of test-scores. We do not disagree with these criticisms of the way that Raven's test results, and similar tests, have been interpreted. Our interest in Raven's is as an index of a valued kind of cognitive ability in a particular cultural context. Our study uses RPM to investigate the relationship between this valued individual cognitive ability and the way children talk together.

We used two similar Raven's tests, the Standard Progressive Matrices (SPM), consisting of 60 problems, and the Coloured Progressive Matrices (CPM) consisting of 36 similar problems. The CPM is appropriate for younger children (up to the age of 12) while the SPM is appropriate for any age. Test scores from CPM and SPM can be translated into a common scale through the use of a table provided (Raven, Court & Raven, 1995, p. 64). We used the SPM for children working together in groups of three, giving each group a single book and a single answer sheet and encouraging them to talk together in solving the problems. We also gave the same children the CPM exactly three days later, asking them to work individually and following the guidelines for the administration of the test given in the manual (Raven, Court & Raven, 1995).

Results

Different approaches were used to investigate the different hypotheses. Our results will therefore be presented in relation to each hypothesis and the methods involved in testing it. Although the main methods used have been described above this results section will also refer to the design of each of these three investigations. Because of the methodological pluralism of the study, rigid divisions between method, results and discussion would not help achieve clarity in our presentation of the findings.

1. Evaluating exploratory talk

Our first hypothesis was that following the ground rules of 'exploratory talk' would help groups of children to jointly solve reasoning problems. The best way to assess this hypothesis is through a detailed analysis of the connection between the type of talk children used and their solving of reasoning test problems.

We selected the target group that had achieved the greatest pre/post-intervention change in group score on the Raven's SPM test to be the focus of this study. The group score had increased from 39 to 47 (the SPM has 60 items and so is measured on a 60 point scale). A comparison between this group score and the scores of individual members on the other version of the same test, the
Raven's CPM (which has 36 items and so is scored out of 36) was made by converting both scores to a common 36 point interval scale, following the procedure described in the Raven's test manual (Raven, Court & Raven, 1995 p. 64). In the pre-intervention tests the group score was lower than the highest individual score (31 to 32). In the post-intervention tests, however, the group score was slightly higher than the highest individual score in the group (34 compared to 33). This suggested that the striking improvement in group score after the intervention could not be accounted for by a change in the quality of reasoning of one individual in the group, but was a product of a change in the way the group reasoned together. To investigate this hypothesis further, we focused our discourse analysis on the talk of the group when they dealt with eight problems of the Raven's SPM test. These were problems that had been answered incorrectly in their pre-intervention attempt, but had been correctly solved by the group after the intervention. Here, to illustrate our analysis, are examples of the group's pre-intervention talk (which failed to produce the correct answer) and also of their post-intervention talk (which succeeded in finding the correct answer) when dealing with the same problem (B12, shown in Figure 1 below).

Figure 1 Problem B12. Raven's SPM
Transcript 1: pre-intervention talk in Group 1 School A1 on Raven's SPM problem B12

(Note: for the sake of intelligibility, punctuation has been added to all transcripts. Contextual information is presented in parentheses)

George: B12
Susan: We haven't done that.
Trisha: (giggles) this is where your Mum can see what you're really like at school George.
Trisha: Square and diamond, it's 2
George: No it's not
Trisha: It is 2
George: No it's not
Trisha: It is
George: No it's not
Susan: It's that one 6
Trisha: It is
George: No it's not it's got to be a square and a circle
Trisha: Its that, it has to be that, it has to be that, it has to be 6 because look they've only got that (pointing to the pictures)
Susan: Look first they are starting with one of them things over (pointing) and then it has to be black
George: Right, 6
Susan: No it isn't George
Trisha: That's number 2 because it goes bigger and bigger and bigger (Trisha is looking at the next question)
Susan: It isn't George look at that one, no it isn't George
(\textit{George is writing '6', which is the wrong answer. Susan pushes him})
George: Susan! All right someone else be the writer then
Susan: Me (forcibly takes paper)
Trisha: No, Susan you have to sit in this chair to be the writer (Trisha takes pen)
Trisha: Give it to George because he's quicker at it
Susan: Give me the pen \textit{(takes the pen)}
Susan: Are we finished on that one, are we on that one now
Trisha: You're not allowed to do it
George: Let Trisha have a go when we get up to E
(Turn to next exercise)
Susan: Well what do you think it is you dur brain? \textit{(addressed to George)}

\textbf{Transcript 2: post-intervention talk in Group 1 School A1 on Raven's SPM problem B12}

Susan: Set C
Trisha: No we ain't done B
Trisha: That has got to be a diamond, a square with a diamond with a circle in that one, number 6, do you agree?
George: No, what do you mean?
Trisha: OK no it's got to be square
Susan: I think it's number 6, that's the one
George: No it ain't
Susan: I think it's number 6
Trisha: No 'cause it's got to swing round every time, so there is a circle in it
Susan: Yes but it hasn't got a circle in there has it and that one has \textit{(indicating)}
(3 sec pause. Concentrated faces)
George: I think it's number ...
Trisha: I think it's number 4 to be honest
George: I don't I think it's number 6
Trisha: I don't, I think it's number 3 look because that one \textit{(pointing)} has that in the middle and it's got a half one in the middle
George: Complicated ain't it?
Susan: No because that one is that, I think it's that one
George: No because look at that and look at that (pointing) they are the same, you can't have two the same and it's got that one on, look Sue, it's got that one on and it's got that one on so its out of them three.

Susan: That one, one, 'cause that's a ..

George: Yes but it's got to be that

Susan: It's that because look that's got a square so it's just got to be empty

George: With no circle in so it's just got to be an empty square

Susan: No they are just normal boxes

Trisha: Look that's got a triangle, that's got a square, look that's got a square with a diamond with a circle in, that's got a square with a diamond in and that's got a square with a circle in so that's got to be a square

George: I don't understand this at all

Trisha: Because look on that they've taken the circle out yes? So on that you are going to take the circle out because they have taken the circle out of that one

George: On this they have taken the circle out and on this they have taken the diamond out and on this they have put them both in, so it should be a blank square because look it goes circle square

Susan: It’s got to be a blank square. Yeah it is. Mrs Dawes is coming. (inaudible whisper)

George: Do you agree on number 5, do you agree on 5?

(george writes '5', which is the correct answer)

George: Who is doing C? Susan, right let's have these here, C1

Commentary

In the pre-intervention talk of Transcript 1, George challenges Trisha's first suggestion ('It is 2') without giving a reason. Trisha offers no further justification for her suggestion. This leads into a series of exchanges typical of the type of talk we call 'disputational', in which participants simply assert their opposed views without reasoning. Susan then suggests 'It is that one 6' and this is taken up by Trisha, and both she and Susan offer reasons. '6' is apparently agreed upon, and George writes it down. However, Susan then appears to change her mind without saying what her new opinion is (or she may be objecting to him writing the answer down before checking properly with her and Trisha: no reason is made explicit). There is then a dispute about who should be writing the answers on the answer sheet.

Transcript 2 illustrates some ways that the talk of the same children changed after doing the TRAC programme. Compared with their pre-intervention talk, there are more long turns at talk, as more elaborate explanations are given. Again, Trisha is the first to propose an answer, but this time she does this not as a statement ('It is 2') but as an elaborated hypothesis with a question encouraging debate ('That has got to be a diamond, a square with a diamond with a circle in that one, number 6,
do you agree?’). George asks for more explanation. Attempting to be explicit, Trisha appears to see that she is wrong and changes her claim. George and Susan again engage in a ‘disputational’ exchange but this is short-lived. After a pause (for thought?) the children return to using language to think explicitly together about the problem. They come to agree that it is a kind of subtraction problem, and so find the correct answer.

Many more of the essential features of exploratory talk - as represented by our ‘ground rules’ - are evident in the post-intervention talk than in the pre-intervention talk. Explicit reasons for claims are given, challenges are offered with reasons, several alternatives are considered before a decision is reached, and the children can be seen seeking to reach agreement together. Explicit reasoning may be represented in talk by the incidence of some specific linguistic forms, and we can see here some ‘key features’: the hypothetical nature of claims is indicated by a preceding ‘I think’, reasons are linked to claims by the use of ‘because’ or ‘cause’ and agreement is sought through the question ‘do you agree?’. Explicit reasoning requires the linking of clauses and leads here to the incidence of more longer utterances in the post intervention talk. This group solved a total of eight new problems in the post-test which they had failed to solve in the pre-test. When we compare talk that led to the group solving these problems correctly and with talk which led to wrong answers, we find that there is a clear association with the relative incidence of these key linguistic features. This can be seen from Table 1 (below) which compares the number of long utterances (where ‘long’ is defined through taking an arbitrary cut-off point of being 100 characters in length or more when transcribed), and the incidence of ‘because’, ‘agree’ and ‘I think’.

<table>
<thead>
<tr>
<th>Key linguistic feature</th>
<th>Incidence in talk leading to incorrect answers</th>
<th>Incidence in talk leading to correct answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>long turns at talk</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>‘Because’ and ‘cause’</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>‘I think’</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>‘agree’</td>
<td>3</td>
<td>18</td>
</tr>
</tbody>
</table>

This specification of some of the differences between successful talk and unsuccessful talk can be used to provide us with an instrument to measure, if only in a very approximate way, features indicative of the effectiveness of the talk of the children. This is applied in the next section where
we look at the evaluation of the effectiveness of the intervention programme in changing the talk of the children.

2. Evaluating the intervention programme

Our second hypothesis was that the incidence of exploratory talk could be increased by the use of specially-designed teacher-led and peer-group activities, and our third was that the effective teaching of exploratory talk could transfer from the school where the programme originated with a dedicated teacher to other schools. These claims for the programme were tested by performing a computer-based discourse analysis of the children’s talk while they worked together on the Raven's SPM test, before and after the TRAC intervention. As is explained below, it was also tested through comparisons of group test scores before and after the intervention.

A method of discourse analysis using computer-based concordancer was applied to exploring the differences between transcripts of the talk of focal groups of children in the four target classes working on the same standard task (Ravens SPM) both before and after the intervention.

Above we saw from an extract of transcript that the exploratory talk which helped the children to solve the problem led to certain indicative linguistic features, those given in Table 1. Elsewhere detailed analysis revealed that modals, 'would', 'could', 'should' and 'might' were sometimes being used, as well as because, to introduce reasoning and so these were included in the list of indicative key words. Generalising these terms to all the transcripts collected of the talk of the focal groups produced the results given in Table 2.

Table 2 Incidence of key linguistic features in the talk of all target focal groups

<table>
<thead>
<tr>
<th>Key feature</th>
<th>School A1 Group 1</th>
<th>School A1 Group 2</th>
<th>School B1 Group 1</th>
<th>School C1 Group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>because</td>
<td>25</td>
<td>100</td>
<td>53</td>
<td>40</td>
</tr>
<tr>
<td>agree</td>
<td>7</td>
<td>87</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>I think</td>
<td>7</td>
<td>87</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>would</td>
<td>1</td>
<td>15</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>should</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>could</td>
<td>2</td>
<td>14</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>might</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Totals</td>
<td>44</td>
<td>306</td>
<td>83</td>
<td>110</td>
</tr>
</tbody>
</table>

Table 3 Incidence of key linguistic features in the talk of all control focal groups

<table>
<thead>
<tr>
<th>Key feature</th>
<th>School A2 Group 1</th>
<th>School A2 Group 2</th>
<th>School B2 Group 1</th>
<th>School C2 Group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>because</td>
<td>28</td>
<td>15</td>
<td>34</td>
<td>25</td>
</tr>
<tr>
<td>agree</td>
<td>15</td>
<td>6</td>
<td>12</td>
<td>20</td>
</tr>
</tbody>
</table>
I think would should could might
 35 18 27 44 3 5 1 3
 7 3 3 1 3 2 5 2
 15 3 16 2 0 0 0 2
 6 9 7 0 9 6 13 4
 3 2 15 1 3 0 2 1
 109 56 114 93 47 30 51 36

Table 4 Incidence of long utterances for all focal groups

<table>
<thead>
<tr>
<th>Condition</th>
<th>School A 1&amp;2 Group 1</th>
<th>School A 1&amp;2 Group 2</th>
<th>School B 1&amp;2 Group 1</th>
<th>School C 1&amp;2 Group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>Target</td>
<td>1</td>
<td>35</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td>Control</td>
<td>13</td>
<td>6</td>
<td>25</td>
<td>16</td>
</tr>
</tbody>
</table>

It can be seen from Table 2 that, overall, key features in the talk of the target focal groups increased after the intervention. The same increase was not found in the talk of the control focal groups shown in Table 3. A two way analysis of variance, with repeated measures on pre- and post- intervention counts, was performed. This showed that the difference between the target and control condition changes in the total count of key features was statistically significant ($F = 3.9$ one-tailed $p = 0.048$). The same analysis on the number of long utterances (Table 4) also showed a significant result ($F = 8.9$ one-tailed $p = 0.012$).

One can see that the increase varies across the schools, with an especially large shift in School A Group 1, and a very small shift in School C Group 1. Key features also have different patterns of incidence in the talk of each group (so for example in School A1 there is a high post-intervention rate of use of ‘agree’ by School A Group 1 but not for School A Group 2). On the basis of our observation of the various teachers carrying out the TRAC programme, we speculate that these differences are at least in part due to teachers placing more or less emphasis on the use of different kinds of language. In School A, class 1, the teacher stressed the importance of reaching agreement and modelled the use of the question: ‘do you agree?’. Our analysis suggests that the teaching in School A, class 2 might have lacked this element.

The use of key terms by the control groups reduced over the period of the intervention. This reduction was not statistically significant. We have no explanation for it. On possible hypothesis is that these groups found the task and situation less engaging the second time around than the first. If this is true it would have implications for the scores on the group Raven’s tests. This unexplained reduction in the use of key terms by the control groups means that our claim to statistical significance for an increased use of key terms by target groups needs to be treated with caution.

It is also noticeable that the control groups began the intervention with a greater use of indicators of exploratory talk than the target groups. These indicators are only a rough guide to the exploratory talk taking place. However this difference in the amount of indicators of exploratory talk was
reflected in higher scores on the group Raven's test in the pre-test for the control focal groups than for the target focal groups.

## Group test scores

### Table 5 Group scores on the Raven's SPM test.

<table>
<thead>
<tr>
<th>School</th>
<th>Number of groups</th>
<th>Mean Pre-intervention score (out of 60)</th>
<th>Mean Post-intervention score (out of 60)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 (target)</td>
<td>12</td>
<td>38.92</td>
<td>44.00</td>
<td>5.08</td>
</tr>
<tr>
<td>A2 (control)</td>
<td>11</td>
<td>43.00</td>
<td>44.45</td>
<td>1.45</td>
</tr>
<tr>
<td>B1 (target)</td>
<td>5</td>
<td>42.40</td>
<td>46.80</td>
<td>4.40</td>
</tr>
<tr>
<td>B2 (control)</td>
<td>5</td>
<td>40.20</td>
<td>39.40</td>
<td>-0.80</td>
</tr>
<tr>
<td>C1 (target)</td>
<td>6</td>
<td>45.66</td>
<td>47.33</td>
<td>1.66</td>
</tr>
<tr>
<td>C2 (control)</td>
<td>9</td>
<td>43.77</td>
<td>46.22</td>
<td>2.40</td>
</tr>
<tr>
<td>Total target</td>
<td>23</td>
<td>41.43</td>
<td>45.48</td>
<td>4.05</td>
</tr>
<tr>
<td>Total control</td>
<td>25</td>
<td>42.72</td>
<td>44.08</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Overall the scores of the target groups increased more that the scores o the control groups which is in line with the hypothesis that coaching exploratory talk will improve scores on reasoning tests. However the control group for school C1 actually increased more than the target school. An analysis of covariance (with post-test scores as the dependent variable and pre-test scores as the covariate) indicates that the difference between the gains of the target and control groups was not significant (F= 2.43; one-tail p = 0.06).

3. Evaluating changes in individual performance

Table 6 gives the increase in individual test scores for both target and control classes. An analysis of covariation (again with post-test scores as the dependent variable and pre-test scores as the covariate) revealed that the gains made by the individual target class children were significantly greater than those made by children in control classes (F= 3.141; one-tail p = 0.04.)

### Table 6 Individual scores on the Raven's CPM test.

<table>
<thead>
<tr>
<th>School</th>
<th>Number of children</th>
<th>Mean Pre-intervention score (out of 36)</th>
<th>Mean Post-intervention score (out of 36)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 (target)</td>
<td>35</td>
<td>28.91</td>
<td>31.31</td>
<td>2.40</td>
</tr>
<tr>
<td>A2 (control)</td>
<td>31</td>
<td>31.39</td>
<td>32.89</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>B1 (target)</td>
<td>15</td>
<td>32.73</td>
<td>34.53</td>
<td>1.80</td>
</tr>
<tr>
<td>B2 (control)</td>
<td>14</td>
<td>31.28</td>
<td>32.21</td>
<td>0.93</td>
</tr>
<tr>
<td>C1 (target)</td>
<td>10</td>
<td>33.10</td>
<td>34.70</td>
<td>1.60</td>
</tr>
<tr>
<td>C2 (control)</td>
<td>19</td>
<td>32.86</td>
<td>33.95</td>
<td>1.09</td>
</tr>
<tr>
<td>Total target</td>
<td>60</td>
<td>30.56</td>
<td>32.68</td>
<td>2.11</td>
</tr>
<tr>
<td>Total control</td>
<td>64</td>
<td>31.87</td>
<td>32.89</td>
<td>1.01</td>
</tr>
</tbody>
</table>

**Discussion**

It might seem strange that the individual increases in reasoning test scores should prove significant when the group increases were not significant. The socio-cultural model of learning reasoning which we are exploring would predict that the teaching exploratory talk should have its first, more direct, impact on group reasoning, and then secondarily on individual reasoning. However the design of our study meant that there were approximately three times as many individuals as groups and this probably explains why the group results did not reach statistical significance while the individual results did. Although the increase in group scores was not significant there is some evidence to support the hypothesis. In the two schools where our observations suggested that the TRAC programme had been most carefully and comprehensively carried out (Schools A1 and B1), group scores on the Raven’s test increased by over 10%.

The purpose of this study was not to show that it is possible to increase Raven's test scores through a training programme in non-verbal reasoning skills. This possibility is well known and has been shown by other studies (for example Riding & Powell, 1985). The difference with our study is that the training programme focused entirely on language use, teaching the ground-rules of exploratory talk without the use of materials or exercises which were in any way similar to a Raven's test. Children in the target classes had no more experience of doing the Raven' test, or any similar kinds of non-verbal problem solving, than those in the control classes. The individual Raven's tests were carried out three days after the group Raven's tests in all cases: yet their individual test results improved more, to a statistically significant extent.

It has been suggested to us (by other researchers) that the statistically significant results of the target classes may reflect the appropriation by target children of strategies for solving non-verbal problems used in the group Raven's tests, rather than simply the appropriation of 'dialogic' ways of reasoning from the use of exploratory talk. However, even if this is the case, it does not necessarily undermine the support for our hypothesis. We have demonstrated that the strategies used by target children to share and account for problem-solving strategies in the group Raven's test did reflect the influence of the teaching programme. If individuals in the target classes were able to appropriate task specific strategies more effectively than those in the control classes, this is therefore likely to be because of the more effective way that they jointly constructed knowledge
when they talked together around the task. Whether the causation was direct or indirect, the results of the experimental study support the view that the statistically significant improvements in the individual test results were a result of our language-based intervention.

The TRAC intervention programme worked well in the school where it originated. Here it was implemented by an enthusiastic teacher-researcher who was committed to the programme and to the study. The transfer to two further schools proved problematic and raises issues for future research. If we judge by the group test scores (Table 3), the intervention programme appears to have had an impact on group process in school B while there is little evidence of an impact in school C. This conclusion was supported by our field notes and is also supported the indicators of exploratory talk (Table 2). This difference may be accounted for by the different motivation of the teachers or by the different pedagogic strategies they adopted. A planned follow-up study will focus more on teacher motivation and on the pedagogy. However another possibility is also suggested by the fact that the schools involved varied in the socio-economic background of their intake. School C and its control school serve a catchment area of predominantly affluent, middle class communities, while the other schools serve populations of lower average income where rates of unemployment are high. One possible explanation for the failure of the intervention programme in School C1 is that this kind of coaching has the most positive effect on children's joint attempts to solve the Raven's problems when the children involved are from lower socio-economic groups. That is, it is possible that the middle class children were already more familiar with the 'ground rules' of exploratory talk than those from the lower income families. This is supported by the observation that a number of individuals in Schools A1 and B1 who got a very low result on the pre-intervention test improved dramatically after coaching in exploratory talk suggesting that their initial failure was due to the lack of quite simple strategies which the intervention course was able to provide them with. Although this was an unexpected result it conforms to a large literature on the impact of socio-economic class on language acquisition and academic achievement, particularly the work of Bernstein (1992; 1996).

We have concentrated in this paper on the particular effects of the intervention on children's talk and reasoning while doing the Raven's test. However we were also concerned to explore the impact of teaching exploratory talk on the way children work together and reason together on tasks directly related to the school curriculum. To this end we investigated the impact of the TRAC intervention programme on children's talk and learning in classroom activities in the area of science and of citizenship. Here we found a similar improvement, from an educational point of view, in the quality of their interaction and their collaborative learning. These findings have been published separately (Wegerif, Mercer & Dawes, 1998).

Conclusion
The four main findings of the study are that children's use of exploratory talk can be increased through teaching, that the programme to teach exploratory talk can transfer from its place of origin to other schools (although this transfer can be difficult to achieve); that exploratory talk helps groups to reason together, and that individual scores on a standard non-verbal reasoning test can be significantly increased by teaching children to use exploratory talk.

We began this paper with the outline of a possible socio-cultural model of learning to reason as induction into the dialogical practice of reasoning together. The study described was designed partly to explore the extent to which this model of learning to reason could provide an account of general cognitive development. The results of our study indicate that, at least in part, the kind of reasoning ability involved in individual non-verbal reasoning tests is mediated by social interaction. This result is interesting because non-verbal reasoning tests have been taken by some to be the paradigmatic context for the assessment of individual reasoning ability independent of culture and language use. However the exact nature of any link between teaching the use of exploratory talk and individual test results needs further elucidation. One plausible hypothesis is that the group experience in the target classes prior to the individual post-intervention testing, in which knowledge was much more explicitly shared and accounted for than in the control classes, helped target class children appropriate specific problem-solving skills (such as considering alternative options before choosing an answer) and develop a greater awareness in general of the nature of the task. However, this interpretation does not undermine our socio-cultural claims of a relationship between the 'intermental' and the 'intramental' in learning and development, as it still supports the view that reasoning skills may be developed through social 'collective thinking' activity. The findings of the study are compatible with the socio-cultural position, that success at any cognitive task is a situated achievement in which many contextual factors play a part.

In the introduction Gauvain was quoted calling for the need to find routes to reconcile what she called 'traditional views of development' with a more culturally based approach. The findings of our study give modest support to the idea that we should place a traditional account of cognitive development within a historical and cultural context in which this account is seen, at least in part, to reflect a process of induction into certain valued ways of using language. The study was not detailed enough, and its findings were not unambiguous enough, for us to claim that they either confirm or deny the strong hypothesis that learning and using dialogical reasoning is the origin of what is described by some as the process of general cognitive development. However pursuing aspects of this hypothesis led us to findings which are highly relevant to these issues, and also led to the development of new methods which may be used in the future to explore further the relationship between language use and cognitive development.

Acknowledgements
We are grateful for the support of the Economic and Social Research Council (ref. R000221868), for the active cooperation of teachers in Milton Keynes, for Karen Littleton's contributions to the project and for Richard Joiner's help with the analysis of the data.

References


