Talking science: the research evidence on the use of small-group discussions in science teaching

Judith Bennett
Department of Educational Studies
The University of York, UK
This report should be cited as:
## Contents

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Why are small-group discussions seen as important?</td>
<td>1</td>
</tr>
<tr>
<td>Research on small-group discussions</td>
<td>2</td>
</tr>
<tr>
<td>What are small-group discussions?</td>
<td>3</td>
</tr>
<tr>
<td>What are systematic reviews?</td>
<td>3</td>
</tr>
<tr>
<td>The origins and aims of systematic reviews</td>
<td>3</td>
</tr>
<tr>
<td>Systematic review methods</td>
<td>3</td>
</tr>
<tr>
<td>The systematic reviews of studies on small-group discussions in science lessons</td>
<td>4</td>
</tr>
<tr>
<td>The review research question</td>
<td>4</td>
</tr>
<tr>
<td>The criteria for including studies</td>
<td>4</td>
</tr>
<tr>
<td>Characterising the studies</td>
<td>5</td>
</tr>
<tr>
<td>An overview (systematic map) of evaluation studies on small-group discussions</td>
<td>5</td>
</tr>
<tr>
<td>The in-depth reviews</td>
<td>8</td>
</tr>
<tr>
<td>The evidence on the use of small-group discussions</td>
<td>9</td>
</tr>
<tr>
<td>The evidence the effects of small-group discussions on understanding of evidence</td>
<td>15</td>
</tr>
<tr>
<td>The evidence on the effects of different stimulus materials on understanding of evidence</td>
<td>16</td>
</tr>
<tr>
<td>Conclusions and implications</td>
<td>17</td>
</tr>
<tr>
<td>Issues in research on small-group discussions in science teaching</td>
<td>17</td>
</tr>
<tr>
<td>A note on design experiments</td>
<td>18</td>
</tr>
<tr>
<td>Implications for further research</td>
<td>18</td>
</tr>
<tr>
<td>Implications for policy</td>
<td>18</td>
</tr>
<tr>
<td>Implications for practice</td>
<td>18</td>
</tr>
<tr>
<td>References</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List of tables, figures and boxes</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1: Nature of the stimulus</td>
<td>6</td>
</tr>
<tr>
<td>Table 1.2: Product of the discussion task</td>
<td>7</td>
</tr>
<tr>
<td>Table 1.3: Reported outcome</td>
<td>7</td>
</tr>
<tr>
<td>Table 1.4: Research strategy</td>
<td>8</td>
</tr>
<tr>
<td>Table 1.5: Nature of data gathered</td>
<td>8</td>
</tr>
<tr>
<td>Table 1.6: Overview of studies included in the review</td>
<td>10</td>
</tr>
<tr>
<td>Table 1.7: Details of the studies included in the review</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.1: The country of study</td>
<td>5</td>
</tr>
<tr>
<td>Figure 1.2: Subject focus</td>
<td>6</td>
</tr>
<tr>
<td>Figure 1.3: Group size</td>
<td>6</td>
</tr>
<tr>
<td>Figure 1.4: The nature of group</td>
<td>6</td>
</tr>
<tr>
<td>Figure 1.5 Length of discussion tasks</td>
<td>6</td>
</tr>
<tr>
<td>Box 1: Techniques advocated for small-group discussions</td>
<td>7</td>
</tr>
<tr>
<td>Box 2: Chief characteristics of research on small-group discussions</td>
<td>9</td>
</tr>
<tr>
<td>Box 3: Summary of key findings on the use of small-group discussions</td>
<td>15</td>
</tr>
<tr>
<td>Box 4: Summary of key findings on effects of small-group discussions on understanding of evidence</td>
<td>16</td>
</tr>
<tr>
<td>Box 5: Summary of key findings on effects of different stimulus materials on understanding of evidence</td>
<td>17</td>
</tr>
</tbody>
</table>

## Acknowledgements

Thanks are due to the following groups and individuals who contributed to the work presented in this report: the Evidence, Policy and Practice Initiative (EPPI)-Centre and the EPPI Review Group for Science at the University of York, particularly Sylvia Hogarth, Fred Lubben and Alison Robinson. The sponsorship of the National Science Learning Centre, through the White Rose Research into Practice awards, is also gratefully acknowledged.
Introduction

Why are small-group discussions being advocated in science teaching?

How are small-group discussions currently used in science teaching?

What are their effects on students’ understanding of science ideas?

How do students respond more generally to their use?

Many people involved in teaching and curriculum development in science believe that small-group discussions are an important tool in science teaching, motivating students and enhancing their learning in science. This booklet examines the research evidence on small-group discussions and their effects.

Aims

The research reviews forming the basis of this booklet had two principal aims:

- to identify the ways in which small-group discussions are currently used in science lessons;
- to look at the effects of small group discussions on students’ understanding of science and attitudes to science.

The term understanding has been taken to encompass science concepts, and ideas about the nature of science and the methods of science. The term attitude includes attitude towards science, attitude towards school science, motivation to learn, interest in science activities and career intentions.

The evidence presented has been gathered using the systematic review methods based on those developed as part of the Evidence, Policy and Practice Initiative (EPPI), a Government-sponsored project in the UK whose aim is to synthesis and disseminate research findings in key areas of education.

Why are small-group discussions seen as important?

The use of small-group discussions in teaching has been advocated for a number of years as one of a range of learner-centred teaching approaches or ‘active learning’ strategies felt to stimulate students’ interest in what they are studying by providing them with a significant degree of autonomy over the learning activity.

Several factors, summarised below, have contributed to the current high levels of interest in small-group discussion work in science. Some have emerged directly from research studies, whilst others appear to draw more loosely on research evidence and take the form of approaches which are being advocated in science teaching, but whose effects have yet to be explored on a more systematic basis.

The development of scientific literacy

The publication of Beyond 2000 (Millar and Osborne, 1998) stimulated discussion and debate over the nature of the school science curriculum and, in particular, the ways in which it might foster the development of scientific literacy. This term embraces the knowledge, understanding and skills young people need to develop in order to think and act appropriately on scientific matters which may affect their lives and the lives of other members of the local, national and global communities of which they are a part. There are now moves in a number of countries to introduce science curricula with a strong emphasis on scientific literacy.

A key aspect of scientific literacy is the ability to participate in informed discussion and debate of scientific issues, and this points to the need for including small-group discussions in the repertoire of activities employed in science lessons.
Ideas and evidence
An area related to the development of scientific literacy is that of ideas about evidence, which involves encouraging students to evaluate, interpret and analyse evidence from primary and secondary sources in science, including stories about how important science ideas were first developed, and then established and finally accepted. This has led to considerations of the role of argument in school science, in the sense of putting forward claims and supporting them with sound and persuasive evidence (Osborne et al., 2001). Small-group discussions have a key role to play here, since the practice of using evidence in argumentation requires interaction with peers.

Formative assessment
The area of formative assessment, or assessment for learning, is receiving considerable attention at present, to a large part resulting from the publication of Inside the Black Box (Black and Wiliam, 1998). Formative assessment relates to the assessment strategies and techniques that take place during teaching to establish progress and diagnose learning needs to support individual students. (This contrasts with summative assessment, which refers to the tests and examinations which take place at the end of courses or blocks of teaching.) A number of approaches have been advocated for increasing the use and effectiveness of formative assessment in science teaching, including the use of peer-review of work through small-group discussions (see, for example, Daws and Singh, 1999).

The constructivist view of learning
One of the most significant research programmes in science education has emerged from the constructivist viewpoint on learning, which has explored in depth the ideas and understanding students bring with them to science lessons and the ways in which some of their ideas may hinder the development of accepted scientific ideas (e.g. Driver et al., 1985). One of the recommendations for practice which has emerged from constructivist research is that small-group discussions should be used in science lessons as a means of helping students explore their ideas and move towards more scientific ideas and explanations. Further impetus for the inclusion of small-group discussions in science lessons has come from the development of ideas about social constructivism (Driver et al., 1994). These draw on the work of Vygotsky, who emphasises the importance of the social dynamics of interactions in fostering learning.

The development of literacy skills
There is a more general drive to improve students’ literacy skills and, in England and Wales, this has been formalised into the National Literacy Strategy (DfEE, 1998). Small-group discussions have been advocated as a means for developing students’ language skills in science (e.g. Newton et al., 1999, and Osborne et al., 2001).

Research on small-group discussions
There is a growing body of evidence that teachers would welcome support and guidance on running small-group discussions. In particular, evaluation of materials with a specific focus on teaching socio-scientific issues and developing scientific literacy, such as the new AS Public Understanding of Science course (Osborne et al., 2002) and the Valuable Lessons project (Levinson and Turner, 2001), established that teachers saw the provision of support and guidance on running small-group discussions as a priority. Whilst the ability to engage in discussion is seen as important, activities aimed at developing this ability are not well known to science teachers. Furthermore, the introduction of small-group discussions in science lessons challenges the established pedagogy of science teaching and places new demands on teachers.
What are small-group discussions?

Reading reports on small-group discussions, it became clear that the term was used in a wide variety of ways. For the purposes of the work reported here, a small-group discussion was taken to be an activity that:

• involves groups of two to six students;
• has a specific stimulus (e.g. a newspaper article, video clip, prepared curriculum materials);
• involves a substantive discussion task of at least two minutes;
• is either synchronous (i.e. face-to-face) or asynchronous (i.e. mainly IT-mediated);
• has a specific purpose (e.g. individual sense-making, or leading to an oral presentation, or to a written product).

What are systematic reviews?

The origins and aims of systematic reviews

Systematic reviews of research studies are a comparatively recent development in education, though they are well established in medical research. They have emerged from the international debate over the nature and purpose of educational research, and how it contributes to maximising the effectiveness of educational provision (e.g. Hargreaves, 1996 and Hillage et al., 1998, in the UK; Shavelson and Towne, 2001, in the USA).

There are several reasons why systematic reviews are being seen as a key strand in educational research. Firstly, there is a growing interest in practical policy-related decision making being linked to evidence in a number of areas, not just in education. Systematic reviews of research literature are seen as having the potential to yield evidence on which policy makers can draw. Secondly, there is a drive towards forging closer links between research, policy and practice. In particular, drawing on research findings in classroom practice is seen as desirable, with teachers being encouraged to engage in what is variously described as ‘evidence-based’, ‘evidence-informed’ or ‘evidence-enriched’ practice.

It was for these reasons that, in 2000 the Government in the UK funded, via the Department for Education and Skills (DFES), the Evidence for Policy and Practice Initiative (EPPI)-Centre to focus on systematic reviews of research evidence in key areas of education. The Centre is based in the Social Science Research Unit at the Institute of Education in London and works in partnership with Review Groups located around the UK. The Review Group for Science is located in the Department of Educational Studies at the University of York.

EPPI-Centre systematic reviews aim to produce high-quality reviews of research findings that provide evidence accessible to a range of different user groups, including teachers, researchers and policy-makers, and for this reason, Review Groups are constituted such that their membership includes policy-makers, teachers, advisory teachers, inspectors, academic researchers, teacher trainers and those involved in curriculum development work, including text-book authors.

An additional round of funding for systematic reviews (using EPPI-Centre methods) has been made available from the Government through the Training and Development Agency for schools (TDA), the group that regulates pre-service and some in-service teacher training in England and Wales.

Systematic review methods

The systemic review process, as developed by the EPPI-Centre, involves several stages:

• identification of review topic area;
• identification of review research question;
• development of inclusion and exclusion criteria for studies in the
review (relating to, for example, aspects such as the age of students, the nature of the research design, and the reported outcomes);

- undertaking of **systematic searches** of electronic data bases and other sources for potentially relevant research studies;
- coding or **keywording** studies against pre-specified and agreed characteristics (some of which are generic to all EPPI reviews, whilst others are developed specifically for each review);
- production of an overview or **systematic map** of studies in the review area, that groups the studies according to their chief characteristics;
- undertaking an **in-depth review** of studies to look in detail at their design and findings and to evaluate the quality of the work reported.

The in-depth review involves a process called **data extraction**, where information from the studies is extracted in a systematic way. Information extracted from the studies includes:

- study aims and rational;
- study research questions;
- study design methods, including selection of groups, sampling, consent of subjects;
- data collection methods;
- data analysis methods;
- reliability and validity of methods of data collection and analysis;
- results and conclusions;
- quality of reporting;
- quality of the study in relation to methods and data.

This information is then used to make judgements about the quality of the weight of evidence presented in the study in relation to the review research question.

Each of these judgements involves a decision about whether the weight of evidence in a study is **high**, **medium** or **low**.

Once the work of undertaking the review has been completed, it is written up as a **review research report**. The structure of these reports is prescribed by the EPPI-Centre so that all reviews are presented in a common format. Draft reports are sent to external referees for peer review, and, once finalised, made available in electronic form through the Research Evidence in Education Library (REEL), an open-access resource available through the EPPI-Centre website (http://eppi.ioe.ac.uk)

The work reported here also extended the EPPI review methods to draw on the guidance and framework for assessing research evidence in qualitative research studies (Spencer et al., 2003).

### The systematic review of studies on small-group discussions in science lessons

#### The review research question

The main review research question was: **How are small-group discussions used in science teaching with students aged 11-18, and what are their effects on students’ understanding in science or attitude to science?**

Within this, three reviews were conducted, focusing on:

- the nature of small-group discussions in science (Review 1);
- the effect of small-group discussions on students’ understanding of evidence in science (Review 2);
- the effect of different stimulus materials on understanding of evidence in science (Review 3).

#### The criteria for including studies

Studies were included in the review if:

- They were about the use of small-group discussions in science lessons.
- They involved groups of two to six students.
- They involved a substantive, structured discussion task of two minutes’ duration or more.
They illustrated how small-group discussions are being used. They addressed aspects of students’ understanding in science or attitudes to science. They focused on students in the 11-18 age range. They had been undertaken in the period 1980-2004. They were published in English.

Student age was restricted to 11-18 because this is where the main focus of interest lies. The start date for the period of publication was selected because this was the time when the use of small-group discussions was first advocated in science teaching.

The searches yielded some 2,290 studies, of which 94 met the inclusion criteria for the review.

Characterising the studies

These studies were then coded against particular characteristics (keyworded) to produce an overview (the systematic map) of evaluation studies on context-based/STS interventions. In producing the map, the following characteristics of studies were scrutinised:

- the country of study;
- the age/level of the students;
- the type of study;
- the science discipline of the study;
- constitution of discussion groups;
- duration of discussion tasks;
- stimulus for discussion tasks;
- product of discussion tasks;
- research strategy used;
- nature of data collected;
- outcomes reported.

An overview (systematic map) of studies on small-group discussions

The period of the review covered 1980 - 2003. 93% of the 94 studies that met the inclusion criteria were from post-1990, and 53% from post 1998, indicating that that the research activity area has been minimal up to fifteen years ago and has been most prolific from the late 1990s onwards.

Study focus

72% of the studies looked at effects on students’ understanding of science. A smaller number of studies (15%) reported on the effect of small-group-discussions on students’ attitude to science.

Country of study

Figure 1.1 shows the countries where studies were undertaken, with 39% of the studies undertaken in the USA, with 13% in the UK (13%) and 12% in Canada. In this context, it should be noted that the review was limited to papers published in English, but did include studies of small-group discussions held in Bahasa Malay, Cantonese, Dutch, Finnish, French, German, Greek, Hebrew, Mandarin, Portuguese and Spanish.

Figure 1.1: Country of study

The nature of the studies

Of the studies, 54% reported on evaluations of small-group discussions, of which 28% adopted some form of experimental design involving a control group, including 13% randomised controlled trials (RCTs). The remaining studies provided either descriptive information about small-group discussions, or explored of relationships between different characteristics of small-group discussions.

The science subject focus

Figure 1.2 shows the subject focus area, revealing that small group discussions were predominantly used in physics and biology teaching. The most likely explanation for this is that most of the small-group discussions
developing skills of decision-making on socio-scientific issues are, traditionally, placed within biology classes, and discussions relating to exploration of difficult ideas are located mainly within physics.

The nature of the groups

Over 90% of the studies focus on secondary school learners in mixed-sex groupings between 11-16 years of age. The majority of studies involved mixed ability classes, grouped in homogeneous ability discussion groups. Figure 1.3 shows group size, showing groups of 3-4 students to be the most frequently-used group size.

A variety of ways was used to constitute groups, as Figure 1.4 illustrates.

Where details were given, mixed-ability and friendship groups predominated. It seems likely that the percentage of same-sex groups will be higher than reported as friendship groups are more likely to be same-sex groups.

In 33% of cases, groups were deliberately constituted by the teacher, with students choosing their groups in the remaining cases. 94% of the studies concerned self-contained and permanent groups. The remaining studies drew on the techniques of ‘snowballing’, and ‘jigsawing’ – two of a range of techniques advocated for small-group discussions (see Box 1).

Length of discussion tasks

Figure 1.5 shows the duration of discussion tasks, indicating that two-thirds of the discussions took a class period or longer.

Stimulus for discussion tasks

Table 1.1 shows the nature of the stimulus material provided for the discussion tasks.

Table 1.1: Nature of the stimulus

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-line oral teacher instruction</td>
<td>2</td>
</tr>
<tr>
<td>Oral context provided by teacher</td>
<td>3</td>
</tr>
<tr>
<td>Newspaper article</td>
<td>1</td>
</tr>
<tr>
<td>Prepared curriculum print materials</td>
<td>66</td>
</tr>
<tr>
<td>Practical work</td>
<td>39</td>
</tr>
<tr>
<td>Computer software</td>
<td>27</td>
</tr>
<tr>
<td>Field trip</td>
<td>1</td>
</tr>
<tr>
<td>Video/TV/film clip</td>
<td>9</td>
</tr>
<tr>
<td>Learner generated</td>
<td>15</td>
</tr>
</tbody>
</table>
Box 1: Techniques advocated for small-group discussions

**Jigsawing:** Students are members of two different groups during the discussion task. The first is the ‘home’ group, in which students work in groups of four to six on some instructional material which has been broken down into sections. Each student in all the home groups is assigned a different portion of the material. The home groups then break apart and reform into ‘expert’ groups in which group members are all focusing on and discussing the same piece of the material to make sure they understand it. Once this has happened, students' groups then break once again and reform back into ‘home groups’ to peer-tutor the home group on the aspect of the material they have studied intensively, and learn from other home group members about the other aspects of the material.

**Envoying:** Students working in two groups. In the first group, they discuss a common task, which differs for each group. Groups then reform, with new groups containing one member of each of the original groups, who act as envoys to report on their particular task.

**Snowballing:** Pairs of students discuss a question or idea and agree on their views, then join with another pair to share what they have discussed, and then finally with another group of four (two pairs) to share thinking for a final time.

**Four corners:** The teacher chooses a topic and the students then brainstorm related sub-topics. Through a process of elimination, four topics are identified and one each is allocated to students grouped into the four corners of the room. The groups then choose a leader, a recorder and a reporter. The topics are discussed in the groups and the reporter then summarises them for the rest of the groups.

Discussion tasks in half of the studies used more than one type of stimulus. In two-thirds of the studies, discussions were based on curriculum print materials, usually a worksheet, a handout with text, specific problems to be solved or issues to be discussed. In more than one-third of the studies, the group discussions centred around practical work, and in just under a quarter of the studies the stimulus was computer software, used either interactively or passively.

The product of the discussion

Table 1.2 shows the intended product of the discussion task.

**Table 1.2: Product of the discussion task**

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Specific Product of the Discussion Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual sense-making</td>
<td>94</td>
</tr>
<tr>
<td>Report group views/presentation orally in class</td>
<td>22</td>
</tr>
<tr>
<td>Support a group position in a class debate/quiz</td>
<td>11</td>
</tr>
<tr>
<td>Present group written project (including poster)</td>
<td>12</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
</tr>
</tbody>
</table>

In a very high proportion of the studies, the product of the discussion task was individual understanding of the science underlying the activity, such as in a practical experiment, the preparation of a poster or a computer-based exercise, in which the learners were engaged. In the majority of cases this understanding was then shared with classmates in different ways: groups might present their findings or views orally or by way of posters or might defend their position in a whole class debate. Those products falling into the other category included either group or individual written reports or posters that were submitted to the teacher or researcher.

Outcomes reported in the studies

Table 1.3 summarises the outcomes reported in each study. These often included more than one aspect per study.

**Table 1.3: Reported outcome**

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Specific Outcome of the Discussion Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual understanding of science</td>
<td>73</td>
</tr>
<tr>
<td>Evidence (methods and nature of science)</td>
<td>33</td>
</tr>
<tr>
<td>Applications of science</td>
<td>3</td>
</tr>
<tr>
<td>Attitudes to (school) science</td>
<td>16</td>
</tr>
<tr>
<td>Skills (communication/collaboration)</td>
<td>61</td>
</tr>
<tr>
<td>Decision-making on socio-scientific issues</td>
<td>11</td>
</tr>
</tbody>
</table>
Nearly three-quarters (73%) of studies focused on the impact of the discussion tasks on the conceptual aspects of science understanding. Not surprisingly, well over half the studies (61%) focused on the communication itself and collaborative skills associated with the discussion tasks given to student groups. A small proportion of studies involved decision-making on socio-scientific issues and very few included aspects relating to the applications of science.

**Research strategy**

Table 1.4 shows the research strategy employed to gather data on the small-group discussions.

<table>
<thead>
<tr>
<th>Table 1.4: Research strategy</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>30</td>
</tr>
<tr>
<td>Survey</td>
<td>16</td>
</tr>
<tr>
<td>Case study</td>
<td>56</td>
</tr>
<tr>
<td>Action research</td>
<td>3</td>
</tr>
<tr>
<td>Ethnography</td>
<td>4</td>
</tr>
</tbody>
</table>

It is not surprising that more than half of the studies employed case studies, as a characteristic of work in the area is a desire to gather detailed information about the nature of discussions. One outcome of the case study approach and the very labour-intensive nature of much of the day collection and analysis is that sample sizes tended to be small – very often one class or one or two groups of students within a class. Studies involving several classes, or classes in more than one school, were comparatively rare.

Experimental approaches were also used in a number of cases, as studies were seeking to explore the effects of small group discussions compared with other approaches.

**The nature of the data gathered**

There was considerable diversity in the nature of the data gathered, as Table 1.5 shows. On average, the studies presented findings based on at least two different types of data.

<table>
<thead>
<tr>
<th>Table 1.5: Nature of data gathered</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test results</td>
<td>43</td>
</tr>
<tr>
<td>External examination results</td>
<td>1</td>
</tr>
<tr>
<td>Written reports/questionnaires</td>
<td>34</td>
</tr>
<tr>
<td>Audio-taped discussions</td>
<td>45</td>
</tr>
<tr>
<td>Video-taped discussions/observation</td>
<td>68</td>
</tr>
<tr>
<td>Likert-type instruments</td>
<td>3</td>
</tr>
<tr>
<td>Self-reports (diaries, interviews)</td>
<td>36</td>
</tr>
<tr>
<td>Presentations</td>
<td>1</td>
</tr>
<tr>
<td>Computer logs</td>
<td>16</td>
</tr>
<tr>
<td>Concept webs</td>
<td>5</td>
</tr>
</tbody>
</table>

The studies established that small-group discussions were being used in a variety of ways in science lessons. However, a characteristic of many of the studies was that small-group discussions tended to be wrapped up within other activities, often characterised as ‘collaborative learning’. This term was used in a variety of ways, often loosely, such that it appeared to include most activities which did not involve teacher exposition.

Box 2 summarises the chief characteristics of research on small-group discussions.

**The in-depth reviews**

Three in-depth reviews were conducted, covering a total of 25 of the studies in the systematic map:

- Review 1 focused on the nature of small-group discussions and was based on 19 of the 25 studies;
- Review 2 focused on the effects of small-group discussions on students’ understanding of evidence and was based on 14 of the 25 studies;
- Review 3 focused on the effects of different stimulus materials on students’ understanding of evidence and was based on 10 of the 25 studies.
Box 2: Chief characteristics of research on small-group discussions

- The majority of the studies report work that has taken place in the USA, the UK and Canada.
- Small-group discussions were used with all ages of students in the secondary age range.
- The majority of work focused on small-group discussions in relation to students’ understanding.
- A diversity of measures was used to assess effects on understanding and attitude.
- Very little research has been done on small-group discussions in relation to the teaching of chemistry.
- Typical small-group discussions involved groups of 3-4 students emerging from friendship ties, and last for at least 30 minutes.
- Typical small-group discussions had individual sense-making as their main aim (as opposed to, for example, leading to a group presentation) and use prepared printed materials as the stimulus for discussion.
- The most common research strategy was that of case study.
- 28 of the 94 studies reviewed used experimental designs.
- The most popular techniques for gathering data were observation, video- and audio-tapes of discussions, interviews, questionnaires and test results.

Table 1.6 provides an overview of the studies, indicating the country of study, which studies were used in each review, and the quality rating for each of the studies. Studies were rated as high (H), medium high (MH), medium (M), medium low (ML) or low (L). Because quality judgements about weight of evidence are made in relation to the focus of each of the reviews, some studies were given different ratings for different reviews.

The remaining discussion focuses on the twenty studies rated as being of medium quality or better, and Table 1.7 summarises the key features of each of these studies.

The evidence on the use of small-group discussions

Nineteen studies were included in the in-depth review, which focused on the nature of small-group discussion work aimed at improving students’ understanding of evidence.

The evidence from the review draws on the findings from fourteen studies weighted as medium-high and medium in overall quality. No studies were rated high in quality.

The review has revealed a number of features of particular interest in relation to the use of small-group discussion work in science. It is clear from the study reports that a complex and interacting set of factors are involved in enabling students to engage in dialogues in a way that could help them draw on evidence to articulate arguments and develop their understanding. Thus a particular characteristic of such studies is detailed description of student interactions.

Although the studies in the in-depth review shared a number of similar characteristics at the broad level, there are considerable differences at the detailed level. For example, there is considerable variety in the specific research questions, the topics used for the discussion tasks, and in the use and interpretation of the term ‘small-group discussion.’ Despite this variety, there is a high degree of consistency in the findings and conclusions. In general, students often struggle to formulate and express coherent arguments during small-group discussions, and demonstrate a low level of engagement with tasks.

The review presents very strong evidence of the need for teachers and students to be given explicit teaching in the skills associated with the development of arguments and the characteristics associated with effective group discussions. Indeed, five of the seven highest quality studies in the review make this recommendation.
Table 1.6: Overview of studies included in the review

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Review 1</th>
<th>Review 2</th>
<th>Review 3</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Vries et al., 2002</td>
<td>France</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Finkel, 1996</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>M (1), MH (2)</td>
</tr>
<tr>
<td>Gayford, 1995</td>
<td>UK</td>
<td>✓</td>
<td></td>
<td></td>
<td>MH</td>
</tr>
<tr>
<td>Hogan, 1999a</td>
<td>USA</td>
<td>✓</td>
<td></td>
<td></td>
<td>MH</td>
</tr>
<tr>
<td>Hogan, 1999b</td>
<td>USA</td>
<td>✓</td>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Jiménez-Aleixandre et al.,</td>
<td>Spain</td>
<td>✓</td>
<td></td>
<td></td>
<td>MH</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jiménez-Aleixandre and</td>
<td>Spain</td>
<td>✓</td>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Pereiro-Muñoz, 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson and Stewart, 2002</td>
<td>USA</td>
<td>✓</td>
<td></td>
<td></td>
<td>ML</td>
</tr>
<tr>
<td>Keys, 1997</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>MH</td>
</tr>
<tr>
<td>Kurth et al., 2002</td>
<td>USA</td>
<td>✓</td>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Lajoie et al., 2001</td>
<td>Canada</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>M</td>
</tr>
<tr>
<td>Lavoie, 1999</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>M</td>
</tr>
<tr>
<td>Meyer and Woodruff, 1997</td>
<td>Canada</td>
<td>✓</td>
<td></td>
<td></td>
<td>ML</td>
</tr>
<tr>
<td>Palincsar et al., 1993</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>ML (1), M (2,3)</td>
</tr>
<tr>
<td>Richmond and Striley, 1996</td>
<td>USA</td>
<td>✓</td>
<td></td>
<td></td>
<td>MH</td>
</tr>
<tr>
<td>Roth and Roychoudhury,</td>
<td>Canada</td>
<td>✓</td>
<td></td>
<td></td>
<td>MH</td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sherman and Klein, 1995a</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>H</td>
</tr>
<tr>
<td>Suthers and Weiner, 1995</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>ML</td>
</tr>
<tr>
<td>Tao, 2001</td>
<td>Hong Kong</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>ML (1), M (2,3)</td>
</tr>
<tr>
<td>Tao, 2003</td>
<td>Hong Kong</td>
<td>✓</td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td>Tolmie and Howe, 1993</td>
<td>UK</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>MH</td>
</tr>
<tr>
<td>Tsai, 1999</td>
<td>Taiwan</td>
<td>✓</td>
<td></td>
<td></td>
<td>ML</td>
</tr>
<tr>
<td>Williams, 1995</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>L</td>
</tr>
<tr>
<td>Woodruff and Meyer, 1997</td>
<td>Canada</td>
<td>✓</td>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Zohar and Nemet, 2002</td>
<td>Israel</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>MH (1), M (2,3)</td>
</tr>
<tr>
<td><strong>Total studies = 25</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The review presents good evidence that groups function better when the stimulus used to promote discussion involves both internal and external conflict, i.e. where a diversity of views and/or understanding are represented within a group (internal conflict) and where an external stimulus presents a group with conflicting views (external conflict).

There is good evidence on group structure. Groups function better when they are specifically constituted such that differing views are represented. There is also evidence to suggest that assigning managerial roles to students (e.g. reflector, regulator, questioner, explainer) as suggested in collaborative learning theory is likely to be counter-productive for poorly-structured tasks.

Some evidence is also presented which suggests single sex groups may function better than mixed sex groups, though overall development of understanding is not affected by the gender of group composition. Group leaders also emerge as having a crucial role: those who were able to adopt an inclusive style, and one which promoted reflection, were the most successful in achieving substantial engagement with the task. An alienating, overly-assertive leadership style generated a lot of ‘off-task’ talk and low levels of engagement.
Table 1.7: Details of the nature of the studies included in the review

<table>
<thead>
<tr>
<th>Study</th>
<th>Review</th>
<th>Sample details</th>
<th>Focus of study/nature of intervention</th>
<th>Data gathered</th>
</tr>
</thead>
</table>
| 1 De Vries et al., 2002 | 1,2    | • one class  
• 14 students  
• age 16-17  
• groups of two  
• asynchronous discussion via computer | The subject focus was physics and the topic of sound. Student discussions (via computer) were logged against 13 categories within explanation, argumentation, problem-solving and management. | • student self-report diaries  
• log of computer dialogue                                                                 |
| 2 Finkel, 1996 | 1,2    | • one class  
• 25 students  
• age 16-18  
• groups of 3-4 | The subject focus was biology and the topic of genetics. Students presented with two basic genetics models and data conflicting with these models. Students required to work in groups to produce a revised model. | • audiotapes of group discussions  
• audiotapes of plenary class presentations and discussions  
• computer logs  
• individual diaries  
• student work                                                                 |
| 3 Gayford, 1995 | 2      | • two classes (control and experimental) from four schools  
• age 16  
• groups of 3-4 | The subject focus was environmental science. Students were presented with material on environmental issues and asked to reach views, distinguishing between evidence and opinion. Control group covered the same material, but through teacher exposition. | • pre- and post-tests of six topic questions.  
• self-completion questionnaire to measure motivation                                                                 |
| 4 Hogan, 1999a | 1      | • one class  
• 24 students  
• age 13-14  
• groups of 3  
• mixed ability, friendship ties | No specific topic focus, but based on series of discussion tasks developed by the researcher. | • one to one interview  
• audio and video tapes of discussions  
• field notes of class observations                                                                 |
| 5 Hogan, 1999b | 1,2    | • 2 schools  
• 8 classes (four control and four experimental)  
• 163 students  
• age 11-16  
• groups of 3-4  
• heterogeneous for gender and ability | No specific topic focus, but based on series of discussion tasks developed by the researcher aimed at developing conceptual understanding and meta-cognitive skills relating to small-group discussions. | • one to one interview  
• audio and video tapes of discussions  
• field notes of class observations  
• tests of conceptual understanding  
• psychological profiles                                                                 |
<table>
<thead>
<tr>
<th></th>
<th>Authors</th>
<th>Setting</th>
<th>Methods</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Jiménez-Aleixandre et al., 2000</td>
<td>1 class 24 students age 14-15 groups of 4</td>
<td>The subject focus was biology and the teaching of genetics. There was no specific intervention, but the class selected for the research was felt to be used to learner-centred activities and small-group discussion work.</td>
<td>Observation, Audiotapes of group discussions</td>
</tr>
<tr>
<td>7</td>
<td>Jiménez-Aleixandre and Pereiro-Muñoz, 2002</td>
<td>1 school 38 students age 15-16 (plus some mature students) groups of 4-6</td>
<td>The focus was students’ ability to construct arguments and engage in decision-making about environmental processes.</td>
<td>Observation, Audio and videotapes of group discussions, Field notes, Student-generated material</td>
</tr>
<tr>
<td>8</td>
<td>Keys, 1997</td>
<td>1 school 6 students age 14-15 groups of 3</td>
<td>The subject focus was chemistry and the topic of elements and bonding. Discussions focused on the development of reasoning strategies and discourse through a collaborative writing tasks.</td>
<td>Videotapes of discussions, Interviews with students, Tests of conceptual understanding, Student work</td>
</tr>
<tr>
<td>9</td>
<td>Kurth et al., 2002</td>
<td>One school 4 students age 11-12 group of 4</td>
<td>The subject focus was physics and the teaching of density, with material being modified from the normal school module to incorporate discussion tasks.</td>
<td>Tests of conceptual understanding, Observation, Self-completion questionnaire, Self-completion diary</td>
</tr>
<tr>
<td>10</td>
<td>Lajoie et al., 2001</td>
<td>1 school 2 classes 40 students age 14-15 groups of 2</td>
<td>The subject focus was biology and the teaching of digestion, through the use of a computer-learning environment, Bioworld.</td>
<td>Audio and video tapes of discussions, Computer log of actions and decisions</td>
</tr>
<tr>
<td>11</td>
<td>Lavoie, 1999</td>
<td>1 school 10 classes (5 control and 5 experimental) 250 students age 15-16 groups of 3-4</td>
<td>The subject focus was biology, with a range of topics taught in a standard way, and though a learning cycle model (exploration, term introduction, concept application).</td>
<td>Daily logs kept by teachers, Observation, Video-recordings of lessons, Pre-post intervention tests of logical thinking, Conceptual understanding and attitude, Post intervention questionnaires to students and teachers</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>
</tbody>
</table>
| **12** | Palincsar et al., 1993 | 1,2,3 | • 2 schools  
• 9 classes  
• 230 students  
• age 11-12  
• groups of 4 | The subject focus was chemistry and the topic of kinetic theory, with discussion tasks aimed in particular at modelling the working of scientific communities. | • tests of conceptual understanding  
• interviews  
• video-recordings of selected groups  
• student logs |
| **13** | Richmond and Striley, 1996 | 1 | • 1 school  
• 1 classes  
• 24 students  
• age 15-16  
• groups of 4  
• mixed ability and gender | The subject focus was science, and the topics relation to cholera epidemics and cystic fibrosis. The aim was to explore the difficulties students encounter when developing scientific arguments and to explore how student interactions shaped arguments. | • observation  
• self-completion student diaries  
• school/college records |
| **14** | Roth and Roychoudhury, 1992 | 1 | • 1 school  
• 6 classes  
• 148 students (but only one group studied in detail)  
• age 15-17  
• groups of 3-4 | The subject focus was physics and the topic of light. The aim was to explore the development of student understanding through engaging in the process of developing concept maps. | • video-recording of discussions  
• self-completion questionnaire  
• concept maps generated in discussions |
| **15** | Sherman and Klein, 1995 | 2,3 | • 1 school  
• number of classes unspecified  
• 231 students  
• age 13-14  
• groups of 2 | The subject focus was investigations in science. Students worked through a computer programme about designing controlled experiments. | • observation  
• self-completion questionnaire  
• pre- and post-tests of understanding |
| **16** | Tao, 2001 | 1,2,3 | • 1 school  
• 1 class  
• 16 students  
• age 17-18  
• groups of 2 | The subject focus was physics. Multiple solutions were presented to students to see if discussions improved their understanding. | • pre- and post-tests of understanding  
• individual interviews  
• audio tapes of discussions |
| **17** | Tao, 2003 | 3 | • 1 school  
• 4 classes, but one used for most data collection  
• 150 students  
• age 15-16  
• groups of 2 | The focus was on the nature of science, and the responses of students’ when discussing the stories in groups. | • pre- and post-tests for whole sample  
• video and audio tapes of 36 students  
• field notes for observations of discussions  
• interviews with 18 students |
### The evidence on the use of small-group discussions (continued)

Some evidence emerged to support the notion that small-group discussion work does improve students’ understanding and use of evidence. Whilst this was not the main focus of this review, all the studies included presented some evidence in this area, as improvement in use of evidence was one of the reasons for using small-group discussions. The detailed effects of small group discussions on students’ understanding of evidence is presented in the next sections.

The review yielded almost no systematic evidence of the effects of small-group discussions on students’ attitudes to science.

The review also helped to provide information on the research strategies adopted to explore aspects of small-group-discussion work. A number of similarities emerged in the approaches adopted in the studies. They tended to make use of opportunistic samples, drawing on the researchers’ personal contacts. Experimental designs are not often used, though studies often made comparisons between discussion groups in the same class or within a discussion group. Data collection methods typically involve audio and/or video recordings, with analysis and reporting drawing heavily on extracts from recorded dialogue. Whilst approaches to gathering data are seldom justified in any detail by the authors, sound procedures appear to be introduced to check the reliability of the data analysis and present the findings in a way which makes them trustworthy.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>School</th>
<th>Classes</th>
<th>Students</th>
<th>Groups</th>
<th>Subject Focus</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tolmie and Howe, 1993</td>
<td>18</td>
<td>1</td>
<td>unspecified</td>
<td>82</td>
<td>2</td>
<td>Physics and forces and motion</td>
<td>Video-tapes, Psychological test, Computer record of joint predictions</td>
</tr>
<tr>
<td>Woodruff and Meyer, 1997</td>
<td>19</td>
<td>1</td>
<td>unspecified</td>
<td>82</td>
<td>3</td>
<td>Physics, shadows, floating and sinking</td>
<td>Audio-recordings, Field notes</td>
</tr>
<tr>
<td>Zohar and Nemet, 2002</td>
<td>20</td>
<td>2</td>
<td>5 (control, 2 experimental)</td>
<td>186</td>
<td>5-7</td>
<td>Biology and genetics</td>
<td>Multiple-choice test of understanding, Pre- and post-test of argumentation skills, Student worksheets, Audio-tapes of four discussions</td>
</tr>
</tbody>
</table>
Box 3: Summary of key findings on the use of small-group discussions

- There is considerable diversity in the topics used to promote small-group discussions.
- Students often struggle to formulate and express coherent arguments.
- Students often demonstrate a low level of engagement with tasks.
- Teachers and students need to be given explicit teaching in the skills associated with the development of arguments and the characteristics associated with effective group discussions.
- Groups function more purposefully when the stimulus used to promote discussion involves both internal and external conflict, i.e. where a diversity of views and/or understanding are represented within a group (internal conflict) and where an external stimulus presents a group with conflicting views (external conflict).
- Groups function more purposefully when specifically constituted such that differing views are represented.
- Single sex groups function more purposefully than mixed sex groups.
- Assigning managerial roles to students (e.g. leader, reflector, questioner, explainer) is counterproductive for poorly-structured tasks.
- Group leaders able to adopt an inclusive style, and one which promoted reflection, were the most successful in achieving engagement with the task.
- Little systematic data has been gathered on the effects of small-group discussions on students' attitudes to science.

A key difference which has emerged concerns the two contrasting approaches to data analysis, with some studies developing grounded theory from the data, and others drawing on existing models to structure their analysis.

The evidence on the effects of small-group discussions on understanding of evidence

Fourteen studies were included in the in-depth review, which focused on the effects of small-group discussions on students' understanding of evidence in science.

The evidence draws primarily on the findings from the twelve studies weighted as, medium-high or better and, to a lesser extent, as medium in quality.

The review suggests that there is reasonable evidence that use of small group discussions based on a combination of internal conflict (i.e. where a diversity of views and/or understanding are represented within a group) and external conflict (where an external stimulus presents a group with conflicting views) resulted in a significant improvement of students' understanding of evidence. Where there was either internal conflict or external conflict, there was some improvement in students' understanding.

Improvement of students' understanding of evidence was not significantly different for members of all-female, all-male or mixed gender pairs. There appeared to some benefit for female students when they were given several opportunities to engage with aspects of tasks related to understanding of evidence.

Improvement of students' understanding of evidence correlated with the initial dissimilarity of the group members in terms of their understanding of the science content of the discussion task, i.e. student groups were constructed such that they contained students with as wide a range of understandings as was possible.

The use of small group discussions supported by a specific programme fostering collaborative reasoning (including evaluating and strengthening of knowledge claims) improved
students' metacognitive knowledge of collaborative reasoning (including their knowledge of reasoning about evidence) significantly more than for students not following such a programme. The improved metacognitive knowledge of collaborative reasoning did not translate in better use of strategies while reasoning, including when dealing with scientific evidence.

The review suggests there is some evidence that the use of small group discussions (together with specific instruction in argumentation skills) improved students' ability to construct more complex arguments.

The effectiveness of small group discussions in producing an improvement in students' understanding of evidence appeared to depend on three types of understanding: understanding of the science domain, of the process by which model-revision takes place, and of metacognition.

The use of small group discussions resulted in a significantly better understanding for students using a cued version (i.e. one which gives students specific instructions on what to include in points they make in discussions) of a Computer-based Instruction (CBI) programme than a non-cued version.

Although outside the specific focus of the in-depth review question, one additional finding worth noting is that there was reasonable evidence to suggest that the gender composition of small discussion groups determined the interaction style for developing students' explanatory understanding: all-male groups confronted differences in their individual predictions and explanations, all-female groups searched for common features of their predictions and explanations across tasks and mixed groups secured progress through taking turns to contribute to the discussion.

The review also established that, where data were gathered of effects of small-group discussions, these tended to focus on aspects of understanding. Little systematic data has been gathered on the effects of small-group discussions on students' attitudes to science.

**Box 4: Summary of key findings on effects of small-group discussions on understanding of evidence**

- Students' understanding of evidence improves when discussion tasks involve both internal and external conflict, i.e. where a diversity of views and/or understanding are represented within a group (internal conflict) and where an external stimulus presents a group with conflicting views (external conflict).
- Improvements in understanding are independent of gender composition of groups, though single-sex groups functioned more purposefully.
- Improvements in understanding are greatest where there was initial dissimilarity in understanding of the science ideas associated with the discussion task.
- Students' understanding of evidence improves when they are provided with specific guidance on how to engage in small-group discussions and/or construct arguments.

---

**The evidence on the effects of different stimulus materials on understanding of evidence**

Ten studies were included in the in-depth review which focused on the effect of using different stimuli on students' understanding of the use of evidence in small group discussions. This discussion focuses on the eight studies of medium quality or better.

Two findings that emerged most strongly from this review.

Firstly, small group discussions focused on understanding the use of evidence, regardless of the prompt stimulus, are
enhanced by giving students some form of guidance on how to use that stimulus effectively. This guidance can be prior training in argumentation that provides instruction on how to use evidence or can be built into the structure or sequence of stimulus-based task.

Secondly, a successful stimulus for students working in small groups to enhance their understanding of evidence has two elements. One requires students to generate their individual prediction, model or hypothesis which they then debate in their small group (internally driven conflict or debate). The second element requires them to test, compare, revise or develop that jointly with further data provided (externally driven conflict or debate).

An additional finding was that prior knowledge can sometimes limit the understanding of evidence and its function. This can, for example, be the use of incorrect or inadequate factual knowledge or an idiosyncratic or inconsistent use of evidence to develop a hypothesis or test a model.

Finally, rich stimuli, such as those that provide complex and open-ended engagement, enhanced opportunities for developing understanding of evidence.

**Conclusions and implications**

**Issues in research on small-group discussions in science teaching**

It is clear from this review that there is considerable variation in the nature of research into small-group discussion work, particularly in relation to its focus, the clarity with which any variables being investigated are specified, and the techniques used to analyse data.

Further variation is introduced through the use of opportunistic samples for data collection. A substantial proportion of the work focuses on descriptive data. This can be very helpful in the early stages of a new research area.

---

**Box 5: Summary of key findings on effects of different stimulus materials on understanding of evidence**

- The effectiveness of small-group discussions is linked more strongly to the provision of specific guidance to students on how to engage in small-group discussions and/or construct arguments, rather than to different types of stimulus material.
- The most successful stimuli generate both internal and external conflict.
- Incorrect or inadequate prior knowledge hinders development of students’ understanding of evidence through small-group discussion.
- Students’ understanding is improved when they are presented with rich stimuli providing complex and open-ended engagement.

However, with increasing interest in the effects of small group discussions – on student learning, understanding, and attitudes – there is a need to consider what strategies and techniques lend themselves best to the gathering and analysis of data that would help explore such effects. A mixed method approach, which marries in-depth qualitative data on the nature of discussions and more quantitative data on student attributes would appear to have much to offer.

The collection of data to establish effects is not without problems in educational settings. Though there is considerable interest at present in the use of experimental designs in education research, practical considerations that may make such designs less feasible. For example, decisions on participation in new teaching strategies can seldom be made by researchers. This means that it is very difficult to allocate students or classes randomly to groups that will or will not receive an intervention. Most often, the research design has to be built around existing class sets in schools. This suggests that the use of
design experiments may prove potentially fruitful.

A note on design experiments
A design experiment in educational contexts involves evaluating the effects of an intervention in a limited number of settings. For example, this might involve selecting teachers who teach roughly comparable groups, but who have different teaching styles, and exploring the effects of the intervention on each group of students. The design experiment would then yield information on the circumstances in which the intervention is likely to be most successful. Design experiments have the advantage of being able to encompass the complexity of educational settings whilst enabling the aims of interventions to be tested systematically.

A particularly notable feature is the two very contrasting approaches to data analysis, with some studies developing grounded theory from the data, and others drawing on existing models to structure their analysis. This finding suggests that research into small-group discussions in science teaching would benefit from a consideration of discourse analysis techniques developed in other subject areas, such as English, to establish what they might have to offer work in science.

Implications for further research
One particularly strong feature which has emerged from the work undertaken for the review is that there is a dearth of systematic research on small-group discussion work and considerable uncertainty on the part of teachers as to what they are required to do to implement good practice. Both these factors point to a pressing need for a medium- to large-scale research study which focuses on the use and effects of a limited number of carefully-structured small-group discussion tasks aimed at developing various aspects of students’ understanding of evidence, linked to a coherent analysis framework.

Implications for policy
Current policy is strongly advocating the use of small-group discussion work. The review does indicate that there could be benefits in pursuing such a policy. However, it is clear from the review that small-group discussion work needs to be supported by the provision of guidance to teachers and students on the development of the skills necessary to make such work effective. Thus, some form of professional development training for teachers would appear to be highly desirable to provide them with guidance on how to maximise the effectiveness of small-group discussions.

Implications for practice
The review suggests that small-group discussion work can provide an appropriate vehicle for assisting students in the development of ideas about using evidence and constructing well-supported arguments. Thus teachers should be encouraged to incorporate such discussions into their teaching, provided that appropriate support is offered to help them develop the necessary skills. The gathering of additional research data on their use and effects would also be essential.
References in text of report


References to the full systematic review technical reports

References for the studies included in the review
Roth W-M, Roychoudhury A (1992) The social construction of scientific concepts or the concept map as conscription device and tool for social thinking in high school science. Science and Education 76: 531-557.


