



HAL
open science

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

Judith Bennett, Sylvia Hogarth, Fred Lubben, Bob Campbell, Alison
Robinson

► To cite this version:

Judith Bennett, Sylvia Hogarth, Fred Lubben, Bob Campbell, Alison Robinson. Talking science: the research evidence on the use of small-group discussions in science teaching. *International Journal of Science Education*, 2009, 32 (01), pp.69-95. 10.1080/09500690802713507 . hal-00544827

HAL Id: hal-00544827

<https://hal.science/hal-00544827>

Submitted on 9 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



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

Journal:	<i>International Journal of Science Education</i>
Manuscript ID:	TSED-2008-0099.R2
Manuscript Type:	Research Paper
Keywords:	learning activities, literature review
Keywords (user):	small group discussion



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

Abstract

This paper reports the findings of two systematic reviews of the use and effects of small-group discussions in high school science teaching. 94 studies were included in an overview (systematic map) of work in the area, and 24 studies formed the basis of the in-depth reviews.

The reviews indicate that there is considerable diversity in the topics used to promote small-group discussions. They also demonstrate that students often struggle to formulate and express coherent arguments, and demonstrate a low level of engagement with tasks. The reviews suggest that groups function more purposefully, and understanding improves most, when specifically constituted such that differing views are represented, when some form of training is provided for students on effective group work, and when help in structuring discussions is provided in the form of 'cues'. Single sex groups function more purposefully than mixed sex groups, though improvements in understanding are independent of gender composition of groups. Finally, the reviews demonstrate very clearly that, for small-group discussions to be effective, 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.

In addition to the substantive findings, the paper also reports on key features of the methods employed to gather and analyse data. Of particular note are the two contrasting approaches to data analysis, one adopting a grounded theory approach and the other drawing on established methods of discourse analysis.

Introduction

The use of small-group discussions in teaching has been advocated for a number of years in science teaching, both to motivate students and to enhance their learning. Such tasks are now appearing with greater frequency in a range of high school science teaching resources.

Despite this increasing popularity, comparatively little is known about the detail of their use and effects. This paper examines the research evidence on the use of small-group discussions

1
2
3 through presenting the findings of two systematic reviews, with a view to making
4
5 recommendations for policy and practice on the use of small-group discussions in science
6
7 teaching.
8
9

10 As originally conceived, the reviews had two principal aims. The first of these was to identify
11 the ways in which small-group discussions are currently used in science lessons. The second
12 was to look at the effects of small group discussions on students' understanding of science
13 and on students' attitudes to science. In practice, the review work established that there is a
14 dearth of studies reporting in any detailed and systematic way on the effects of small-group
15 discussions on students' attitudes to science, so this paper reports the review findings on the
16 use of small-group discussions in science teaching and their effects on students'
17 understanding of science ideas.
18
19
20
21
22
23
24
25

26 Several factors have contributed to the current high levels of interest in small-group
27 discussion work in science. The use of small-group discussion in teaching has its origins in
28 learner-centred teaching approaches, and is one of a range of 'active learning' strategies
29 aimed at stimulating students' interest in what they are studying by providing them with a
30 significant degree of autonomy over the learning activity (e.g. Bentley and Watts, 1992;
31 Kyriacou, 1998). A number of people have advocated the use of discussions in science
32 lessons. Lemke (1990) argues that "learning science means learning to talk science" (p1), and
33 that this means moving away from science lessons dominated by teacher talk in which the
34 teacher asks a question, then invites and evaluates a student response. Lemke refers to this as
35 'triadic dialogue', which is similar to the initiation-response-feedback (IRF) sequence
36 characterised earlier by Sinclair and Coulthard (1975). Lemke and others (e.g. Sutton, 1992,
37 and Wellington and Osborne, 2001) criticise the approach for leading to talk that focuses on
38 what the teacher wants to hear, rather than promoting genuine communication, and all argue
39 for increased use of discussion work in science lessons.
40
41
42
43
44
45
46
47
48
49
50
51
52

53 Support for the use of small-group discussions in science teaching has also emerged in the
54 recommendations from other areas of work in science education. For example, research on
55 alternative conceptions has explored in depth the ideas and understanding students bring with
56 them to science lessons and the ways in which some of their ideas may hinder the
57 development of accepted scientific ideas (e.g. Driver *et al.*, 1985). Small-group discussions
58 have been suggested as a means of helping students explore their ideas and move from
59
60

1
2
3 understandings that may often be naïve to towards more valid scientific ideas and
4 explanations. Further impetus for the inclusion of small-group discussions in science lessons
5 has come from the development of ideas about *social constructivism* (Driver *et al.*, 1994),
6 which draws on the work of Vygotsky and the importance to learning of social dynamics of
7 interaction (Scott, 1998).
8
9
10
11

12
13
14 The publication of *Beyond 2000* (Millar and Osborne, 1998) in the UK stimulated discussion and
15 debate over the nature of the school science curriculum and, in particular, the ways in which it
16 might foster the development of *scientific literacy*. This term embraces the knowledge,
17 understanding and skills young people need to develop in order to think and act appropriately on
18 scientific matters which may affect their lives and the lives of other members of the local,
19 national and global communities of which they are a part. Many of the materials developed
20 include small-group discussions in the repertoire of activities employed in science lessons in
21 order to encourage students to participate in informed discussion and debate of scientific issues
22 (e.g. Millar, 2006). Linked to the development of scientific literacy is that of *ideas about*
23 *evidence* (Millar and Osborne, 1998), which involves encouraging students to evaluate, interpret
24 and analyse evidence from primary and secondary sources in science, including stories about how
25 important science ideas were first developed, and then established and finally accepted. This has
26 led to considerations of the role of *argument* in school science, in the sense of putting forward
27 claims and supporting them with sound and persuasive evidence (e.g. Newton *et al.*, 1999;
28 Osborne *et al.*, 2001). Small-group discussions are viewed as having a key role to play here,
29 since the practice of using evidence in argumentation requires interaction with peers.
30
31
32
33
34
35
36
37
38
39
40
41
42

43
44 In the UK, two other developments have also served to raise the profile of small-group
45 discussions. Firstly, the publication of *Inside the Black Box* (Black and Wiliam, 1998) has
46 resulted in considerable attention being paid to formative assessment, or assessment for
47 learning. Small-group discussions are one approach that has been advocated for increasing
48 the use and effectiveness of formative assessment in science teaching (see, for example, Daws
49 and Singh, 1999). Secondly, there is a more general drive to improve students' *literacy skills*.
50 In England and Wales, this has been formalised into the National Literacy Strategy (DfEE,
51 1998), where small-group discussions have been advocated as a means for developing
52 students' oral communication skills in science.
53
54
55
56
57
58
59
60

1
2
3 Whilst the preceding discussion point to a general recognition of the potential value of small-
4 group discussion in science teaching, their use challenges the established pedagogy of science
5 teaching and place new demands on teachers (see, for example, Bentley and Watts, 1992).
6
7 There is a growing body of evidence, both anecdotal and more systematic, that many teachers
8 lack skills and do not feel confident with small-group discussions. In particular, evaluation of
9 materials with a specific focus on teaching socio-scientific issues and developing scientific
10 literacy, such as *AS Public Understanding of Science* course (Osborne *et al.*, 2002) and the
11 *Valuable Lessons* project (Levinson and Turner, 2001), raised particular concerns about
12 teacher efficacy in the use of small-group discussions. For example, Osborne *et al.* (2002)
13 comment, “Our view, given the poor quality of much of the teaching involving discussion, is
14 that training is essential. Teachers ... need an opportunity to interact with experienced
15 humanities teachers, to observe the strategies they use for fostering and stimulating discussion
16 ...” (p75). In a similar vein, Levinson and Turner recommended that, “Teacher training
17 courses should provide prospective science teachers with more opportunities in the area of
18 initiating and managing discussions ... practicing teachers should improve and update such
19 skills through CPD [continuing professional development].” (p21).
20
21
22
23
24
25
26
27
28
29
30
31
32

33 It is clear from the discussion above that there is considerable interest in the use of small-
34 group discussion in science teaching. Some of this interest has emerged directly from
35 research studies, whilst, in other areas, it draws more loosely on research evidence and take
36 the form of approaches which are being advocated in science teaching, but whose effects have
37 yet to be explored on a more systematic basis. There also appears to be a comparative lack of
38 guidance for teachers. The evidence presented in this paper provides a number of insights
39 into the form such guidance might take.
40
41
42
43
44
45
46
47
48
49

50 **The origins and purposes of systematic reviews**

51 Systematic reviews of research studies are a comparatively recent development in education,
52 though they are well established in medical research. They have emerged from the
53 international debate over the nature and purpose of educational research, and how it
54 contributes to maximising the effectiveness of educational provision (*e.g.* Hargreaves, 1996
55 and Hillage *et al.*, 1998, in the UK; Shavelson and Towne, 2001, in the USA).
56
57
58
59
60

1
2
3 There are several reasons why systematic reviews are being seen as a key strand in
4 educational research. Firstly, there is a growing interest in practical policy-related decision
5 making being linked to evidence in a number of areas, not just in education. Systematic
6 reviews of research literature are seen as having the potential to yield evidence on which
7 policy makers can draw (Davies, 2000; Torgerson and Torgerson, 2001). Secondly, there is a
8 drive towards forging closer links between research, policy and practice (see, for example,
9 Hargreaves, 1996; OECD, 2002; Oakley, 2002). In particular, drawing on research findings
10 in classroom practice is seen as desirable, with teachers being encouraged to engage in what is
11 variously described as ‘evidence-based’, ‘evidence-informed’ or ‘evidence-enriched’ practice
12 (e.g. Millar *et al.* 2006).
13
14
15
16
17
18
19
20
21
22

23 In 2000 the Government in the UK funded, via the Department for Education and Skills
24 (DfES), the Evidence for Policy and Practice Initiative (EPPI)-Centre to focus on undertaking
25 systematic reviews of research evidence in key areas of education, and reporting these in a
26 form accessible to a range of different user groups, including teachers, researchers and policy-
27 makers. The advantages and disadvantages of systematic reviews when compared with other
28 forms of review, and of reviews employing the EPPI methodology have been rehearsed
29 elsewhere (Bennett *et al.*, 2005).
30
31
32
33
34
35
36

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

- 38 • identification of review topic area and review research question or questions;
- 39 • development of *inclusion and exclusion criteria* for studies in the review (relating to, for
40 example, aspects such as the age of students, the nature of the research design, and the
41 reported outcomes);
- 42 • undertaking of *systematic searches* of electronic data bases and other sources for
43 potentially relevant research studies;
- 44 • refining the search through *screening* the potentially relevant studies against the inclusion
45 criteria;
- 46 • coding or *keywording* studies against pre-specified and agreed characteristics (some of
47 which are generic to all EPPI reviews, whilst others are developed specifically for each
48 review);
- 49 • production of an overview or *systematic map* of studies in the review area, that groups the
50 studies according to their chief characteristics;
- 51
52
53
54
55
56
57
58
59
60

- undertaking an *in-depth review* of studies to look in detail at their design and findings and to evaluate the quality of the evidence reported.

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*. Full details of the EPPI methodology may be found in the EPPI Review Group Manual (EPPI-Centre, 2002).

The EPPI review methods tend to be more tailored to quantitative research. As much of the work on small-group discussion draws extensively on qualitative approaches, the work reported here extended the EPPI review methods to draw on the guidance and framework for assessing research evidence in qualitative research studies (Spencer *et al.*, 2003). This was seen as particularly important as the majority of the studies included in the review made use of qualitative methods. As such, they ran the risk of being characterized as low in quality had judgments been informed only by the EPPI criteria.

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

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?* Within this, two reviews were conducted. The purpose of the first review was to gain an overview of the nature of small-group discussions used in science lessons in order to assess the levels of use, establish any patterns in use, such as in topic focus, group structure and the dynamics of group interactions. As much of the support for the use of small-group discussions is linked to their potential benefits associated with to improved understanding of science ideas, the second review focused specifically on the effects of small-group discussions on understanding.

A note on terminology

1
2
3 The reports on small-group discussions scrutinised for the systematic reviews made it clear
4 that the term was used in a wide variety of ways. For the purposes of the reviews, a small-
5 group discussion was taken to be an activity that:
6
7

- 8 • involves groups of two to six students;
- 9 • has a specific stimulus (e.g. a newspaper article, video clip, prepared curriculum
10 materials; structured teacher input);
- 11 • involves a substantive discussion task of at least two minutes (i.e. did not simply involve a
12 student talking to a neighbour briefly about an idea or to agree and answer to a question);
- 13 • is either *synchronous* (i.e. face-to-face) or *asynchronous* (i.e. mainly IT-mediated);
- 14 • has a specific purpose (e.g. individual sense-making, or leading to an oral presentation, or
15 to a written product).

16
17
18
19
20
21
22
23
24
25 The term *understanding* has been taken to encompass understanding of science concepts,
26 understanding of ideas about the nature of science and understanding of the methods of
27 science.
28

29 30 31 *Studies in the review*

32
33
34
35 The reviews focused on research on teaching at high school level, undertaken in the period
36 1980-2005, and published in English. Student age was restricted to 11-18 because this is the
37 age range covered by the majority of reported studies, pointing to this being the school age
38 range where the use of small-group discussions has been promoted most actively. The start
39 date for the period of publication was selected because this was the time when the use of
40 small-group discussions started to become more prominent in science teaching. The inclusion
41 criteria for studies in the review were: (i) they were about the use of small-group discussions
42 in science lessons, (ii) they involved groups of two to six students, (iii) they focused on a
43 substantive, structured discussion task of two minutes' duration or more, and, for the second
44 review, (iv) they addressed aspects of students' understanding in science.
45
46
47
48
49
50
51
52

53
54
55 94 studies were identified that met the above inclusion criteria. The identification of the
56 studies is a two-step process. Firstly, systematic searches are undertaken, principally through
57 the use of electronic search strings. Electronic searching inevitably means that large numbers
58 of studies emerge in the initial stage, and some 2,290 studies matched the search terms. The
59 second step involves refining the search through careful screening of abstracts (or full copies
60

1
2
3 of reports if there is insufficient information in the abstract) against the inclusion criteria. The
4
5 94 studies identified were then coded for particular characteristics (*keyworded*) to produce an
6
7 overview (*the systematic map*) of studies of small-group discussion interventions. In
8
9 producing the map, the following characteristics of studies scrutinised included the country of
10
11 study, the age/level of the students, the type of study, the science discipline of the study,
12
13 constitution of discussion groups, the duration of discussion tasks, the stimulus provided,
14
15 organisational features of discussion tasks, the product of discussion tasks, the research
16
17 strategy used to gather data, the nature of data collected, and the outcomes reported.

18 19 **An overview (systematic map) of studies on small-group discussions**

20
21
22
23 This section presents a brief overview of the key features to emerge from the 94 research
24
25 reports on the use and effects of small-group discussions.

26
27
28 Although the period of the review covered 1980-2005, over 90% were from post-1990,
29
30 indicating that that the research activity area has been minimal up to fifteen years ago and has
31
32 been most prolific from the late 1990s onwards. The majority of the reported work has been
33
34 undertaken in North America (USA = 39%, Canada = 12%) and the UK (13%). Other work
35
36 has been undertaken in Australia, The Netherlands and Germany. However, these figures
37
38 need to be set in the context of the review being limited to reports published in English,
39
40 though the review does include reports of studies of small-group discussions held in Bahasa
41
42 Malay, Cantonese, Dutch, Finnish, French, German, Greek, Hebrew, Mandarin, Portuguese
43
44 and Spanish.

45
46 Relatively little detail was given about how groups were constituted, but where this was
47
48 provided, mixed-ability and friendship groups predominated. In one-third of cases, groups
49
50 were deliberately constituted by the teacher, with students choosing their groups in the
51
52 remaining cases. Slightly over half the studies used a group size of 3-4 students, with a
53
54 further quarter of studies using pairs of students for groups. Over 80% of the studies
55
56 concerned self-contained and permanent groups. The remaining studies drew on the
57
58 techniques of 'snowballing', 'envoying' and 'jigsawing'. Snowballing discussions involve
59
60 progressively larger groups of students discussing a question or idea and agreeing on their
views. Discussion starts with pairs, who then join together and so on. Envoying discussions
involve groups of students discussing a common task. When the discussion is completed, one

1
2
3 member of each group moves to another group to report on the discussion of the original
4 group and to hear about the discussion in the second group. The envoy then returns to the
5 original group to report back. Jigsawing discussions have students first working in groups
6 where each has a different task and then moving into different groups that comprise of all the
7 people who have focused on the same aspect of the task.
8
9
10
11

12
13
14 Two-thirds of the discussions took a class period or longer, with approximately two-fifths
15 taking place in lessons described as 'science', a similar proportion in physics lessons, around
16 one-fifth in biology lessons and only 4% in chemistry lessons. One possible explanation for
17 this very small figure for chemistry is that most of the small-group discussions relating to
18 exploration of difficult ideas are located mainly within physics, and those developing skills of
19 decision-making on socio-scientific issues are more commonly placed within biology classes.
20
21
22
23
24
25

26
27 The most common stimulus material provided for the discussion tasks was printed
28 worksheets, which were used in over two-thirds of the studies. Practical work and computer
29 software provided the stimulus for around two-fifths and a quarter of discussions respectively.
30 Some discussion made use of more than one stimulus. Rather surprisingly, only one study
31 used a newspaper article as a stimulus for discussion.
32
33
34
35
36

37
38 In a very high proportion of the studies (94%), the main aim of the discussion task was
39 individual understanding of the science underlying the activity, such as in a practical
40 experiment, the preparation of a poster or a computer-based exercise, in which the learners
41 were engaged. In the majority of cases this understanding was then shared with classmates in
42 different ways: groups might present their findings or views orally (20%) or by way of posters
43 (10%) or might defend their position in a whole class debate (5%). There were surprisingly
44 few examples of written products being generated directly as an outcome of the discussion
45 (6%).
46
47
48
49
50
51

52
53 Around three-fifths of the studies reported on evaluations of small-group discussions with the
54 remainder providing descriptive information about the use of small-group discussions. Case
55 studies featured prominently, with extensive use being made of video and audio recordings in
56 order to gather detailed information about the nature of discussions. One outcome of the very
57 labour-intensive nature of much of the data collection and analysis was that sample sizes
58
59
60

1
2
3 tended to be small – very often one class or one or two groups of students within a class.
4
5 Studies involving several classes, or classes in more than one school, were rare.
6
7

8
9 The chief characteristics of research on small-group discussions are summarised in Table 1.
10

11 [Table 1 about here.]
12
13

14 15 16 **The in-depth reviews** 17

18
19 Two in-depth reviews were conducted, covering a total of 24 of the 94 studies in the
20 systematic map. These 24 studies were selected because (i) they reported in detail on the use
21 of small-group discussions in science lessons (19 studies), and/or (ii) they reported
22 evaluations of interventions aimed at developing aspects of students' understanding in science
23 (14 studies).
24
25
26
27

28
29 Studies were rated as high (H), medium high (MH), medium (M), medium low (ML) or low
30 (L). As quality judgements in systematic reviews are made in relation to the specific focus of
31 each of the reviews, some studies were given different ratings in each review. Ratings were
32 based on the extent to which the studies reported met a range of criteria relating to (i) the
33 nature of the sample and how it was selected; (ii) the nature of any control group (for
34 evaluations); (iii) the extent to which small-group discussions formed the main feature or
35 variable being investigated (for evaluations); (iv) the level of detail provided about the
36 discussion task; (v) the steps taken to establish the reliability and validity of the data
37 collection tools and processes, and the data analysis; (vi) the trustworthiness and reliability
38 of the data collection and analysis for qualitative data; (vii) the representativeness of the data
39 collection situation to normal classroom situations; (viii) the quality of the reporting. No
40 studies met all the relevant criteria, so none were judged to be high quality. The studies
41 rated as 'medium high' met most of the relevant criteria. Studies rated at the lower end of the
42 scale displayed one or more of the following characteristics: they provided comparatively
43 little detail of the discussion tasks, they tended to be weaker in the reporting of steps taken to
44 enhance the reliability and validity of data collection and analysis, and/or they were overly-
45 descriptive at the expense of analysis and discussion. The categorisation process is described
46 in detail in the full technical reports (see References section).
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Table 2 lists the studies and their quality rating for each review.
4
5

6
7 [Table 2 about here.]
8
9

10 The remaining discussion focuses on the nineteen studies rated as 'medium' or better, as these
11 provide the stronger evidence, and Table 3 summarises the key features of each of these
12 studies.
13
14

15
16
17 [Table 3 about here.]
18
19

20 21 **The detailed evidence from the in-depth reviews** 22

23
24 The consolidated evidence presented draws primarily on the findings of the fourteen studies
25 in the Review 1 and the twelve studies in Review 2 weighted as medium or better in overall
26 quality, as these studies generated the better quality evidence. Of these studies, nine were
27 common to both reviews, i.e. they focused on aspects of the use of small-group discussions
28 and the development of understanding. The features discussed below are considered to be
29 those that offer the strongest evidence as they emerged from three or more studies. To avoid
30 undue repetition, evidence from both reviews is considered together under the following five
31 headings: focus of the discussion topic, group structures and interactions, negotiating and
32 agreeing meaning through discussion, effects on understanding, factors promoting effective
33 discussion to enhance understanding
34
35
36
37
38
39
40
41
42

43 44 *Focus of the discussion topic* 45

46
47 The studies were based on a range of science topics, as Table 3 demonstrates. Seven studies
48 addressed aspects of understanding of science topics. These focused on light (Roth and
49 Roychoudhury, 1992), kinetic theory (Palincsar *et al.*, 1993), genetics (Finkel, 1996; Jiménez-
50 Aleixandre *et al.*, 2000), two physics topics [shadows, floating and sinking] (Woodruff and
51 Meyer, 1997), mechanics (Tao, 2001), and density (Kurth *et al.*, 2002). Four looked
52 primarily at aspects of what could be termed scientific method: hypotheses on the diagnosis of
53 disease (Richmond and Striley, 1996; Lajoie, 2001), designing controlled experiments
54 (Sherman and Klein, 1995), and building theories and models from primary evidence on
55 elements and bonding (Keys, 1997). Three had a specific focus on socio-scientific issues in
56
57
58
59
60

1
2
3 relation to the greenhouse effect (Gayford, 1995), genetic engineering (Zohar and Nemet,
4 2002), and environmental science (Jiménez-Aleixandre and Pereiro-Muñoz (2002). Three
5 involved making predictions based on evidence presented in the topics of sound (De Vries *et*
6 *al.*, 2002), a range of biology topics (Lavoie, 1999), and mechanics (Tolmie and Howe,
7 1993). The studies by Hogan (1999a and 1999b) had no specific topic focus but were based
8 on a series of discussion task developed by the researchers.
9
10
11
12

13
14
15
16 The studies in the in-depth review reflect the patterns noted in the systematic map, where the
17 bulk of the discussion topics lay in the areas of physics and biology. Although there was
18 diversity in topic focus, the common link between the discussions focusing on the
19 development of understanding was that they all required students to draw on evidence to
20 support a particular hypothesis, theory or point of view.
21
22
23

24 25 26 *Group structures and interactions* 27

28
29
30 The principal evidence on group structure and interaction came from five studies: Tolmie and
31 Howe (1993), and Richmond and Striley (1996); Keys (1997); Hogan (1999a); De Vries *et*
32 *al.*, 2002.
33
34

35
36
37 Group leadership emerged as crucial in promoting effective discussion. Richmond and
38 Striley (1996) and Kurth *et al.* (2002) established the need for a leader to adopt an inclusive
39 style and share tasks equitably around a group, as this promoted more substantial engagement
40 in the discussion by a number of participants, and increased the quality of the discussion.
41 This, in turn, permitted most members to develop their understanding. Non-inclusive
42 leadership generated much off-task talk and engagement was generally low. Hogan (1999a)
43 found that at least one group member had to act in a way which promoted reflection in the
44 group for understanding of science ideas to be developed.
45
46
47
48
49
50

51
52
53 The evidence on assigning specific roles to students was mixed. Kurth *et al.* (2002)
54 advocated assigning particular roles to pupils in groups as a means of achieving effective
55 discussion. However, Richmond and Striley, (1996) and (Hogan, 1999a) reported allocating
56 roles to have benefits when tasks were well-structured but counterproductive in poorly-
57 structured tasks, adding to students' difficulties in engaging with the task. Hogan identified
58 eight sociocognitive roles in group reasoning processes which took place in small-group
59
60

1
2
3 discussions, all of which were persistent over time, and independent of particular formal roles
4 that might have been allocated to students in advance of a discussion. Four of Hogan's roles
5 were positive (promoter of reflection, contributor of content knowledge, creative model
6 builder, mediator of social interaction and ideas) and four were negative (promoter of
7 distraction, of acrimony, of simple task completion, reticent participant).
8
9

10
11
12
13 Aspects of group composition were explored in four studies (Tolmie and Howe, 1993; Keys,
14 1997; Hogan, 1999a; De Vries *et al.*, 2002). Tolmie and Howe found improved
15 understanding to be independent of group composition (male, female, mixed), with biggest
16 improvements being noted when groups contained members with a high degree of
17 dissimilarity in their initial predictions and explanations. In common with Keys and De Vries
18 *et al.*, Tolmie and Howe identified clear differences in interactional styles with all-male
19 groups confronting differences in their individual predictions and explanations, whilst all-
20 female groups searched for common features of their predictions and explanations in order to
21 avoid conflict. Mixed groups interacted in a more constrained way than single-gender groups,
22 though they also tended to avoid conflict and look for common patterns in contributions.
23 Tolmie and Howe suggested that both male pairs and female groups demonstrated qualities
24 one would want to see in the development of arguments but did not see this as a reason for
25 promoting the use of mixed gender groups, as their study suggested the best of the all-male
26 and all-female group interactions was lost in mixed pairs. Hogan also found that friendship
27 groups, which were generally single-sex, functioned more effectively and promoted better
28 development of understanding than mixed or teacher-constituted groups.
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43
44 The studies suggest that both the behavioural characteristics and gender composition of
45 groups need careful consideration if groups are to function purposefully during discussion
46 tasks.
47
48

49 50 51 *Negotiating and agreeing meaning through discussion* 52

53
54 Three of the studies (Roth and Roychoudhury, 1992; Keys, 1997; Jiménez-Aleixandre *et al.*,
55 2000) reported in detail on the ways in which meanings were negotiated and agreed by
56 groups.
57
58
59
60

1
2
3 Keys (1997) identified three common characteristics of reasoning in discussions: recognising
4 that prior ideas (models) may be incorrect; evaluating new observations for consistency with
5 current ideas and using evidence to modify ideas; and coordinating all mutually consistent
6 knowledge propositions into a coherent model. A similar pattern was described by Roth and
7 Roychoudhury (1992) who found discussions usually involved positions being stated and
8 contested, with views either accepted or temporarily or permanently rejected as positions
9 finally stabilised into shared meaning. Less positively, they found that students tended not to
10 engage very often in processes which fostered meaning. Rather they would reach agreement
11 on the basis of finding something agreeable to all group members. Agreements were often
12 reached by one or more group members exerting authority, and on the basis of 'majority rule',
13 rather than agreed shared understanding. These findings were echoed by Jiménez-Aleixandre
14 *et al.* (2000) who reported that a large proportion of student talk related to what they termed
15 'doing the lesson', or interactions referring to the rules of the task, rather than talk related to
16 the focus of the task. Jiménez-Aleixandre *et al.* also noted that arguments were frequently
17 developed by a subset within the group and, though agreement was generally reached, this
18 was often for social reasons, rather than because of agreement over the outcome of the
19 discussion.
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34
35 Whilst the above findings suggest that groups can work together to develop common
36 understandings, they also suggest that a number of factors can influence the way in which
37 these understandings are reached, and relative lack of engagement with the science content,
38 coupled with the influence of students prepared to express views very strongly means that the
39 gains in understanding may not be that high.
40
41
42
43
44

45 *Effects on understanding*

46
47
48

49 All the studies that looked at development of understanding reported benefits from small-
50 group discussion work, though this finding does need to be interpreted with some caution as
51 the majority of those undertaking the research were advocates of the approach. The following
52 provide examples of the effects that were reported.
53
54
55
56

57
58 Roth and Roychoudhury (1992) established that group discussion over the construction of a
59 concept map provided a vehicle for negotiation of meaning and understanding of concepts
60 and their relationships, thus providing a structure through which students were able to learn

1
2
3 the language patterns of science and use these to construct scientific knowledge. Tolmie and
4 Howe (1993) reported significant improvements in students' predictions of the trajectories of
5 falling objects through the use of a computer-based simulation. Richmond and Striley (1996)
6 noted increasing levels of sophistication and increased use of subject knowledge in the
7 arguments students developed in discussion of socio-scientific issues. Similarly, Zohar and
8 Nemet (2002) reported substantial changes in the quality of student arguments in the context
9 heredity and genetics.

16 17 *Factors promoting effective discussion to enhance understanding*

20
21 The factors that emerged that contributed to effective discussions and enhancement in
22 understanding are of particular interest. Findings from four studies pointed to improvements
23 in understanding being greatest for discussion tasks where there was dissimilarity or conflict
24 in understanding or views. This might take the form of either *internal conflict*, or differences
25 held by individual group members (Tolmie and Howe, 1993; De Vries *et al.*, 2002), or
26 *external conflict* where an external stimulus presents a group with conflicting views (Tolmie
27 and Howe, 1993; Gayford, 1995; Finkel, 1996). In some of the studies the discussion topic
28 was selected to provide opportunities for both internal and external conflict. For example,
29 Tolmie and Howe (1993) required students to make individual predictions about aspects of
30 forces and motion, then engage in a task which required a joint prediction (internal conflict)
31 and finally to compare this with an actual situation to reach an explanation of any
32 discrepancies (external conflict). Whilst other studies did not comment specifically on the
33 need for dissimilarity, it was clear from some of the accounts (e.g. Zohar and Nemet, 2002)
34 that internal and external conflict were built into the discussion tasks. It seems likely that the
35 dissimilarity in views provides a very clear and immediate focus to engage students in
36 discussion.

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51 Two studies offered comments on the nature of the data provided to students for the
52 discussion. Jiménez-Aleixandre *et al.* (2000) indicated that hypothetical, unquestionable data
53 (provided by the teacher) generates different patterns of discussion than empirical, uncertain
54 data, perhaps gathered by students themselves, with the former leading to greater gains in
55 understanding. In a similar vein, Roth and Roychoudhury (1992) found discussion to be more
56 productive if students were provided with a fixed set of concepts to delimit the content of the
57 discourse.
58
59
60

1
2
3
4
5 Three studies pointed to improved understanding when students were given specific
6 instructions on how to construct arguments or cues to guide them in the points they needed to
7 include (Sherman and Klein, 1995; De Vries *et al.*, 2002; Zohar and Nemet, 2002). This
8 finding was reflected in the more general observations in two further studies that scaffolding
9 routines, or structuring discussion through the provision of interim targets, also improved
10 students' understanding (Palincsar *et al.*, 1993; Finkel, 1996).
11
12
13
14
15
16

17 Although gains in understanding were reported, the studies also suggested that students often
18 struggled to formulate and express coherent views during small-group discussions, and
19 demonstrated a low level of engagement with tasks. It is therefore not surprising that a
20 number of the studies made recommendations relating to the need for students and/or teachers
21 to be given explicit teaching in the skills associated with the development of arguments and
22 the characteristics associated with effective group discussions.
23
24
25
26
27
28
29

30 Three studies (Richmond and Striley, 1996; Hogan, 1999a; Zohar and Nemet, 2002)
31 recommended training for both students and teachers in the skills needed for handling and
32 participating in group discussions. Richmond and Striley indicated that productive learning
33 was unlikely to take place on a large scale through the use of small-group discussions until
34 students acquire the skills associated with inclusive leadership and are thus able to foster a
35 climate of equitable participation. Hogan (1999a) argued that guiding students towards taking
36 constructive roles in discussions could be achieved through metacognitive training, i.e.
37 knowledge about the nature of collaborative learning, effective group learning strategies, and
38 awareness of what constitutes progress. Two studies (Jiménez-Aleixandre *et al.*, 2000 and
39 Roth and Roychoudhury, 1992) recommended coaching in argumentation skills for both
40 teachers and students.
41
42
43
44
45
46
47
48
49
50

51 Jiménez-Aleixandre *et al.* (2000) suggested that effective discussions are only likely to take
52 place when linked to specific, inquiry-focused tasks where help is given to students to
53 develop their understanding through the construction of arguments. Similarly, Roth and
54 Roychoudhury (1992) reported that students frequently struggled with language, often making
55 short utterances, and appeared to find it difficult to clarify their understanding through
56 explanations, justifications and elaborations. This led them to conclude that a major outcome
57
58
59
60

1
2
3 of their study was the recognition that understanding was only likely to be improved if
4 students were given help in constructing arguments.
5
6
7

8
9 One study (Zohar and Nemet, 2002) did involve incorporating explicit instruction about
10 argumentation into their intervention. One introductory lesson involved arguments being
11 defined and their structure explained, together with providing examples of characteristics of
12 good arguments. Students then practised the principles through several concrete examples.
13 Zohar and Nemet concluded that argumentation skills were enhanced by explicit instruction
14 and several opportunities for students to take part in discussions to help develop their skills.
15
16
17
18
19

20
21 The nature of the evidence presented in a substantial number of the studies points to the
22 importance of training for teachers and students. This is all the more important as many of
23 the researchers involved in the studies were committed to the use of small-group discussions
24 and had developed some proficiency in designing discussion tasks. This review finding also
25 resonates with the recommendations of Levinson and Turner (2001) and Osborne *et al.* (2002)
26 in their evaluations of two programmes focusing on teaching of socio-scientific issues, and
27 noted earlier in this paper.
28
29
30
31
32
33

34 35 *Methodological considerations* 36

37
38 Although the primary focus of the reviews was to gather substantive data on the use of small-
39 group discussions, some of the methodological aspects have a bearing on the nature of
40 evidence yielded.
41
42
43
44

45
46 Positive features of the data collection included the use of multiple data sources with all
47 studies drawing on at least two different kinds of data to increase trustworthiness. Whilst
48 virtually all studies used video recording and/or audio recording to make verbatim records of
49 discussions, these were supported by direct observation to record field notes, interviews,
50 products of student tasks, such as concept maps, student questionnaires and measures of
51 student knowledge were obtained. Although the methods used were rarely justified, the
52 picture gained was one of studies collecting extensive data in an attempt to get as detailed a
53 picture as possible of students' dialogue.
54
55
56
57
58
59
60

1
2
3 There were a number of limitations to the data collection. All the studies used a convenience
4 sample for the identification of schools, in many cases using schools where access had been
5 secured through previous involvement of the researcher. With one exception (Zohar and
6 Nemet, 2002), the studies were based in one school and often within one class. A
7 characteristic of much of the work was the use of retrospective sampling, i.e. data were
8 gathered on a number of groups, but reports presented on only a sample of the groups within
9 this, depending on characteristics of the discussion which emerged in the analysis. Such
10 sampling methods are probably realistic for research studies fitting in with practice, and
11 requiring extensive periods of data collection and thus a high degree of co-operation with the
12 class teachers involved. However, retrospective sampling does confer the option on the
13 researcher of exercising a high degree of selectivity in relation to the data presented.
14
15
16
17
18
19
20
21
22
23

24 Three sizeable studies (Gayford, 1995; Lavoie, 1999; Zohar and Nemet, 2002) utilised an
25 experimental design, making comparisons between a control group who experienced a more
26 conventional teaching approach with a group which received some form of intervention
27 related to small-group discussion work. One study (Tolmie and Howe, 1993) specifically set
28 up groups where gender was a variable to be explored. However, the emphasis of the
29 majority of the studies was on describing and interpreting the nature of student discussions
30 and their effects on students' understanding, sometimes making detailed comparisons
31 between groups participating in their studies. Two factors may contribute to the absence of a
32 control group in the studies. Firstly, those undertaking the research might see no need to
33 design their studies to include a control group in what were largely qualitative and
34 interpretative studies. Secondly, the practicalities associated with collecting and analysing
35 extensive in-depth data from a much larger sample in order to make such comparisons would
36 place prohibitive resource constraints on the studies.
37
38
39
40
41
42
43
44
45
46
47
48

49 Data analysis was characterised positively by the presentation and discussion of rich and
50 detailed data in the form of extracts from students' discourse, with all studies adopting
51 procedures to increase the trustworthiness of the analysis by having two or more people
52 involved. However, given that the studies were largely gathering qualitative data, there was a
53 surprising lack of contextual detail. Data also tended to be presented in a rather convergent
54 manner, with few examples of data being presented which might disprove assertions or report
55 on unintended outcomes.
56
57
58
59
60

1
2
3 Of particular interest were the two very contrasting overall analysis strategies apparent in the
4 studies. The first strategy, adopted by the majority of the studies, was to develop grounded
5 theory from the data through the development of categories then used to characterise the
6 interactions between participants. Hogan (1999a) referred to her analysis as ‘ethnographic
7 interaction analysis’, whilst Roth and Roychoudhury (1992) used what they described as the
8 techniques used by anthropologists when analysing interactive behaviours. The second
9 strategy, whose use was more limited, was to draw on existing work on discourse analysis or
10 discourse analysis classifications. Such an approach involves trying to identify themes in
11 what people say by looking at sentences, groups of sentences or sentence fragments. These
12 might, for example, relate to attempts to cite others to support a view, or use of evidence to
13 support an account of an event. Discourse analysis techniques were employed by Jiménez-
14 Aleixandre *et al.* (2000), who drew on the work of Bloome *et al.* (1989) to do the initial
15 coding of exchanges between students, and then used Toulmin’s (1958) work on argument to
16 classify the interactions where students were talking about science aspects in the discussion.
17 Keys (1997) drew on elements of a framework developed by Kuhn (1993) to code students’
18 verbal interactions relating to scientific reasoning. However, there was a notable absence of
19 justification in the studies for the approach adopted for analysis, with the development of
20 grounded theory appearing to be seen as an unproblematic choice in the majority of cases. It
21 may be the case that the choice of approach reflects the personal views of the researchers on
22 the role and purposes of data in research. However, the lack of reference made to discourse
23 analysis techniques suggests that these approaches may be unfamiliar to some researchers
24 working in the area of small-group discussions, which, in turn, may be limiting the nature of
25 the analysis. There would appear to be a good case for those researching the effects of small-
26 group discussions to gain a greater familiarity with discourse analysis techniques.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 Table 4 summarises the key findings to emerge from studies on the use and effects of small-
49 group discussions in science.
50

51
52
53 [Table 4 about here.]
54
55

56 **Conclusions and recommendations**

57
58
59

60 The reviews reported here have yielded insights on both the substantive focus (small-group discussions and their effects) and on the methods employed to gather the data. The review

1
2
3 has revealed a number of features of particular interest in relation to the use of small-group
4 discussion work in science. It is clear from the review that a complex and interacting set of
5 factors are involved in enabling students to engage in dialogue in a way that could help them
6 draw on evidence to develop and articulate their understanding of science ideas.
7
8
9

10
11
12 Two particularly strong features which have emerged from the work undertaken for the
13 review is that there is a relatively little good quality systematic research on the effects of
14 small-group discussion work, and considerable uncertainty on the part of teachers as to what
15 they are required to do to implement good practice. Current policy is strongly advocating the
16 use of small-group discussion work in science, and the reviews do indicate that there could be
17 benefits arising from this, as small-group discussion work can provide an appropriate vehicle
18 for assisting in the development of students' understanding of science ideas. Thus teachers
19 should be encouraged to incorporate such discussions into their teaching. However, it is also
20 clear that small-group discussion work needs to be supported by the provision of support and
21 guidance for teachers and students on the development of the skills necessary to make such
22 work effective.
23
24
25
26
27
28
29
30
31
32

33 One feature, notable by its absence, was the dearth of systematic evidence on the effects of
34 the use of small-group discussions on students' attitudes to their science lessons or science
35 more widely. The absence of such data was very surprising, as the motivational effects of
36 small-group discussions are often cited as a reason for their inclusion in science teaching.
37
38
39
40
41

42 It is clear from this review that there is considerable variation in the nature of research into
43 small-group discussion work, particularly in relation to its focus, the clarity with which any
44 variables being investigated are specified, the use of opportunistic samples for data collection,
45 and the techniques used to analyse data. Particularly striking are the two very contrasting
46 approaches to data analysis, with some studies developing grounded theory from the data, and
47 others drawing on existing models to structure their analysis. A substantial proportion of the
48 work also focuses on descriptive data. This can be very helpful in the early stages of a new
49 research area. However, with increasing interest in the *effects* of small group discussions – on
50 student learning, understanding, and attitudes – there is a need to consider what strategies and
51 techniques lend themselves best to the gathering and analysis of data that would help explore
52 such effects.
53
54
55
56
57
58
59
60

1
2
3 Taken together, these findings suggest there are four key features that need to characterise any
4 further developments in the use of small-group discussions in science teaching and research
5 into their effects. Firstly, some form of professional development training for teachers is
6 highly desirable to provide guidance on how to maximise the effectiveness of small-group
7 discussions. Secondly, further research into the effects of small-group discussions should
8 include a consideration of the extent to which analysis of the data might benefit from
9 established discourse analysis techniques developed in other subject areas (e.g. Barnes and
10 Todd, 1997; Mercer and Littleton, 2007), to establish what they might have to offer work in
11 science. Thirdly, the area would benefit from a more detailed exploration of the effects of
12 small-group discussions on attitudinal effects. Finally, in relation to providing evidence of
13 the *effects* of small-group discussions, there would appear to be potential benefits associated
14 with adopting a mixed method approach to data collection, marrying in-depth qualitative data
15 on the nature of discussions with more quantitative data on student attributes.
16
17
18
19
20
21
22
23
24
25
26
27

28 **Acknowledgements**

29
30
31 The work reported in this paper was funded by the Department for Education and Skills
32 (DfES) as part of the *Evidence, Policy and Practice Initiative* (EPPI).
33
34
35
36

37 **Electronic searches conducted for the review**

38 Electronic searches of the following databases were conducted:
39
40
41

42 British Education Index (BEI)

43 Educational Resources Information Center (ERIC)

44 PsycINFO

45 Social Science Citation Index (SSCI)
46
47
48
49
50

51 **References in text of report**

52
53
54
55 Barnes, D. & Todd, F. (1997) *Communication and learning in small groups*. London:
56 Routledge and Kegan Paul.

57
58 Bennett, J., Lubben, F., Hogarth, S. & Campbell, B. (2005) Systematic reviews of research
59 in science education: rigour or rigidity? *International Journal of Science Education*, 27
60 (4), 387-406.

- 1
2
3 Bentley, D & Watts, M. (1992) *Communicating in school science*. London: Falmer Press.
- 4
5 Black, P. & Wiliam, D. (1998) *Inside the black box: raising standards through classroom*
6
7 *assessment*. London: King's College.
- 8
9 Davies, P. (2000) The relevance of systematic reviews to educational policy and practice.
10
11 *Oxford Review of Education*, 26, 365-378.
- 12
13 Daws, N. & Singh, B. (1996) Formative assessment: to what extent is its potential to enhance
14
15 pupils' science being realised? *School Science Review*, 77 (281), 93-100.
- 16
17 Department for Education and Employment (DfEE) (1998) *The National Literacy Strategy*.
18
19 London: DfEE.
- 20
21 Driver, R. & Bell, B. (1985) Students' thinking and the learning of science: a constructivist
22
23 view. *School Science Review*, 67 (240), 443-456.
- 24
25 Driver, R., Asoko, H., Leach, J., Mortimer, E. & Scott, P. (1994) Constructing scientific
26
27 knowledge in the classroom. *Educational Researcher*, 23 (7), 5-12.
- 28
29 EPPI-Centre (2002) *EPPI-Centre Review Group Manual*. London: Social Science Research
30
31 Unit, Institute of Education.
- 32
33 Hargreaves, D. (1996) *Teaching as a research-based profession: possibilities and prospects*.
34
35 London: The Teacher Training Agency (TTA).
- 36
37 Hillage, L., Pearson, R., Anderson, A. & Tamkin, P. (1998) *Excellence in research on*
38
39 *schools*. Brighton: Institute for Employment Studies.
- 40
41 Kuhn, D. (1993) Science as argument: implications for teaching and learning scientific
42
43 thinking. *Science Education*, 77 (2), 319-337.
- 44
45 Kyriacou, C. (1998) *Essential teaching Skills* (2nd edition) Cheltenham: Stanley Thornes.
- 46
47 Lemke, J. (1990) *Talking science*. New York: Ablex Publishing Corporation.
- 48
49 Levinson, R. & Turner, S. (2001) *Valuable Lessons: Engaging with the Social Context of*
50
51 *Science in Schools*. London: The Wellcome Trust.
- 52
53 Mercer, N. & Littleton, K. (2007) *Dialogue and the Development of children's thinking: a*
54
55 *sociocultural approach*. Oxford: Routledge.
- 56
57 Millar, R. (2006). *Twenty First Century Science: Insights from the design and*
58
59 *implementation of a scientific literacy approach in school science*. *International Journal*
60
of Science Education, 28 (13), 1499-1521.
- 61
62 Millar, R. & Osborne, J. (eds) (1998) *Beyond 2000: Science Education for the Future*.
63
64 London: King's College/The Nuffield Foundation.
- 65
66 Millar, R., Leach, J., Osborne, J. & Ratcliffe, M. (2006) *Improving subject teaching: lessons*
67
68 *from research in science education*. London: Routledge.

- 1
2
3 Newton, P., Driver, R. & Osborne, J. (1999) The place of argumentation in the pedagogy of
4 school science. *International Journal of Science Education*, 21 (5), 553-576.
5
6 Oakley, A. (2002) Social science and evidence-based everything: the case of education.
7
8
9 *Educational Review*, 54(3), 21-33.
10
11 OECD (2002) *Educational research and development in England: examiners' report*.
12
13 Organisation for Economic Co-operation and Development. www.oecd.org/dataoecd/17/56
14
15 (accessed 2 January 2008)
16
17
18 Osborne, J. & Collins, S. (2001) Pupils' views of the role and value of the science curriculum.
19
20 *International Journal of Science Education*, 23 (5), 441-467.
21
22 Osborne, J., Duschl, R. & Fairbrother, R. (2002) *Breaking the Mould? Teaching Science for*
23
24 *Public Understanding*. London: The Nuffield Foundation.
25
26 Scott, P. (1998) Teacher talk and meaning-making in science classrooms: a Vygotskian
27
28 analysis and review. *Studies in Science Education*, 32 45-80.
29
30 Shavelson, R. & Towne, L. (2001) (eds) *Scientific enquiry in education*. Washington DC:
31
32 National Academy Press.
33
34 Sinclair, J. and Coulthard, R. (1975) *Towards an Analysis of Discourse*, OUP.
35
36 Spencer, L., Ritchie, J., Lewis, J. & Dillon, L. (2003). *Quality in Qualitative Evaluation: a*
37
38 *framework for assessing research evidence*. London: The Strategy Unit.
39
40 Sutton, C. (1992) *Words, science and learning*. Buckingham: Open University Press.
41
42 Toulmin, S. (1958). *The use of argument*. Cambridge: Cambridge University Press.
43
44 Torgerson, C. (2003) *Systematic Reviews*. London: Continuum.
45
46 Torgerson, C. & Torgerson, D. (2001) The need for randomised controlled trials in
47
48 educational research. *British Journal of Educational Studies*, 49 (3), 316-328.
49
50 Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*.
51
52 Buckingham: Open University Press.

References to the full systematic review technical reports

- 53
54
55
56 Author 1, Author 3, Author 2, Author 4, & Author 5. (2004) A systematic review of the
57
58 nature of small-group discussions in science teaching aimed at improving students'
59
60 understanding of evidence. In: *Research Evidence in Education Library*. London: EPPI-

Centre, Social Science Research Unit, Institute of Education.

<http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=321>

Author 1, Author 3, Author 2 & Author 4. (2004) A systematic review of the use of small-group discussions in science teaching with students aged 11-18, and their effects on students' understanding in science or attitude to science. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education. <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=326>

References for the studies included in the review

De Vries E, Lund K, Baker M (2002) Computer-mediated epistemic dialogue: explanation and argumentation as vehicles for understanding scientific notions. *Journal of the Learning Sciences* 11: 63-103.

Finkel EA (1996) Making sense of genetics: students' knowledge use during problem solving in a high school genetics class. *Journal of Research in Science Teaching* 33: 345-368.

Gayford C (1995) Science education and sustainability: a case-study in discussion-based learning. *Research in Science & Technological Education* 13: 135-45.

Hogan K (1999a) Sociocognitive roles in science group discourse. *International Journal of Science Education* 21: 855-882.

Hogan K (1999b) Thinking aloud together: a test of an intervention to foster students' collaborative scientific reasoning. *Journal of Research in Science Teaching* 36: 1085-1109.

Jiménez-Aleixandre MP, Pereiro-Muñoz C (2002) Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education* 24: 1171-1190.

Jiménez-Aleixandre MP, Rodriguez AB, Duschl RA (2000a) 'Doing the lesson' or 'doing science': argument in high school genetics. *Science and Education* 84: 757-92.

Johnson SK, Stewart J (2002) Revising and assessing explanatory models in a high school genetics class: a comparison of unsuccessful and successful performance. *Science and Education* 86: 463-480.

Keys CW (1997) An investigation of the relationship between scientific reasoning, conceptual knowledge and model formulation in a naturalistic setting. *International Journal of Science Education* 19: 957-970.

- 1
2
3 Kurth LA, Anderson CW, Palincsar AS (2002) The case of Carla: dilemmas of helping all
4 students to understand science. *Science Education* 86: 287-313.
5
6 Lajoie SP, Lavigne NC, Guerrero C, Munsie SD (2001) Constructing knowledge in the
7 context of BioWorld. *Instructional Science* 29: 155-186.
8
9 Lavoie DR (1999) Effects of emphasizing hypothetico-predictive reasoning within the science
10 learning cycle on high school student's process skills and conceptual understandings in
11 biology. *Journal of Research in Science Teaching* 36: 1127-1147.
12
13 Meyer K, Woodruff E (1997) Consensually driven explanation in science teaching. *Science*
14 *Education* 81: 173-192.
15
16 Palincsar AS, Anderson C, David YM (1993) Pursuing scientific literacy in the middle grades
17 through collaborative problem solving. *Elementary School Journal* 93: 643-658.
18
19 Richmond G, Striley J (1996) Making meaning in classrooms: social processes in small-group
20 discourse and scientific knowledge building. *Journal of Research in Science Teaching* 33:
21 839-858.
22
23 Roth W-M, Roychoudhury A (1992) The social construction of scientific concepts or the
24 concept map as conscription device and tool for social thinking in high school science.
25 *Science and Education* 76: 531-557.
26
27 Sherman GP, Klein JD (1995) The effects of cued interaction and ability grouping during
28 cooperative computer-based science instruction. *Educational Technology Research and*
29 *Development* 43: 5-24.
30
31 Suthers D, Weiner A (1995) Groupware for developing critical discussion skills. In: Schnase
32 JL, Cunniss EL *Proceedings of CSCL '95: The First International Conference on*
33 *Computer Support for Collaborative Learning*. New Jersey: Lawrence Erlbaum
34 Associates Inc, pages 341-348.
35
36 Tao PK (2001) Developing understanding through confronting varying views: the case of
37 solving qualitative physics problems. *International Journal of Science Education* 23:
38 1201-1218.
39
40 Tao PK (2003) Eliciting and developing junior secondary students' understanding of the
41 nature of science through a peer collaboration instruction in science stories. *International*
42 *Journal of Science Education* 25: 147-171.
43
44 Tolmie A, Howe C (1993) Gender and dialogue in secondary school physics. *Gender and*
45 *Education* 5: 191-209.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Tsai C-C (1999) 'Laboratory exercises help me memorize the scientific truths': a study of
4 eighth graders' scientific epistemological views and learning in laboratory activities.
5
6 *Science and Education* 83: 654-674.
7

8
9 Williams A (1995) Long-distance collaboration: a case study of science teaching and
10 learning. In: Spiegel SA *Perspectives from Teachers' Classrooms. Action Research.*
11 *Science FEAT (Science for Early Adolescence Teachers).*
12

13
14 Woodruff E, Meyer K (1997) Explanations from intra- and inter-group discourse: students
15 building knowledge in the science classroom. *Research in Science Education* 27: 25-39.
16

17 Zohar A, Nemet F (2002) Fostering students' knowledge and argumentation skills through
18 dilemmas in human genetics. *Journal of Research in Science Teaching* 39: 35-62.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

Abstract

This paper reports the findings of two systematic reviews of the use and effects of small-group discussions in high school science teaching. 94 studies were included in an overview (systematic map) of work in the area, and 24 studies formed the basis of the in-depth reviews.

The reviews indicate that there is considerable diversity in the topics used to promote small-group discussions. They also demonstrate that students often struggle to formulate and express coherent arguments, and demonstrate a low level of engagement with tasks. The reviews suggest that groups function more purposefully, and understanding improves most, when specifically constituted such that differing views are represented, when some form of training is provided for students on effective group work, and when help in structuring discussions is provided in the form of 'cues'. Single sex groups function more purposefully than mixed sex groups, though improvements in understanding are independent of gender composition of groups. Finally, the reviews demonstrate very clearly that, for small-group discussions to be effective, 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.

In addition to the substantive findings, the paper also reports on key features of the methods employed to gather and analyse data. Of particular note are the two contrasting approaches to data analysis, one adopting a grounded theory approach and the other drawing on established methods of discourse analysis.

Introduction

The use of small-group discussions in teaching has been advocated for a number of years in science teaching, both to motivate students and to enhance their learning. Such tasks are now appearing with greater frequency in a range of high school science teaching resources. Despite this increasing popularity, comparatively little is known about the detail of their use and effects. This paper examines the research evidence on the use of small-group discussions

1
2 through presenting the findings of two systematic reviews, with a view to making
3 recommendations for policy and practice on the use of small-group discussions in science
4 teaching.
5
6

7
8 As originally conceived, the reviews had two principal aims. The first of these was to identify
9 the ways in which small-group discussions are currently used in science lessons. The second
10 was to look at the effects of small group discussions on students' understanding of science
11 and on students' attitudes to science. In practice, the review work established that there is a
12 dearth of studies reporting in any detailed and systematic way on the effects of small-group
13 discussions on students' attitudes to science, so this paper reports the review findings on the
14 use of small-group discussions in science teaching and their effects on students'
15 understanding of science ideas.
16
17
18
19

20
21 Several factors have contributed to the current high levels of interest in small-group
22 discussion work in science. The use of small-group discussion in teaching has its origins in
23 learner-centred teaching approaches, and is one of a range of 'active learning' strategies
24 aimed at stimulating students' interest in what they are studying by providing them with a
25 significant degree of autonomy over the learning activity (e.g. Bentley and Watts, 1992;
26 Kyriacou, 1998). A number of people have advocated the use of discussions in science
27 lessons. Lemke (1990) argues that "learning science means learning to talk science" (p1), and
28 that this means moving away from science lessons dominated by teacher talk in which the
29 teacher asks a question, then invites and evaluates a student response. Lemke refers to this as
30 'triadic dialogue', which is similar to the initiation-response-feedback (IRF) sequence
31 characterised earlier by Sinclair and Coulthard (1975). Lemke and others (e.g. Sutton, 1992,
32 and Wellington and Osborne, 2001) criticise the approach for leading to talk that focuses on
33 what the teacher wants to hear, rather than promoting genuine communication, and all argue
34 for increased use of discussion work in science lessons.
35
36
37
38
39
40
41
42

43 Support for the use of small-group discussions in science teaching has also emerged in the
44 recommendations from other areas of work in science education. For example, research on
45 alternative conceptions has explored in depth the ideas and understanding students bring with
46 them to science lessons and the ways in which some of their ideas may hinder the
47 development of accepted scientific ideas (e.g. Driver *et al.*, 1985). Small-group discussions
48 have been suggested as a means of helping students explore their ideas and move from
49
50
51
52

Deleted: the constructivist view of learning

1 understandings that may often be naïve to towards more valid scientific ideas and
 2 explanations. Further impetus for the inclusion of small-group discussions in science lessons
 3 has come from the development of ideas about *social constructivism* (Driver *et al.*, 1994),
 4 which draws on the work of Vygotsky and the importance to learning of social dynamics of
 5 interaction ([Scott, 1998](#)).
 6
 7

8
 9
 10 The publication of *Beyond 2000* (Millar and Osborne, 1998) [in the UK](#) stimulated discussion and
 11 debate over the nature of the school science curriculum and, in particular, the ways in which it
 12 might foster the development of *scientific literacy*. This term embraces the knowledge,
 13 understanding and skills young people need to develop in order to think and act appropriately on
 14 scientific matters which may affect their lives and the lives of other members of the local,
 15 national and global communities of which they are a part. ~~Many of the materials developed~~
 16 include small-group discussions in the repertoire of activities employed in science lessons in
 17 order to encourage students to participate in informed discussion and debate of scientific issues
 18 [\(e.g. Millar, 2006\)](#). ~~Linked to the development of scientific literacy is that of ideas about~~
 19 [evidence \(Millar and Osborne, 1998\)](#), which involves encouraging students to evaluate, interpret
 20 and analyse evidence from primary and secondary sources in science, including stories about how
 21 important science ideas were first developed, and then established and finally accepted. This has
 22 led to considerations of the role of *argument* in school science, in the sense of putting forward
 23 claims and supporting them with sound and persuasive evidence (e.g. Newton *et al.*, 1999;
 24 Osborne *et al.*, 2001). Small-group discussions are viewed as having a key role to play here,
 25 since the practice of using evidence in argumentation requires interaction with peers.
 26
 27
 28
 29
 30
 31
 32
 33
 34

Formatted: Right: -18 pt

Deleted: There are now moves in a number of countries to introduce science curricula with a strong emphasis on scientific literacy.

Deleted: .

Deleted: .

Formatted: Font: (Default) Arial, 11 pt

35
 36 In the UK, two other developments have also served to raise the profile of small-group
 37 discussions. Firstly, the publication of *Inside the Black Box* (Black and Wiliam, 1998) has
 38 resulted in considerable attention being paid to formative assessment, or assessment for
 39 learning. Small-group discussions are one approach that has been advocated for increasing
 40 the use and effectiveness of formative assessment in science teaching (see, for example, Daws
 41 and Singh, 1999). Secondly, there is a more general drive to improve students' *literacy skills*.
 42 In England and Wales, this has been formalised into the National Literacy Strategy (DfEE,
 43 1998), ~~where~~ small-group discussions have been advocated as a means for developing
 44 students' [oral communication skills](#) in science.
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

Deleted: . S

Deleted: language skills

Deleted: (e.g. Newton *et al.*, 1999, and Osborne *et al.*, 2001).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Whilst the preceding discussion point to a general recognition of the potential value of small-group discussion in science teaching, their use challenges the established pedagogy of science teaching and place new demands on teachers (see, for example, Bentley and Watts, 1992).

There is a growing body of evidence, both anecdotal and more systematic, that many teachers lack skills and do not feel confident with small-group discussions. In particular, evaluation of materials with a specific focus on teaching socio-scientific issues and developing scientific literacy, such as *AS Public Understanding of Science* course (Osborne *et al.*, 2002) and the *Valuable Lessons* project (Levinson and Turner, 2001), raised particular concerns about teacher efficacy in the use of small-group discussions. For example, Osborne *et al.* (2002) comment, “Our view, given the poor quality of much of the teaching involving discussion, is that training is essential. Teachers ... need an opportunity to interact with experienced humanities teachers, to observe the strategies they use for fostering and stimulating discussion ...” (p75). In a similar vein, Levinson and Turner recommended that, “Teacher training courses should provide prospective science teachers with more opportunities in the area of initiating and managing discussions ... practicing teachers should improve and update such skills through CPD [continuing professional development].” (p21).

It is clear from the discussion above that there is considerable interest in the use of small-group discussion in science teaching. Some of this interest has emerged directly from research studies, whilst, in other areas, it draws 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. There also appears to be a comparative lack of guidance for teachers. The evidence presented in this paper provides a number of insights into the form such guidance might take.

Deleted: widespread

The origins and purposes 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).

Deleted: ¶

A note on terminology¶

¶
The reports on small-group discussions scrutinised for the systematic reviews made it clear that the term was used in a wide variety of ways. For the purposes of the reviews, 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; structured teacher input);¶
<#>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).¶
¶
The term *understanding* has been taken to encompass understanding of science concepts, understanding of ideas about the nature of science and understanding of the methods of science.¶

1
2 There are several reasons why systematic reviews are being seen as a key strand in
3 educational research. Firstly, there is a growing interest in practical policy-related decision
4 making being linked to evidence in a number of areas, not just in education. Systematic
5 reviews of research literature are seen as having the potential to yield evidence on which
6 policy makers can draw (Davies, 2000; Torgerson and Torgerson, 2001). Secondly, there is a
7 drive towards forging closer links between research, policy and practice (see, for example,
8 Hargreaves, 1996; OECD, 2002; Oakley, 2002). In particular, drawing on research findings
9 in classroom practice is seen as desirable, with teachers being encouraged to engage in what is
10 variously described as 'evidence-based', 'evidence-informed' or 'evidence-enriched' practice
11 (e.g. Millar *et al.* 2006).

Formatted: Font: Italic

12
13
14
15
16
17
18 In 2000 the Government in the UK funded, via the Department for Education and Skills
19 (DfES), the Evidence for Policy and Practice Initiative (EPPI)-Centre to focus on undertaking
20 systematic reviews of research evidence in key areas of education, and reporting these in a
21 form accessible to a range of different user groups, including teachers, researchers and policy-
22 makers. The advantages and disadvantages of systematic reviews when compared with other
23 forms of review, and of reviews employing the EPPI methodology have been rehearsed
24 elsewhere (Bennett *et al.*, 2005).

Formatted: Font: Italic

25
26
27
28
29
30 The systemic review process, as developed by the EPPI-Centre, involves several stages:

- 31 • identification of review topic area and review research question or questions;
- 32 • development of *inclusion and exclusion criteria* for studies in the review (relating to, for
33 example, aspects such as the age of students, the nature of the research design, and the
34 reported outcomes);
- 35 • undertaking of *systematic searches* of electronic data bases and other sources for
36 potentially relevant research studies;
- 37 • refining the search through *screening* the potentially relevant studies against the inclusion
38 criteria;
- 39 • coding or *keywording* studies against pre-specified and agreed characteristics (some of
40 which are generic to all EPPI reviews, whilst others are developed specifically for each
41 review);
- 42 • production of an overview or *systematic map* of studies in the review area, that groups the
43 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 evidence reported.

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*

Full details of the EPPI methodology may be found in the EPPI Review Group Manual (EPPI-Centre, 2002).

The EPPI review methods tend to be more tailored to quantitative research. As much of the work on small-group discussion draws extensively on qualitative approaches, the work reported here extended the EPPI review methods to draw on the guidance and framework for assessing research evidence in qualitative research studies (Spencer *et al.*, 2003). This was seen as particularly important as the majority of the studies included in the review made use of qualitative methods. As such, they ran the risk of being characterized as low in quality had judgments been informed only by the EPPI criteria.

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

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?* Within this, two reviews were conducted. The purpose of the first review was to gain an overview of the nature of small-group discussions used in science lessons in order to assess the levels of use, establish any patterns in use, such as in topic focus, group structure and the dynamics of group interactions. As much of the support for the use of small-group discussions is linked to their potential benefits associated with to improved understanding of science ideas, the second review focused specifically on the effects of small-group discussions on understanding.

A note on terminology

Deleted: in relation

Deleted: students'

Formatted: Font: Not Bold, Italic

The reports on small-group discussions scrutinised for the systematic reviews made it clear that the term was used in a wide variety of ways. For the purposes of the reviews, 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; structured teacher input);
- involves a substantive discussion task of at least two minutes (i.e. did not simply involve a student talking to a neighbour briefly about an idea or to agree and answer to a question);
- 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).

Formatted: Bullets and Numbering

The term *understanding* has been taken to encompass understanding of science concepts, understanding of ideas about the nature of science and understanding of the methods of science.

Studies in the review

Formatted: Font: Italic

Deleted: ¶

The reviews focused on research on teaching at high school level, undertaken in the period 1980-2005, and published in English. Student age was restricted to 11-18 because this is the age range covered by the majority of reported studies, pointing to this being the school age range where the use of small-group discussions has been promoted most actively. The start date for the period of publication was selected because this was the time when the use of small-group discussions started to become more prominent in science teaching. The inclusion criteria for studies in the review were: (i) they were about the use of small-group discussions in science lessons, (ii) they involved groups of two to six students, (iii) they focused on a substantive, structured discussion task of two minutes' duration or more, and, for the second review, (iv) they addressed aspects of students' understanding in science.

Deleted: is

Deleted: ,

Deleted: and continues to be,

Deleted: , and hence is the age range covered by the majority of reported studies.

Deleted: or attitudes to science

94 studies were identified that met the above inclusion criteria. The identification of the studies is a two-step process. Firstly, systematic searches are undertaken, principally through the use of electronic search strings. Electronic searching inevitably means that large numbers of studies emerge in the initial stage, and some 2,290 studies matched the search terms. The second step involves refining the search through careful screening of abstracts (or full copies

Deleted: re

1
2 of reports if there is insufficient information in the abstract) against the inclusion criteria. The
3 94 studies identified were then coded for particular characteristics (*keyworded*) to produce an
4 overview (*the systematic map*) of studies of small-group discussion interventions. In
5 producing the map, the following characteristics of studies scrutinised included the country of
6 study, the age/level of the students, the type of study, the science discipline of the study,
7 constitution of discussion groups, the duration of discussion tasks, the stimulus provided,
8 organisational features of discussion tasks, the product of discussion tasks, the research
9 strategy used to gather data, the nature of data collected, and the outcomes reported.
10
11
12
13
14

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

16

17
18 This section presents a brief overview of the key features to emerge from the 94 research
19 reports on the use and effects of small-group discussions.
20
21

22
23 Although the period of the review covered 1980-2005, over 90% were from post-1990,
24 indicating that that the research activity area has been minimal up to fifteen years ago and has
25 been most prolific from the late 1990s onwards. The majority of the reported work has been
26 undertaken in North America (USA = 39%, Canada = 12%) and the UK (13%). Other work
27 has been undertaken in Australia, The Netherlands and Germany. However, these figures
28 need to be set in the context of the review being limited to reports published in English,
29 though the review does include reports of studies of small-group discussions held in Bahasa
30 Malay, Cantonese, Dutch, Finnish, French, German, Greek, Hebrew, Mandarin, Portuguese
31 and Spanish.
32
33
34
35

36
37 Relatively little detail was given about how groups were constituted, but where this was
38 provided, mixed-ability and friendship groups predominated. In one-third of cases, groups
39 were deliberately constituted by the teacher, with students choosing their groups in the
40 remaining cases. Slightly over half the studies used a group size of 3-4 students, with a
41 further quarter of studies using pairs of students for groups. Over 80% of the studies
42 concerned self-contained and permanent groups. The remaining studies drew on the
43 techniques of 'snowballing', 'envoying' and 'jigsawing'. Snowballing discussions involve
44 progressively larger groups of students discussing a question or idea and agreeing on their
45 views. Discussion starts with pairs, who then join together and so on. Envoying discussions
46 involve groups of students discussing a common task. When the discussion is completed, one
47
48
49
50
51
52

1
2 member of each group moves to another group to report on the discussion of the original
3 group and to hear about the discussion in the second group. The envoy then returns to the
4 original group to report back. Jigsawing discussions have students first working in groups
5 where each has a different task and then moving into different groups that comprise of all the
6 people who have focused on the same aspect of the task.
7
8
9

10
11 Two-thirds of the discussions took a class period or longer, with approximately two-fifths
12 taking place in lessons described as 'science', a similar proportion in physics lessons, around
13 one-fifth in biology lessons and only 4% in chemistry lessons. One possible explanation for
14 this very small figure for chemistry is that most of the small-group discussions relating to
15 exploration of difficult ideas are located mainly within physics, and those developing skills of
16 decision-making on socio-scientific issues are more commonly placed within biology classes.
17
18
19

20
21 The most common stimulus material provided for the discussion tasks was printed
22 worksheets, which were used in over two-thirds of the studies. Practical work and computer
23 software provided the stimulus for around two-fifths and a quarter of discussions respectively.
24 Some discussion made use of more than one stimulus. Rather surprisingly, only one study
25 used a newspaper article as a stimulus for discussion.
26
27
28

29
30 In a very high proportion of the studies (94%), the main aim of the discussion task was
31 individual understanding of the science underlying the activity, such as in a practical
32 experiment, the preparation of a poster or a computer-based exercise, in which the learners
33 were engaged. In the majority of cases this understanding was then shared with classmates in
34 different ways: groups might present their findings or views orally (20%) or by way of posters
35 (10%) or might defend their position in a whole class debate (5%). There were surprisingly
36 few examples of written products being generated directly as an outcome of the discussion
37 (6%).
38
39
40
41
42

43
44 Around three-fifths of the studies reported on evaluations of small-group discussions with the
45 remainder providing descriptive information about the use of small-group discussions. Case
46 studies featured prominently, with extensive use being made of video and audio recordings in
47 order to gather detailed information about the nature of discussions. One outcome of the very
48 labour-intensive nature of much of the data collection and analysis was that sample sizes
49
50
51
52

1
2 tended to be small – very often one class or one or two groups of students within a class.
3
4 Studies involving several classes, or classes in more than one school, were rare.
5

6
7 The chief characteristics of research on small-group discussions are summarised in Table 1.
8

9
10 [Table 1 about here.]
11

12 **The in-depth reviews**

13
14
15 Two in-depth reviews were conducted, covering a total of 24 of the 94 studies in the
16 systematic map. These 24 studies were selected because (i) they reported in detail on the use
17 of small-group discussions in science lessons (19 studies), and/or (ii) they reported
18 evaluations of interventions aimed at developing aspects of students' understanding in science
19 (14 studies).
20
21
22

23
24 Studies were rated as high (H), medium high (MH), medium (M), medium low (ML) or low
25 (L). As quality judgements in systematic reviews are made in relation to the specific focus of
26 each of the reviews, some studies were given different ratings in each review. Ratings were
27 based on the extent to which the studies reported met a range of criteria relating to (i) the
28 nature of the sample and how it was selected; (ii) the nature of any control group (for
29 evaluations); (iii) the extent to which small-group discussions formed the main feature or
30 variable being investigated (for evaluations); (iv) the level of detail provided about the
31 discussion task; (v) the steps taken to establish the reliability and validity of the data
32 collection tools and processes, and the data analysis; (vi) the trustworthiness and reliability
33 of the data collection and analysis for qualitative data; (vii) the representativeness of the data
34 collection situation to normal classroom situations; (viii) the quality of the reporting. No
35 studies met all the relevant criteria, so none were judged to be high quality. The studies
36 rated as 'medium high' met most of the relevant criteria. Studies rated at the lower end of the
37 scale displayed one or more of the following characteristics: they provided comparatively
38 little detail of the discussion tasks, they tended to be weaker in the reporting of steps taken to
39 enhance the reliability and validity of data collection and analysis, and/or they were overly-
40 descriptive at the expense of analysis and discussion. The categorisation process is described
41 in detail in the full technical reports (see References section).
42
43
44
45
46
47
48
49
50
51
52

1
2 Table 2 lists the studies and their quality rating for each review.
3
4

5 [Table 2 about here.]
6
7

8 The remaining discussion focuses on the nineteen studies rated as 'medium' or better, as these
9 provide the stronger evidence, and Table 3 summarises the key features of each of these
10 studies.
11

12
13 [Table 3 about here.]
14
15

16 17 **The detailed evidence from the in-depth reviews** 18

19
20 The consolidated evidence presented draws primarily on the findings of the fourteen studies
21 in the Review 1 and the twelve studies in Review 2 weighted as medium or better in overall
22 quality, as these studies generated the better quality evidence. Of these studies, nine were
23 common to both reviews, i.e. they focused on aspects of the use of small-group discussions
24 and the development of understanding. The features discussed below are considered to be
25 those that offer the strongest evidence as they emerged from three or more studies. To avoid
26 undue repetition, evidence from both reviews is considered together under the following five
27 headings: focus of the discussion topic, group structures and interactions, negotiating and
28 agreeing meaning through discussion, effects on understanding, factors promoting effective
29 discussion to enhance understanding
30
31
32
33
34

35 36 *Focus of the discussion topic* 37 38

39 The studies were based on a range of science topics, as Table 3 demonstrates. Seven studies
40 addressed aspects of understanding of science topics. These focused on light (Roth and
41 Roychoudhury, 1992), kinetic theory (Palincsar *et al.*, 1993), genetics (Finkel, 1996; Jiménez-
42 Aleixandre *et al.*, 2000), two physics topics [shadows, floating and sinking] (Woodruff and
43 Meyer, 1997), mechanics (Tao, 2001), and density (Kurth *et al.*, 2002). Four looked
44 primarily at aspects of what could be termed scientific method: hypotheses on the diagnosis of
45 disease (Richmond and Striley, 1996; Lajoie, 2001), designing controlled experiments
46 (Sherman and Klein, 1995), and building theories and models from primary evidence on
47 elements and bonding (Keys, 1997). Three had a specific focus on socio-scientific issues in
48
49
50
51
52

1
2 relation to the greenhouse effect (Gayford, 1995), genetic engineering (Zohar and Nemet,
3 2002), and environmental science (Jiménez-Aleixandre and Pereiro-Muñoz (2002). Three
4 involved making predictions based on evidence presented in the topics of sound (De Vries *et*
5 *al.*, 2002), a range of biology topics (Lavoie, 1999), and mechanics (Tolmie and Howe,
6 1993). The studies by Hogan (1999a and 1999b) had no specific topic focus but were based
7 on a series of discussion task developed by the researchers.
8
9
10

11
12 The studies in the in-depth review reflect the patterns noted in the systematic map, where the
13 bulk of the discussion topics lay in the areas of physics and biology. Although there was
14 diversity in topic focus, the common link between the discussions focusing on the
15 development of understanding was that they all required students to draw on evidence to
16 support a particular hypothesis, theory or point of view.
17
18
19

20 21 *Group structures and interactions*

22
23 The principal evidence on group structure and interaction came from five studies: Tolmie and
24 Howe (1993), and Richmond and Striley (1996); Keys (1997); Hogan (1999a); De Vries *et*
25 *al.*, 2002.
26
27
28

29
30 Group leadership emerged as crucial in promoting effective discussion. Richmond and
31 Striley (1996) and Kurth *et al.* (2002) established the need for a leader to adopt an inclusive
32 style and share tasks equitably around a group, as this promoted more substantial engagement
33 in the discussion by a number of participants, and increased the quality of the discussion.
34 This, in turn, permitted most members to develop their understanding. Non-inclusive
35 leadership generated much off-task talk and engagement was generally low. Hogan (1999a)
36 found that at least one group member had to act in a way which promoted reflection in the
37 group for understanding of science ideas to be developed.
38
39
40
41

42
43 The evidence on assigning specific roles to students was mixed. Kurth *et al.* (2002)
44 advocated assigning particular roles to pupils in groups as a means of achieving effective
45 discussion. However, Richmond and Striley, (1996) and (Hogan, 1999a) reported allocating
46 roles to have benefits when tasks were well-structured but counterproductive in poorly-
47 structured tasks, adding to students' difficulties in engaging with the task. Hogan identified
48 eight sociocognitive roles in group reasoning processes which took place in small-group
49
50
51
52

1
2 discussions, all of which were persistent over time, and independent of particular formal roles
3 that might have been allocated to students in advance of a discussion. Four of Hogan's roles
4 were positive (promoter of reflection, contributor of content knowledge, creative model
5 builder, mediator of social interaction and ideas) and four were negative (promoter of
6 distraction, of acrimony, of simple task completion, reticent participant).
7
8
9

10
11 Aspects of group composition were explored in four studies (Tolmie and Howe, 1993; Keys,
12 1997; Hogan, 1999a; De Vries *et al.*, 2002). Tolmie and Howe found improved
13 understanding to be independent of group composition (male, female, mixed), with biggest
14 improvements being noted when groups contained members with a high degree of
15 dissimilarity in their initial predictions and explanations. In common with Keys and De Vries
16 *et al.*, Tolmie and Howe identified clear differences in interactional styles with all-male
17 groups confronting differences in their individual predictions and explanations, whilst all-
18 female groups searched for common features of their predictions and explanations in order to
19 avoid conflict. Mixed groups interacted in a more constrained way than single-gender groups,
20 though they also tended to avoid conflict and look for common patterns in contributions.
21 Tolmie and Howe suggested that both male pairs and female groups demonstrated qualities
22 one would want to see in the development of arguments but did not see this as a reason for
23 promoting the use of mixed gender groups, as their study suggested the best of the all-male
24 and all-female group interactions was lost in mixed pairs. Hogan also found that friendship
25 groups, which were generally single-sex, functioned more effectively and promoted better
26 development of understanding than mixed or teacher-constituted groups.
27
28
29
30
31
32
33
34
35

36 The studies suggest that both the behavioural characteristics and gender composition of
37 groups need careful consideration if groups are to function purposefully during discussion
38 tasks.
39
40
41

42 *Negotiating and agreeing meaning through discussion*

43
44

45 Three of the studies (Roth and Roychoudhury, 1992; Keys, 1997; Jiménez-Aleixandre *et al.*,
46 2000) reported in detail on the ways in which meanings were negotiated and agreed by
47 groups.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 Keys (1997) identified three common characteristics of reasoning in discussions: recognising
3 that prior ideas (models) may be incorrect; evaluating new observations for consistency with
4 current ideas and using evidence to modify ideas; and coordinating all mutually consistent
5 knowledge propositions into a coherent model. A similar pattern was described by Roth and
6 Roychoudhury (1992) who found discussions usually involved positions being stated and
7 contested, with views either accepted or temporarily or permanently rejected as positions
8 finally stabilised into shared meaning. Less positively, they found that students tended not to
9 engage very often in processes which fostered meaning. Rather they would reach agreement
10 on the basis of finding something agreeable to all group members. Agreements were often
11 reached by one or more group members exerting authority, and on the basis of 'majority rule',
12 rather than agreed shared understanding. These findings were echoed by Jiménez-Aleixandre
13 *et al.* (2000) who reported that a large proportion of student talk related to what they termed
14 'doing the lesson', or interactions referring to the rules of the task, rather than talk related to
15 the focus of the task. Jiménez-Aleixandre *et al.* also noted that arguments were frequently
16 developed by a subset within the group and, though agreement was generally reached, this
17 was often for social reasons, rather than because of agreement over the outcome of the
18 discussion.
19
20
21
22
23
24
25
26
27

28 Whilst the above findings suggest that groups can work together to develop common
29 understandings, they also suggest that a number of factors can influence the way in which
30 these understandings are reached, and relative lack of engagement with the science content,
31 coupled with the influence of students prepared to express views very strongly means that the
32 gains in understanding may not be that high.
33
34
35
36

37 *Effects on understanding*

38
39
40 All the studies that looked at development of understanding reported benefits from small-
41 group discussion work, though this finding does need to be interpreted with some caution as
42 the majority of those undertaking the research were advocates of the approach. The following
43 provide examples of the effects that were reported.
44
45
46
47

48 Roth and Roychoudhury (1992) established that group discussion over the construction of a
49 concept map provided a vehicle for negotiation of meaning and understanding of concepts
50 and their relationships, thus providing a structure through which students were able to learn
51
52

1
2 the language patterns of science and use these to construct scientific knowledge. Tolmie and
3 Howe (1993) reported significant improvements in students' predictions of the trajectories of
4 falling objects through the use of a computer-based simulation. Richmond and Striley (1996)
5 noted increasing levels of sophistication and increased use of subject knowledge in the
6 arguments students developed in discussion of socio-scientific issues. Similarly, Zohar and
7 Nemet (2002) reported substantial changes in the quality of student arguments in the context
8 heredity and genetics.

13 14 *Factors promoting effective discussion to enhance understanding*

15
16
17 The factors that emerged that contributed to effective discussions and enhancement in
18 understanding are of particular interest. Findings from four studies pointed to improvements
19 in understanding being greatest for discussion tasks where there was dissimilarity or conflict
20 in understanding or views. This might take the form of either *internal conflict*, or differences
21 held by individual group members (Tolmie and Howe, 1993; De Vries *et al.*, 2002), or
22 *external conflict* where an external stimulus presents a group with conflicting views (Tolmie
23 and Howe, 1993; Gayford, 1995; Finkel, 1996). In some of the studies the discussion topic
24 was selected to provide opportunities for both internal and external conflict. For example,
25 Tolmie and Howe (1993) required students to make individual predictions about aspects of
26 forces and motion, then engage in a task which required a joint prediction (internal conflict)
27 and finally to compare this with an actual situation to reach an explanation of any
28 discrepancies (external conflict). Whilst other studies did not comment specifically on the
29 need for dissimilarity, it was clear from some of the accounts (e.g. Zohar and Nemet, 2002)
30 that internal and external conflict were built into the discussion tasks. It seems likely that the
31 dissimilarity in views provides a very clear and immediate focus to engage students in
32 discussion.

33
34
35 Two studies offered comments on the nature of the data provided to students for the
36 discussion. Jiménez-Aleixandre *et al.* (2000) indicated that hypothetical, unquestionable data
37 (provided by the teacher) generates different patterns of discussion than empirical, uncertain
38 data, perhaps gathered by students themselves, with the former leading to greater gains in
39 understanding. In a similar vein, Roth and Roychoudhury (1992) found discussion to be more
40 productive if students were provided with a fixed set of concepts to delimit the content of the
41 discourse.

1
2
3
4 Three studies pointed to improved understanding when students were given specific
5 instructions on how to construct arguments or cues to guide them in the points they needed to
6 include (Sherman and Klein, 1995; De Vries *et al.*, 2002; Zohar and Nemet, 2002). This
7 finding was reflected in the more general observations in two further studies that scaffolding
8 routines, or structuring discussion through the provision of interim targets, also improved
9 students' understanding (Palincsar *et al.*, 1993; Finkel, 1996).
10
11

12
13
14 Although gains in understanding were reported, the studies also suggested that students often
15 struggled to formulate and express coherent views during small-group discussions, and
16 demonstrated a low level of engagement with tasks. It is therefore not surprising that a
17 number of the studies made recommendations relating to the need for students and/or teachers
18 to be given explicit teaching in the skills associated with the development of arguments and
19 the characteristics associated with effective group discussions.
20
21
22

23
24 Three studies (Richmond and Striley, 1996; Hogan, 1999a; Zohar and Nemet, 2002)
25 recommended training for both students and teachers in the skills needed for handling and
26 participating in group discussions. Richmond and Striley indicated that productive learning
27 was unlikely to take place on a large scale through the use of small-group discussions until
28 students acquire the skills associated with inclusive leadership and are thus able to foster a
29 climate of equitable participation. Hogan (1999a) argued that guiding students towards taking
30 constructive roles in discussions could be achieved through metacognitive training, i.e.
31 knowledge about the nature of collaborative learning, effective group learning strategies, and
32 awareness of what constitutes progress. Two studies (Jiménez-Aleixandre *et al.*, 2000 and
33 Roth and Roychoudhury, 1992) recommended coaching in argumentation skills for both
34 teachers and students.
35
36
37
38
39
40
41

42 Jiménez-Aleixandre *et al.* (2000) suggested that effective discussions are only likely to take
43 place when linked to specific, inquiry-focused tasks where help is given to students to
44 develop their understanding through the construction of arguments. Similarly, Roth and
45 Roychoudhury (1992) reported that students frequently struggled with language, often making
46 short utterances, and appeared to find it difficult to clarify their understanding through
47 explanations, justifications and elaborations. This led them to conclude that a major outcome
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 of their study was the recognition that understanding was only likely to be improved is
3 students were given help in constructing arguments.
4

5
6 One study (Zohar and Nemet, 2002) did involve incorporating explicit instruction about
7
8 argumentation into their intervention. One introductory lesson involved arguments being
9 defined and their structure explained, together with providing examples of characteristics of
10 good arguments. Students then practised the principles through several concrete examples.
11 Zohar and Nemet concluded that argumentation skills were enhanced by explicit instruction
12 and several opportunities for students to take part in discussions to help develop their skills.
13
14
15

Deleted: g

16
17 The nature of the evidence presented in a substantial number of the studies points to the
18 importance of training for teachers and students. This is all the more important as many of
19 the researchers involved in the studies were committed to the use of small-group discussions
20 and had developed some proficiency in designing discussion tasks. This review finding also
21 resonates with the recommendations of Levinson and Turner (2001) and Osborne *et al.* (2002)
22 in their evaluations of two programmes focusing on teaching of socio-scientific issues, and
23 noted earlier in this paper.
24
25
26
27

28 *Methodological considerations*

29

30
31 Although the primary focus of the reviews was to gather substantive data on the use of small-
32 group discussions, some of the methodological aspects have a bearing on the nature of
33 evidence yielded.
34
35

36
37 Positive features of the data collection included the use of multiple data sources with all
38 studies drawing on at least two different kinds of data to increase trustworthiness. Whilst
39 virtually all studies used video recording and/or audio recording to make verbatim records of
40 discussions, these were supported by direct observation to record field notes, interviews,
41 products of student tasks, such as concept maps, student questionnaires and measures of
42 student knowledge were obtained. Although the methods used were rarely justified, the
43 picture gained was one of studies collecting extensive data in an attempt to get as detailed a
44 picture as possible of students' dialogue.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 There were a number of limitations to the data collection. All the studies used a convenience
3 sample for the identification of schools, in many cases using schools where access had been
4 secured through previous involvement of the researcher. With one exception (Zohar and
5 Nemet, 2002), the studies were based in one school and often within one class. A
6 characteristic of much of the work was the use of retrospective sampling, i.e. data were
7 gathered on a number of groups, but reports presented on only a sample of the groups within
8 this, depending on characteristics of the discussion which emerged in the analysis. Such
9 sampling methods are probably realistic for research studies fitting in with practice, and
10 requiring extensive periods of data collection and thus a high degree of co-operation with the
11 class teachers involved. However, retrospective sampling does confer the option on the
12 researcher of exercising a high degree of selectivity in relation to the data presented.
13
14
15
16
17
18

19
20 Three sizeable studies (Gayford, 1995; Lavoie, 1999; Zohar and Nemet, 2002) utilised an
21 experimental design, making comparisons between a control group who experienced a more
22 conventional teaching approach with a group which received some form of intervention
23 related to small-group discussion work. One study (Tolmie and Howe, 1993) specifically set
24 up groups where gender was a variable to be explored. However, the emphasis of the
25 majority of the studies was on describing and interpreting the nature of student discussions
26 and their effects on students' understanding, sometimes making detailed comparisons
27 between groups participating in their studies. Two factors may contribute to the absence of a
28 control group in the studies. Firstly, those undertaking the research might see no need to
29 design their studies to include a control group in what were largely qualitative and
30 interpretative studies. Secondly, the practicalities associated with collecting and analysing
31 extensive in-depth data from a much larger sample in order to make such comparisons would
32 place prohibitive resource constraints on the studies.
33
34
35
36
37
38
39

40 Data analysis was characterised positively by the presentation and discussion of rich and
41 detailed data in the form of extracts from students' discourse, with all studies adopting
42 procedures to increase the trustworthiness of the analysis by having two or more people
43 involved. However, given that the studies were largely gathering qualitative data, there was a
44 surprising lack of contextual detail. Data also tended to be presented in a rather convergent
45 manner, with few examples of data being presented which might disprove assertions or report
46 on unintended outcomes.
47
48
49
50
51
52
53

1
2 Of particular interest were the two very contrasting overall analysis strategies apparent in the
3 studies. The first strategy, adopted by the majority of the studies, was to develop grounded
4 theory from the data through the development of categories then used to characterise the
5 interactions between participants. Hogan (1999a) referred to her analysis as 'ethnographic
6 interaction analysis', whilst Roth and Roychoudhury (1992) used what they described as the
7 techniques used by anthropologists when analysing interactive behaviours. The second
8 strategy, whose use was more limited, was to draw on existing work on discourse analysis or
9 discourse analysis classifications. Such an approach involves trying to identify themes in
10 what people say by looking at sentences, groups of sentences or sentence fragments. These
11 might, for example, relate to attempts to cite others to support a view, or use of evidence to
12 support an account of an event. Discourse analysis techniques were employed by Jiménez-
13 Aleixandre *et al.* (2000), who drew on the work of Bloome *et al.* (1989) to do the initial
14 coding of exchanges between students, and then used Toulmin's (1958) work on argument to
15 classify the interactions where students were talking about science aspects in the discussion.
16 Keys (1997) drew on elements of a framework developed by Kuhn (1993) to code elements'
17 verbal interactions relating to scientific reasoning. However, there was a notable absence of
18 justification in the studies for the approach adopted for analysis, with the development of
19 grounded theory appearing to be seen as an unproblematic choice in the majority of cases. It
20 may be the case that the choice of approach reflects the personal views of the researchers on
21 the role and purposes of data in research. However, the lack of reference made to discourse
22 analysis techniques suggests that these approaches may be unfamiliar to some researchers
23 working in the area of small-group discussions, which, in turn, may be limiting the nature of
24 the analysis. There would appear to be a good case for those researching the effects of small-
25 group discussions to gain a greater familiarity with discourse analysis techniques.

Comment [JMB1]: My understanding of the term 'grounded theory' can be used generically to apply to a situation where researchers develop their own idea from their data. Thus it does describe the process which took place in a number of the studies included in the review, even if the researchers themselves did not employ the term.

Deleted: For example,

Comment [JMB2]: We have taken the description of the work used by the researchers in their paper.

Deleted: Both approaches have their merits.

Formatted: Font: Not Italic

26 Table 4 summarises the key findings to emerge from studies on the use and effects of small-
27 group discussions in science.

28 [Table 4 about here.]

29 Conclusions and recommendations

30 The reviews reported here have yielded insights on both the substantive focus (small-group
31 discussions and their effects) and on the methods employed to gather the data. The review

1
2 has revealed a number of features of particular interest in relation to the use of small-group
3 discussion work in science. It is clear from the review that a complex and interacting set of
4 factors are involved in enabling students to engage in dialogue in a way that could help them
5 draw on evidence to develop and articulate their understanding of science ideas.
6
7

8
9
10 Two particularly strong features which have emerged from the work undertaken for the
11 review is that there is a relatively little good quality systematic research on the effects of
12 small-group discussion work, and considerable uncertainty on the part of teachers as to what
13 they are required to do to implement good practice. Current policy is strongly advocating the
14 use of small-group discussion work in science, and the reviews do indicate that there could be
15 benefits arising from this, as small-group discussion work can provide an appropriate vehicle
16 for assisting in the development of students' understanding of science ideas. Thus teachers
17 should be encouraged to incorporate such discussions into their teaching. However, it is also
18 clear that small-group discussion work needs to be supported by the provision of support and
19 guidance for teachers and students on the development of the skills necessary to make such
20 work effective.
21
22
23
24
25
26

27 One feature, notable by its absence, was the dearth of systematic evidence on the effects of
28 the use of small-group discussions on students' attitudes to their science lessons or science
29 more widely. The absence of such data was very surprising, as the motivational effects of
30 small-group discussions are often cited as a reason for their inclusion in science teaching.
31
32
33

34 It is clear from this review that there is considerable variation in the nature of research into
35 small-group discussion work, particularly in relation to its focus, the clarity with which any
36 variables being investigated are specified, the use of opportunistic samples for data collection,
37 and the techniques used to analyse data. Particularly striking are the two very contrasting
38 approaches to data analysis, with some studies developing grounded theory from the data, and
39 others drawing on existing models to structure their analysis. A substantial proportion of the
40 work also focuses on descriptive data. This can be very helpful in the early stages of a new
41 research area. However, with increasing interest in the *effects* of small group discussions – on
42 student learning, understanding, and attitudes – there is a need to consider what strategies and
43 techniques lend themselves best to the gathering and analysis of data that would help explore
44 such effects.
45
46
47
48
49
50
51
52

1
2 Taken together, these findings suggest there are four key features that need to characterise any
3 further developments in the use of small-group discussions in science teaching and research
4 into their effects. Firstly, some form of professional development training for teachers is
5 highly desirable to provide guidance on how to maximise the effectiveness of small-group
6 discussions. Secondly, further research into the effects of small-group discussions should
7 include a consideration of the extent to which analysis of the data might benefit from
8 established discourse analysis techniques developed in other subject areas (e.g. Barnes and
9 Todd, 1997; Mercer and Littleton, 2007), to establish what they might have to offer work in
10 science. Thirdly, the area would benefit from a more detailed exploration of the effects of
11 small-group discussions on attitudinal effects. Finally, in relation to providing evidence of
12 the *effects* of small-group discussions, there would appear to be potential benefits associated
13 with adopting a mixed method approach to data collection, marrying in-depth qualitative data
14 on the nature of discussions with more quantitative data on student attributes.

Deleted: would benefit from a consideration of

Acknowledgements

The work reported in this paper was funded by the Department for Education and Skills (DfES) as part of the *Evidence, Policy and Practice Initiative* (EPPI).

Electronic searches conducted for the review

Electronic searches of the following databases were conducted:

British Education Index (BEI)

Educational Resources Information Center (ERIC)

PsycINFO

Social Science Citation Index (SSCI)

References in text of report

Barnes, D. & Todd, F. (1997) Communication and learning in small groups. London:

Routledge and Kegan Paul.

Bennett, J., Lubben, F., Hogarth, S. & Campbell, B. (2005) Systematic reviews of research

in science education: rigour or rigidity? *International Journal of Science Education*, 27

(4), 387-406.

- 1
2 Bentley, D & Watts, M. (1992) *Communicating in school science*. London: Falmer Press.
- 3
4 Black, P. & Wiliam, D. (1998) *Inside the black box: raising standards through classroom*
5 *assessment*. London: King's College.
- 6
7 Davies, P. (2000) The relevance of systematic reviews to educational policy and practice.
8 *Oxford Review of Education*, 26, 365-378.
- 9
10 Daws, N. & Singh, B. (1996) Formative assessment: to what extent is its potential to enhance
11 pupils' science being realised? *School Science Review*, 77 (281), 93-100.
- 12
13 Department for Education and Employment (DfEE) (1998) *The National Literacy Strategy*.
14 London: DfEE.
- 15
16 Driver, R. & Bell, B. (1985) Students' thinking and the learning of science: a constructivist
17 view. *School Science Review*, 67 (240), 443-456.
- 18
19 Driver, R., Asoko, H., Leach, J., Mortimer, E. & Scott, P. (1994) Constructing scientific
20 knowledge in the classroom. *Educational Researcher*, 23 (7), 5-12.
- 21
22 [EPPI-Centre \(2002\) EPPI-Centre Review Group Manual. London: Social Science Research](#)
23 [Unit, Institute of Education.](#)
- 24
25 Hargreaves, D. (1996) *Teaching as a research-based profession: possibilities and prospects*.
26 London: The Teacher Training Agency (TTA).
- 27
28 Hillage, L., Pearson, R., Anderson, A. & Tamkin, P. (1998) *Excellence in research on*
29 *schools*. [Brighton: Institute](#) for Employment Studies.
- 30
31 Kuhn, D. (1993) Science as argument: implications for teaching and learning scientific
32 thinking. *Science Education*, 77 (2), 319-337.
- 33
34 Kyriacou, C. (1998) *Essential teaching Skills* (2nd edition) Cheltenham: Stanley Thornes.
- 35
36 Lemke, J. (1990) *Talking science*. New York: Ablex Publishing Corporation.
- 37
38 Levinson, R. & Turner, S. (2001) *Valuable Lessons: Engaging with the Social Context of*
39 *Science in Schools*. London: The Wellcome Trust.
- 40
41 [Mercer, N. & Littleton, K. \(2007\) Dialogue and the Development of children's thinking: a](#)
42 [sociocultural approach. Oxford: Routledge.](#)
- 43
44 [Millar, R. \(2006\). Twenty First Century Science: Insights from the design and](#)
45 [implementation of a scientific literacy approach in school science. International Journal](#)
46 [of Science Education, 28 \(13\), 1499-1521.](#)
- 47
48 Millar, R. & Osborne, J. (eds) (1998) *Beyond 2000: Science Education for the Future*.
49 London: King's College/The Nuffield Foundation.
- 50
51 [Millar, R., Leach, J., Osborne, J. & Ratcliffe, M. \(2006\) Improving subject teaching: lessons](#)
52 [from research in science education. London: Routledge.](#)

Deleted: Brighton:Institute

- 1
2 Newton, P., Driver, R. & Osborne, J. (1999) The place of argumentation in the pedagogy of
3 school science. *International Journal of Science Education*, 21 (5), 553-576.
4
5 Oakley, A. (2002) Social science and evidence-based everything: the case of education.
6
7 *Educational Review*, 54(3), 21-33.
8
9 OECD (2002) *Educational research and development in England: examiners' report*.
10
11 Organisation for Economic Co-operation and Development. www.oecd.org/dataoecd/17/56
12
13 (accessed 2 January 2008)
14
15 Osborne, J. & Collins, S. (2001) Pupils' views of the role and value of the science curriculum.
16 *International Journal of Science Education*, 23 (5), 441-467.
17
18 Osborne, J., Duschl, R. & Fairbrother, R. (2002) *Breaking the Mould? Teaching Science for*
19 *Public Understanding*. London: The Nuffield Foundation.
20
21 Scott, P. (1998) Teacher talk and meaning-making in science classrooms: a Vygotskyan
22 analysis and review. *Studies in Science Education*, 32 45-80.
23
24 Shavelson, R. & Towne, L. (2001) (eds) *Scientific enquiry in education*. Washington DC:
25 National Academy Press.
26
27 Sinclair, J. and Coulthard, R. (1975) *Towards an Analysis of Discourse*, OUP.
28
29 Spencer, L., Ritchie, J., Lewis, J. & Dillon, L. (2003). *Quality in Qualitative Evaluation: a*
30 *framework for assessing research evidence*. London: The Strategy Unit.
31
32 Sutton, C. (1992) *Words, science and learning*. Buckingham: Open University Press.
33
34 Toulmin, S. (1958). *The use of argument*. Cambridge: Cambridge University Press.
35
36 Torgerson, C. (2003) *Systematic Reviews*. London: Continuum.
37
38 Torgerson, C. & Torgerson, D. (2001) The need for randomised controlled trials in
39 educational research. *British Journal of Educational Studies*, 49 (3), 316-328.
40
41 Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*.
42 Buckingham: Open University Press.

References to the full systematic review technical reports

- 43
44
45
46 Author 1, Author 3, Author 2, Author 4, & Author 5. (2004) A systematic review of the
47 nature of small-group discussions in science teaching aimed at improving students'
48 understanding of evidence. In: *Research Evidence in Education Library*. London: EPPI-
49
50
51
52
53
54
55
56
57
58
59
60

Centre, Social Science Research Unit, Institute of Education.

<http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=321>

Author 1, Author 3, Author 2 & Author 4. (2004) A systematic review of the use of small-group discussions in science teaching with students aged 11-18, and their effects on students' understanding in science or attitude to science. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education. <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=326>

References for the studies included in the review

De Vries E, Lund K, Baker M (2002) Computer-mediated epistemic dialogue: explanation and argumentation as vehicles for understanding scientific notions. *Journal of the Learning Sciences* 11: 63-103.

Finkel EA (1996) Making sense of genetics: students' knowledge use during problem solving in a high school genetics class. *Journal of Research in Science Teaching* 33: 345-368.

Gayford C (1995) Science education and sustainability: a case-study in discussion-based learning. *Research in Science & Technological Education* 13: 135-45.

Hogan K (1999a) Sociocognitive roles in science group discourse. *International Journal of Science Education* 21: 855-882.

Hogan K (1999b) Thinking aloud together: a test of an intervention to foster students' collaborative scientific reasoning. *Journal of Research in Science Teaching* 36: 1085-1109.

Jiménez-Aleixandre MP, Pereiro-Muñoz C (2002) Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education* 24: 1171-1190.

Jiménez-Aleixandre MP, Rodríguez AB, Duschl RA (2000a) 'Doing the lesson' or 'doing science': argument in high school genetics. *Science and Education* 84: 757-92.

Johnson SK, Stewart J (2002) Revising and assessing explanatory models in a high school genetics class: a comparison of unsuccessful and successful performance. *Science and Education* 86: 463-480.

Keys CW (1997) An investigation of the relationship between scientific reasoning, conceptual knowledge and model formulation in a naturalistic setting. *International Journal of Science Education* 19: 957-970.

- 1
2 Kurth LA, Anderson CW, Palincsar AS (2002) The case of Carla: dilemmas of helping all
3 students to understand science. *Science Education* 86: 287-313.
4
5 Lajoie SP, Lavigne NC, Guerrero C, Munsie SD (2001) Constructing knowledge in the
6 context of BioWorld. *Instructional Science* 29: 155-186.
7
8 Lavoie DR (1999) Effects of emphasizing hypothetico-predictive reasoning within the science
9 learning cycle on high school student's process skills and conceptual understandings in
10 biology. *Journal of Research in Science Teaching* 36: 1127-1147.
11
12 Meyer K, Woodruff E (1997) Consensually driven explanation in science teaching. *Science*
13 *Education* 81: 173-192.
14
15 Palincsar AS, Anderson C, David YM (1993) Pursuing scientific literacy in the middle grades
16 through collaborative problem solving. *Elementary School Journal* 93: 643-658.
17
18 Richmond G, Striley J (1996) Making meaning in classrooms: social processes in small-group
19 discourse and scientific knowledge building. *Journal of Research in Science Teaching* 33:
20 839-858.
21
22 Roth W-M, Roychoudhury A (1992) The social construction of scientific concepts or the
23 concept map as conscription device and tool for social thinking in high school science.
24 *Science and Education* 76: 531-557.
25
26 Sherman GP, Klein JD (1995) The effects of cued interaction and ability grouping during
27 cooperative computer-based science instruction. *Educational Technology Research and*
28 *Development* 43: 5-24.
29
30 Suthers D, Weiner A (1995) Groupware for developing critical discussion skills. In: Schnase
31 JL, Cunniss EL *Proceedings of CSCL '95: The First International Conference on*
32 *Computer Support for Collaborative Learning*. New Jersey: Lawrence Erlbaum
33 Associates Inc, pages 341-348.
34
35 Tao PK (2001) Developing understanding through confronting varying views: the case of
36 solving qualitative physics problems. *International Journal of Science Education* 23:
37 1201-1218.
38
39 Tao PK (2003) Eliciting and developing junior secondary students' understanding of the
40 nature of science through a peer collaboration instruction in science stories. *International*
41 *Journal of Science Education* 25: 147-171.
42
43 Tolmie A, Howe C (1993) Gender and dialogue in secondary school physics. *Gender and*
44 *Education* 5: 191-209.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2 Tsai C-C (1999) 'Laboratory exercises help me memorize the scientific truths': a study of
3 eighth graders' scientific epistemological views and learning in laboratory activities.
4 *Science and Education* 83: 654-674.
5
6 Williams A (1995) Long-distance collaboration: a case study of science teaching and
7 learning. In: Spiegel SA *Perspectives from Teachers' Classrooms. Action Research.*
8 *Science FEAT (Science for Early Adolescence Teachers).*
9
10 Woodruff E, Meyer K (1997) Explanations from intra- and inter-group discourse: students
11 building knowledge in the science classroom. *Research in Science Education* 27: 25-39.
12
13 Zohar A, Nemet F (2002) Fostering students' knowledge and argumentation skills through
14 dilemmas in human genetics. *Journal of Research in Science Teaching* 39: 35-62.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1

An overview of the chief characteristics of research on small-group discussions

Substantive features

- 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, though they are most widely used with the 11-16 age range.
- The majority of work focused on small-group discussions in relation to students' understanding.
- 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, with groups based on friendship ties, and the discussions lasting 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.
- There were very few instances of students being asked to generate written products from their discussions.

Methodological features

- The most common research strategy used to gather data on small-group discussions was that of case study.
 - The most popular techniques for gathering data were video- and audio-tapes of discussions, supported by observation, interviews, questionnaires and test results.
-

Table 2

Overview of studies included in the review, together with their quality ratings

Note 1: Quality ratings are H = high, MH = medium high, M = medium, ML = medium low, L = low.

Note 2: Some studies have different ratings for different reviews because quality judgements about weight of evidence are made in relation to the focus of each of the reviews.

Study	Country	Review 1	Review 2
1 De Vries <i>et al.</i> , 2002	France	M	M
2 Finkel, 1996	USA	M	MH
3 Gayford, 1995	UK	-	MH
4 Hogan, 1999a	USA	MH	-
5 Hogan, 1999b	USA	M	M
6 Jiménez-Aleixandre <i>et al.</i> , 2000	Spain	MH	-
7 Jiménez-Aleixandre and Pereiro-Muñoz, 2002	Spain	M	-
8 Johnson and Stewart, 2002	USA	ML	-
9 Keys, 1997	USA	MH	MH
10 Kurth <i>et al.</i> , 2002	USA	M	-
11 Lajoie <i>et al.</i> , 2001	Canada	M	M
12 Lavoie, 1999	USA	-	M
13 Meyer and Woodruff, 1997	Canada	ML	-
14 Palincsar <i>et al.</i> , 1993	USA	ML	M
15 Richmond and Striley, 1996	USA	MH	-
16 Roth and Roychoudhury, 1992	Canada	MH	-
17 Sherman and Klein, 1995a	USA	-	H
18 Suthers and Weiner, 1995	USA	-	ML
19 Tao, 2001	Hong Kong	ML	M
20 Tolmie and Howe, 1993	UK	MH	MH
21 Tsai, 1999	Taiwan	ML	-
22 Williams, 1995	USA	-	L
23 Woodruff and Meyer, 1997	Canada	M	-
24 Zohar and Nemet, 2002	Israel	MH	M
<i>Total</i>		<i>19</i>	<i>14</i>

Table 3: Details of the nature of the studies included in the review

Note 1: Quality ratings are H = high, MH = medium high, M = medium, ML = medium low, L = low.

Study	Review	Quality rating	Sample details	Focus of study/nature of intervention	Data gathered
1 De Vries <i>et al.</i> , 2002	1	M	• one class	Physics: sound	• student self-report diaries
	2	M	• 14 students • age 16-17 • groups of two • asynchronous discussion via computer	Student discussions (via computer) logged against 13 categories associated with explanation, argumentation, problem-solving and management	• log of computer dialogue
2 Finkel, 1996	1	M	• one class	Biology: genetics	• audiotapes of group discussions
	2	MH	• 25 students • age 16-18 • groups of 3-4	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 plenary class presentations and discussions • computer logs • individual diaries • student work
3 Gayford, 1995	2	MH	• two classes (control and experimental) from four schools • age 16 • groups of 3-4	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	MH	• 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	M	• 2 schools	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
	2	M	• 8 classes (four control and four experimental) • 163 students • age 11-16 • groups of 3-4 • heterogeneous for gender and ability		• audio and video tapes of discussions • field notes of class observations • tests of conceptual understanding • psychological profiles
6 Jiménez-Aleixandre <i>et al.</i> , 2000	1	MH	• 1 class • 24 students • age 14-15 • groups of 4	Biology: 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	• observation • audiotapes of group discussions
7 Jiménez-Aleixandre	1	M	• 1 school	Environmental science	• observation

1	and Pereiro-Muñoz, 2002			<ul style="list-style-type: none"> • 38 students • age 15-16 (plus some mature students) • groups of 4-6 	Exploring students' ability to construct arguments and engage in decision-making about environmental processes	<ul style="list-style-type: none"> • audio and videotapes of group discussions • field notes • student-generated material 	
2							
3							
4							
5	8	Keys, 1997	1 2	MH MH	<ul style="list-style-type: none"> • 1 school • 6 students • age 14-15 • groups of 3 	Chemistry: elements and bonding Discussions focused on the development of reasoning strategies and discourse through a collaborative writing tasks	<ul style="list-style-type: none"> • videotapes of discussions • interviews with students • tests of conceptual understanding • student work
6							
7							
8							
9							
10	9	Kurth <i>et al.</i> , 2002	1	M	<ul style="list-style-type: none"> • one school • 4 students • age 11-12 • group of 4 	Physics: density Material modified from the normal school module to incorporate discussion tasks	<ul style="list-style-type: none"> • tests of conceptual understanding • observation • self-completion questionnaire • self-completion diary
11							
12							
13							
14	10	Lajoie <i>et al.</i> , 2001	1 2	M M	<ul style="list-style-type: none"> • 1 school • 2 classes • 40 students • age 14-15 • groups of 2 	Biology: digestion Discussion facilitated through the use of a computer-learning environment, <i>Bioworld</i> .	<ul style="list-style-type: none"> • audio and video tapes of discussions • computer log of actions and decisions
15							
16							
17							
18							
19							
20	11	Lavoie, 1999	2	M	<ul style="list-style-type: none"> • 1 school • 10 classes (5 control and 5 experimental) • 250 students • age 15-16 • groups of 3-4 	Biology: several topics Topics taught in a standard way, and though a learning cycle model (exploration, term introduction, concept application)	<ul style="list-style-type: none"> • 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
21							
22							
23							
24							
25							
26							
27							
28							
29	12	Palincsar <i>et al.</i> , 1993	1 2	ML M	<ul style="list-style-type: none"> • 2 schools • 9 classes • 230 students • age 11-12 • groups of 4 	Chemistry: kinetic theory Discussion tasks aimed at modelling the working of scientific communities	<ul style="list-style-type: none"> • tests of conceptual understanding • interviews • video-recordings of selected groups • student logs
30							
31							
32							
33							
34							
35	13	Richmond and Striley, 1996	1	MH	<ul style="list-style-type: none"> • 1 school • 1 class • 24 students • age 15-16 • groups of 4 • mixed ability and gender 	Science: cholera epidemics and cystic fibrosis Aimed to explore difficulties students encounter when developing scientific arguments and how student interactions shaped arguments	<ul style="list-style-type: none"> • observation • self-completion student diaries • school/college records
36							
37							
38							
39							
40							
41							
42	14	Roth and Roychoudhury, 1992	1	MH	<ul style="list-style-type: none"> • 1 school • 6 classes 	Physics: light Aimed to explore the development of student understanding	<ul style="list-style-type: none"> • video-recording of discussions • self-completion questionnaire
43							
44							
45							
46							
47							

1				• 148 students (but only one group studied in detail)	through engaging in the process of developing concept maps	• concept maps generated in discussions	
2				• age 15-17			
3				• groups of 3-4			
4							
5							
6	15	Sherman and Klein, 1995	2	H	• 1 school • number of classes unspecified • 231 students • age 13-14 • groups of 2	Science: investigations Computer programme about designing controlled experiments	• observation • self-completion questionnaire • pre- and post-tests of understanding
7							
8							
9							
10							
11	16	Tao, 2001	1 2	ML	• 1 school • 1 class • 16 students • age 17-18 • groups of 2	Physics: several topics Multiple solutions presented to students to see if discussions improved their understanding	• pre- and post-tests of understanding • individual interviews • audio tapes of discussions
12							
13							
14							
15							
16							
17	17	Tolmie and Howe, 1993	1 2	MH MH	• 1 school • number of classes unspecified • 82 students • age 12 to 15 • groups of 2	Physics: forces and motion Exploration of gender differences in group discussions	• Assessment of understanding • Video-tapes of discussions • Psychological test • Computer record of joint predictions
18							
19							
20							
21							
22							
23	18	Woodruff and Meyer, 1997	1	M	• 1 school • 3 classes • sample size not stated – probably one class • age 11-13 • groups of 3-4	Physics: shadows, and floating and sinking. Findings reported from three studies on interactions in discussions within and between groups	• audio-recordings of discussions • field notes
24							
25							
26							
27							
28							
29	19	Zohar and Nemet, 2002	1 2	MH M	• 2 school • 9 classes (5 control, 2 experimental) • 186 students • age 13-14 • groups of 5-7	Biology: genetics Exploration of the effects of a unit teaching argumentation skills	• multiple-choice test of understanding • pre- and post-test of argumentation skills • student worksheets • audio-tapes of four discussions
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							
45							
46							
47							

Table 4

Summary of key findings on the use and effects of small-group discussions

Substantive features

- 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.
- Groups function more purposefully when specifically constituted such that differing views are represented, and improvements in understanding are greatest where there is initial *dissimilarity* in understanding of the science ideas associated with the discussion task.
- Groups function more purposefully, and students' understanding improves, 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, and students' understanding improves, when some form of training is provided for students on aspects of small-group discussion work, and when help in structuring discussions is provided in the form of 'cues'.
- Single sex groups function more purposefully than mixed sex groups, though improvements in understanding are independent of gender composition of groups.
- Group leaders able to adopt an inclusive style, and one which promoted reflection, are the most successful in achieving engagement with the task.
- Incorrect or inadequate prior knowledge hinders development of students' understanding through small-group discussion.
- 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. The effectiveness of small-group discussions, and their effects on students' understanding of evidence, is linked more strongly to the provision of such guidance, rather than to any particular type of stimulus material.
- Little systematic data has been gathered on the effects of small-group discussions on students' affective responses to science.

Methodological features

- All studies generated large data sets, and used multiple data sources to enhance the reliability and/or trustworthiness of data.
 - With one exception all the studies were based on single schools, many on single classes, with schools being identified through convenience sampling.
 - Relatively few studies made use of experimental designs.
 - Two very contrasting approaches to analysis were adopted: one developing theory from the data ('grounded theory') and the other using established techniques for discourse analysis.
-