ENHANCING THE QUALITY OF ARGUMENTATION IN SCHOOL SCIENCE

Jonathan Osborne*, Sibel Erduran*& Shirley Simon**

** King's College London*
***Institute of Education, London*

Abstract

The research reported in this paper focussed on the design of learning environments that support the teaching and learning of argumentation in a scientific context. The research took place over two years between 1999 and 2001 in junior high schools in the greater London area. The research was conducted in two phases. In the first developmental phase, working with a group of 12 science teachers, the main emphasis was to develop sets of materials and strategies to support argumentation in the classroom and to assess teachers' development with teaching argumentation. Data were collected by videoing and audio recording the teachers attempts to implement these lessons at the beginning and end of the year. During this phase, analytical tools for evaluating the quality of argumentation were developed based on Toulmin's argument pattern. Analysis of the data shows that there was significant development in the majority of teachers use of argumentation across the year. Results indicate that the pattern of use of argumentation is teacher specific, as is the nature of the change.

In the second phase of the project, teachers taught the experimental groups a minimum of nine lessons which involved socioscientific or scientific argumentation. In addition, these teachers taught similar lessons to a control group at the beginning and end of the year. Here the emphasis lay on assessing the progression in student capabilities with argumentation. Hence data were collected from several lessons of two groups of students engaging in argumentation. Using a framework for evaluating the nature of the discourse and its quality, the findings show that there was an improvement in the quality of students' argumentation. In addition, the research offers methodological developments for work in this field.

Introduction & Background

Curriculum innovations in science like those sponsored by the Nuffield Foundation in the UK and the National Science Foundation in the USA in the 60s and 70s, have had little impact on the practices of science teachers (Welch, 1979). Four decades after Joseph Schwab's introduction of the idea that science should be taught as an 'enquiry into enquiry' and almost a century since John Dewey advocated classroom learning be a student-centred process of enquiry, we still find ourselves struggling to achieve such practices in the science classroom. Witness the publication of the AAAS edited volume on inquiry (Minstrell & Van Zee, 2000), the recent release of *Inquiry and the National Science Education Standards* (National Research Council, 2000) and the inclusion of 'scientific enquiry' as a separate strand in the English and Welsh science national curriculum (Department for Education and Employment, 1999). These three works serve as signposts to an ideological commitment that teaching science needs to accomplish much more than simply detailing what we know. Of growing importance is the need to educate our pupils and citizens about how we know and why we believe in the scientific world view e.g., science as a way of knowing (Driver, Leach, Millar, & Scott, 1996; Duschl, 1990; Millar & Osborne, 1998). Such a shift requires a new focus on (1) how evidence is used in science for the construction of explanations – that is on the arguments that form the links between data and the theories that science has constructed; and (2), the development of criteria used in science to evaluate the selection of evidence and the construction of explanations. Central to this perspective is a recognition that language is not an adjunct to science but an essential constitutive element (Norris & Phillips, 2001; Osborne, 2002). In particular, that the construction of argument, and its critical evaluation is a core discursive activity of science.

Whilst the consideration of the important role language, conversation and discussion have in science learning can be traced back 3 or 4 decades (Scheffler, 1960; Bruner, 1961; Lansdown, Blackwood & Brandwein, 1971), it was not until the 1980s that serious discussion of the role of language in science learning began (c.f., (Aikenhead, 1991; Gee, 1996; Lemke, 1990; Sutton, 1992)). Only recently, has the field turned its attention to that discourse which specifically addresses argumentation – a case that was strongly articulated in Newton, Driver and Osborne (1999) and Driver, Newton and Osborne (2000). The general point made in these papers is that argumentation, i.e., the coordination of evidence and theory to support or refute an explanatory conclusion, model or prediction (Suppe, 1998) is a critically important epistemic task and discourse process in science. Likewise, situating argumentation as a central element in the learning of sciences has two functions: one is as a heuristic to engage learners in the coordination of conceptual and epistemic goals and the second is to make student scientific thinking and reasoning visible to enable formative assessment by teachers or instructors. Striving for epistemic goals like developing, evaluating and revising scientific arguments represents, therefore, an essential element of any contemporary science education.

For contemporary science impinges directly upon many aspects of people's lives. Individuals and societies have to make personal and ethical decisions about a range of socio-scientific issues (e.g, genetic engineering, reproductive technologies, food safety) based on information available through the press and other media. Often accounts of new developments in science report equivocal findings or contested claims. Evaluating such reports is not straightforward as it requires the ability to assess whether the evidence is valid and reliable, to distinguish correlations from causes, observations from inferences, and to assess the degree of risk (Millar & Osborne, 1998; Monk & Osborne, 1997). Within the context of a society where scientific issues increasingly dominate the cultural landscape, where social practices are constantly examined and reformed in the light of scientific evidence, and where the public maintain an attitude of ambivalence (Giddens, 1990) or anxiety about science (Beck, 1992), there is an urgent need to improve the quality of young people's understanding of the nature of argument in general and argument in a scientific context, in particular. For developments in scientific knowledge increasingly pose a set of moral and ethical dilemmas that require a careful and considered response. Moreover, within science education itself, it is ironic that a discipline which presents itself as the epitome of rationality, so singularly fails to educate its students about the epistemic basis of belief relying instead on authoritative modes of discourse (Scott, 1998) that leave students with naïve images of science (Driver et al., 1996)

An important task for science education, therefore, is to expose the epistemic core of science – the construction of argument and explanation and develop children's ability to understand and practice scientifically valid ways of arguing, enabling them to recognise not only the strengths of scientific argument, but also its *limitations* (Osborne & Young, 1998). Hence, the research presented in this paper, seeks to study whether young people's quality of 'argument' about scientific issues and their critical capabilities can be enhanced in science lessons.

Previous research on argument

Over the past few decades certain influential educational projects have all laid foundations for the work on argumentation in science lessons. These projects have promoted independent thinking, the importance of discourse in education and the significance of co-operative and collaborative group work (e.g, Rudduck, 1983; Barnes, 1977; Cowie and Rudduck, 1990; Solomon, 1990, Ratcliffe, 1996). In addition to these projects, a body of relatively unintegrated research concerning argumentative discourse in science education has begun to emerge (Alverman, Qian, & Hynd, 1995; Boulter & Gilbert, 1995; Geddis, 1991; Hammer, 1995; Herrenkohl, Palinscar, DeWater, & Kawasaki, 1999; Herrenkohl & Guerra, 1995; Jiménex-Aleixandre, Bugallo-Rodríguez, & Duschl, 1997; Jiménex-Aleixandre, Rodríguez, & Duschl, 2000; Mason, 1996; Means & Voss, 1996; Russell, 1983). Perhaps the most significant contribution to this literature has come from Kuhn (e.g, (Kuhn, 1991)) who explored the basic capacity of individuals to use reasoned argument. Kuhn investigated the responses of children and adults to questions concerning problematic social issues. She concluded that many children and adults (especially the less well educated) are very poor at co-ordinating and constructing a relationship between evidence (data) and theory (claim) that is essential to a valid argument. More recent work by Hogan and Maglienti (2001) exploring the differences between the reasoning ability of scientists, students and non-scientists found, likewise, that the performance of the latter two groups were significantly inferior.

Koslowski (1996), who is critical of Kuhn's emphasis on co-variation, was less doubtful of young people's ability to reason pointing to the fact that theory and data are both crucial to reasoning and interdependent. Hence, lack of knowledge of any relevant theory or concepts often constrains young people's ability to reason effectively. Whilst this is an important point (and one to which we will return latter), what it suggests is that scientific rationality requires a knowledge of scientific theories, a familiarity with their supporting evidence and the opportunity to construct and/or evaluate their inter-relationship. Nevertheless, Kuhn's research is important because it highlights the fact that, for the overwhelming majority, the use of valid argument does not come naturally and is acquired through practice. The implication that we draw from the work of Kuhn and others is that argument is a form of discourse that needs to be appropriated by children and *explicitly taught* through suitable instruction, task structuring and modelling. Just giving students scientific or controversial socio-scientific issues to discuss will not prove sufficient to ensure the practice of valid argument. Similar conclusions were reached by Zohar & Nemet (2002), Kelly et al (2001) and Hogan and Maglienti (2001). The latter argued that 'students need to participate over time in explicit discussions in the norms and criteria that underlie scientific work'. Hence our focus has been upon developing pedagogical practices that support argumentation and foster students' epistemological development. And, whilst general advice concerning how to structure successful discussion and argumentation can be found in the literature (e.g, (Dillon, 1994)) or in other disciplines (Andrews, 1995) – only a little has been situated within the specific context of the science classroom.

A significant problem confronting the development of argumentation in the science classroom is that it is fundamentally a dialogic event carried out among two or more individuals. Scott (1998), in a significant review of the nature of classroom discourse shows how it lies on a continuum from 'authoritative', which is associated with closed questioning and IRE dialogue – to 'dialogic' which is associated with extended student contributions and uncertainty. However, the combination of the power relationship that exists between a science teacher and student; the rhetorical project of the science teacher which seeks to establish the consensually agreed scientific world-view with his or her students; and the authoritarian, dogmatic nature of the discipline means that opportunities for dialogic discourse are minimised. For instance, Newton et al. (1999) found that deliberative discussions commonly occupied only 2% of all science lessons in junior high schools. Hence, introducing argumentation requires a shift in the normative nature of classroom discourse. Change requires, however, that science teachers be convinced that argumentation is an essential component for the learning of science. In addition, they require a range of pedagogical strategies that will both initiate and support argumentation if they are to adopt and integrate argumentation into the classroom.

At the core of such strategies is the requirement to consider not singular explanations of phenomena but *plural* accounts (Monk & Osborne, 1997, Driver, Newton & Osborne, 2000). Students must, at the very least spend time considering not only the scientific theory but also an alternative such as the common lay misconception, i.e. that all objects fall with the same acceleration v the notion that heavier things fall faster. Such contexts can also be social considerations of the application of science such as the use of animals for drug testing, problem-based learning situations, or computer mediated situations such as the material developed by the WISE project (Bell & Linn, 2000).

The evidence that does exist suggests that argumentation is fostered by a context in which student-student interaction is permitted and encouraged (Alexopoulou & Driver, 1997; Alverman et al., 1995; Herrenkohl et al., 1999; Jiménex-Aleixandre et al., 2000; Kuhn, Shaw, & Felton, 1997; Thorley & Treagust, 1986; Zohar & Nemet, 2002). For instance, Kuhn, Shaw, and Felton (Kuhn et al., 1997), in testing the hypothesis that engagement in thinking about a topic enhances the quality of reasoning about the topic, found that dyadic interaction significantly increased the quality of argumentative reasoning in both early adolescence and young adults. Likewise, the work of Eichinger et al. (1991) & Herrenkohl et al. (1999) found that bringing scientific discourse to the classroom required the adoption of instructional designs that permit students to work collaboratively in problem solving groups. Some of the research on discourse points, too, to the importance of establishing procedural guidelines for the students (Herrenkohl, Palincsar, DeWater, & Kawasaki, 1999). The point to make is that *both* epistemological *and* social structures in the classrooms are important factors for designing activities that foster argumentation. One element, therefore, is the need to provide students access to not a singular world-view but to plural accounts of phenomena and the evidence that can be deployed in an argument. Of itself, however, that is not sufficient as a second requirement is a context that fosters dialogic discourse. This we have seen as the use of techniques such as student presentations, small-group discussions coupled with guidelines and assistance that support the appropriation of argumentation skills and discourse.

Once space for argumentation is established, then it is possible that it will lead to cognitive development. For instance, Zohar (2002) found that engaging in argumentation about dilemmas posed by human genetics led, not only to an improvement in students argumentation skills, but also to their knowledge and understanding of the topic. Focussing on the epistemic and social nature of classroom activities is, we believe, an essential precursor to cognitive development for, as Billig (1996) argues, 'humans do not converse because they have inner thoughts to express, but they have thoughts because they are able to converse.' And, hence, learning to think is learning to argue. Consequently, in developing materials and strategies for argumentation we have used these elements as guiding principles which underlie the approach and design of all that we have sought to do.

At this point, it is worth noting that, in this paper, we draw a distinction between 'argument' and 'argumentation'. The former we see as a referent to the claim, data, warrants and backings that form the substance or *content* of an argument. The latter, in contrast, we see as a referent to the *process* of arguing. The focus of this work has been to explore those strategies which scaffold and support 'argumentation' and to develop frameworks for the assessment of its quality.

Research Objectives

We believe that promoting the practice of 'argumentation' in science lessons requires the development of appropriate pedagogical strategies that offer practical guidance for teachers. Furthermore, the benefit of such guidance needs to be assessed through empirical studies. Our research was seeking, therefore to:

(i) identify the pedagogical strategies necessary to promote 'argument' skills in young people in science lessons;
(ii) trial the pedagogical strategies and determine the extent to which their implementation enhances teachers' pedagogic practice with 'argument';
(iii) determine the extent to which lessons which follow these pedagogical strategies lead to enhanced quality in pupils' arguments.

To investigate these objectives, we have chosen to work in two contexts – a socio-scientific context and a scientific context. The former we have seen as important because many of the debates surrounding science in the public domain are of this nature. Moreover, many such issues draw on existing knowledge and resources of which young children already have some knowledge. Scientific arguments are, however, important as they are they expose the justification for belief in the scientific world-view and the underlying rationality that lies at the heart of science. Understanding the role of argument in constructing the link between ideas and their evidence, and students' ability to critically evaluate such arguments is a fundamental epistemic ability that any science education should seek to develop.

In conducting our research we have worked initially with a group of 12 teachers to explore and develop their practice at initiating argumentation in the classroom, and then in the second year with a subset of 6 teachers to explore what effect such activities had on the classroom discourse and student use of argument.

Our analytic perspective upon argument

Assuming, as the research evidence suggests, that a context that fosters and develops students' use of argumentation can be established, then what can teachers learn by listening to student discussion and how can they foster and improve the quality of argumentation? Essentially, how can they respond formatively to assist their students and develop their reasoning? How, for instance, can they identify the essential features of an argument? How are they to judge that one argument is better than another? And how should they model arguments of quality to their students? Before we could ask teachers to engage their students in argumentation and use the information they acquire from the process to plan subsequent lessons or evaluate students learning, it was essential to provide some theoretical guidance to answer such questions. Thus, an important component of this research has been the need to adopt and develop a set of criteria to analyse both the content and the form of children's arguments.

In our work, we have chosen to use the analytic framework developed by Toulmin (1958). His model of argument, referred to here as Toulmin's argument pattern (TAP), was one of the first to challenge the 'truth' seeking role of argument and to consider, instead, the rhetorical elements of argumentation and their function. For Toulmin, the essential elements of argument are claims, data, warrants and backings. Normatively, any argument relies on an evidential base that consists of supporting data whose relationship to the claim is elaborated through the warrant, which in turn, may be dependent on a set of underlying theoretical presumptions or backings. Arguments may be hedged with qualifications to show the limits of their validity and are commonly challenged by querying the data, warrants or backings. In practice, arguments are field dependent, as the warrants and backings used to make claims are shaped by the guiding conceptions and values of the field. An alternative framework for the analysis of argument is that developed by Walton (1996) which characterises argument in terms of a schema of 25 common forms of reasoning. Our view, however, was that this framework gave more emphasis to the content of an argument which was not the essential focus of our work.

Toulmin's model has been used as a basis for characterising argumentation in science lessons (Russell, 1983) and is implicit in a coding system of others (Kuhn et al., 1997; Pontecorvo, 1987). In addition, following Pontecorvo, we have focussed on the epistemic and argumentative operations adopted by students — that is their reasoning functions and strategies. These are the salient cognitive operations, produced by the speaker, which correspond to strategies which are more or less effective for constructing valid argument. Features which we have concentrated on, therefore, in the analysis of argumentation in both scientific and socio-scientific contexts, include: the extent to which students have made use of data, claims, warrants, backings and qualifiers; and the extent to which they have engaged in claiming, elaborating, reinforcing or opposing the arguments of each other.

The Research Programme

General features of the research

A group of teachers interested in collaborating with us was initially established for some preliminary work in the area. From this group, 12 were selected - our principal criteria being the experience and confidence of the teachers, as the work would involve a degree of risk on their part. The schools chosen for this work were located in the Greater London area and ranged from urban to suburban settings with mixed ethnic groups. Three schools were all-girls schools, one school was a private school, and 11 were public schools ranging from inner-city urban to suburban middle-class catchment areas. Thus the work was conducted with a broadly representative sample of pupils of varying academic ability. . Our discussions with teachers led to the choice of students in Grade 8 (age 12-13) as the most suitable because of the freedom from examination constraints.

The research has been conducted in essentially two phases. In the first year (Sept 99 – Sept 2000), we sought to focus on developing the skills of the teacher and the materials for use in argument-based lessons. During the first year, the teachers also attended 6 half day meetings, held at King's College London, to discuss and share pedagogical strategies for teaching such lessons, develop materials and to develop their understanding of our theoretical perspective on argument. A fuller description of the approach taken in these workshops can be found in Simone, Erduran and Osborne (submitted, 2002).

In each phase, the teachers involved in the study incorporated a series of nine argument-based lessons, approximately once a month over the course of one year, involving focussed discussions relevant to the science national curriculum for England during the first year. The first and ninth lessons were devoted to discussion of a socio-scientific issue of whether zoos should be permitted whilst the remaining lessons have been devoted solely to discussion and argument of a *scientific* nature. To support these lessons, teachers were initially provided with a set of materials drawn from a trawl of the literature and our own ideas for use with students. These aimed to develop their knowledge and capabilities with scientific reasoning by examining evidence for/against a theory, e.g. the particle hypothesis, the explanation of day and night. Other activities have focussed on sets of data, their interpretation and the conclusions that can be drawn from them. Resources for teaching all of these lessons were also developed separately by teachers.

To assess the teachers' progress, we have video and audio-recorded the teacher at the beginning of year 1 and year 2 and systematically analysed these transcripts to evaluate the characteristics of their approach to argumentation, to see if there is an identifiable measure of their progress. We have also taped and transcribed two groups in each class to develop a schema for evaluating the quality of their argumentation.

In the second phase of the project (Sept 2000 – Sept 2001), we worked with a subset of 6 teachers asking them to repeat the process. Support in this phase was reduced to three half-day meetings across the year and feedback provided *in situ* whenever a visit was made for the purpose of data collection. In addition, each teacher taught a class of the same year, as similar as possible in aptitude and ability, to provide some basis for comparison and a control. The focus of our analysis in this second stage has been on the recordings and transcripts of the discussions by pupils to see if there was any improvement in the quality or quantity of argument. What follows is a summary of the salient findings that have emerged from the work of the project and an exploration of their implications.

Materials and Support for Argument

One of the features of this work has been to try and develop materials that could be used for supporting argumentation in the classroom. The essential precursor to initiating argument in any context is the generation of difference or plural theoretical interpretations. Hence, a common framework for most of the materials we have developed has taken the form of presenting or generating competing theories for students to examine, discuss and evaluate. A universal requirement has been the opportunity for pupils to meet in small groups and discuss these ideas, the evidence for them and to construct arguments justifying the case for one or other theory. However, initiating argument also requires a resource or data to enable the construction of argument. Hence, commonly, competing theories have been accompanied by evidence which students are asked to use to decide whether the evidence presented supports theory 1, theory 2, both or neither. Using these ideas as the essential principles for initiating argument in the science classroom, drawing on the literature, we developed nine generic frameworks for promoting argument in the classroom. An outline of these is provided in Table 1 and Fig 1 gives one example in more detail.

Table 1: Generic Frameworks for Materials for supporting and facilitating Argumentation in the Science Classroom

Framework	Description
1. Table of Statements	Students are given a table of statements on a particular science topic. They are asked to say if they agree or disagree with the statement and argue for their choices. This idea has been developed from the work on discussing instances of physical phenomena (Gilbert & Watts, 1983)
2. Concept Map of Student Ideas	Students are given a concept map of statements derived from student conceptions of a science topic derived from the research literature. They are then asked to discuss the concepts and links individually and as a group to decide whether they are scientifically correct or false, providing reasons and arguments for their choice. This was an adaptation of the common use of concept mapping (Osborne, 1997)
3. A Report of a Science Experiment Undertaken by Students	Students are given a record of another student's experiment and their conclusions. The experiment is written in a way to intentionally include information that is lacking or in a manner could be improved, so as to stimulate disagreement. Students are to provide answers to what they think the experiment and its conclusions could be improved, and why. This idea was drawn from the work of Goldsworthy, Watson and Wood-Robinson (2000)
4. Competing Theories – Cartoons	Students are presented with two or more competing theories in the form of a cartoon. They are asked to state which they believe in and argue why they think they are correct. The work of Keogh and Naylor (Keogh & Naylor, 1999; Naylor & Keogh, 2000) has been valuable in developing a resource which is an excellent stimulus to engaging children with scientific thinking.
5. Competing Theories – Story	Students are presented competing theories in the form of an engaging story reported in a newspaper. They are then asked to provide evidence for which theory they believe in and why.

Framework	Description
6. Competing Theories – Ideas and Evidence	In this approach, students are introduced to a physical phenomenon and then offered two or more, but generally two, competing explanations. In addition, a range of statements of evidence that may support one theory, the other, both or neither are provided. In small groups, students are then asked to consider each piece of evidence and evaluate its role and significance. Finally, they must use the evidence to argue for one idea or another. This idea has been adapted from the work of Solomon et al.(Solomon, 1991; Solomon, Duveen, & Scott, 1992).
7 Constructing an Argument	Students are given an explanation of a physical phenomenon ie day and night are caused by a spinning Earth, and a number of data statements (typically 4). They then have to discuss which data statements provides the strongest explanation for the phenomenon and provide an argument why. This is an idea that has been adapted from the innovative work of Garratt et al. (1999) in undergraduate chemistry.
8. Predicting, Observing and Explaining	This activity, drawn from the work of White and Gunstone (1992), involves introducing a phenomenon to children without demonstrating it and asking students to discuss in small groups what they think will happen when the phenomenon is initiated, and justify their reasoning. The phenomenon is then demonstrated and if what happens is the antithesis of that expected, students are then asked to reconsider and re-evaluate their initial arguments. Discussion focuses on the theory that they advance for their prediction and the evidence to support it.
9. Designing an Experiment	Students are asked to work in pairs to design an experiment to test a hypothesis i.e. that a silver kettle cools faster. Their design needs to specify not only what variable should be measured but how often and what steps should be taken to ensure that the data obtained are reliable. Pairs then meet to discuss their design, to propose alternative procedures and to argue for their relative merits.

Example 1: Competing Theories A

Theory 1: Light rays travel from our eyes onto the objects and enable us to see them.

Theory 2: Light rays are produced by a source of light and reflect off objects into our eyes so we can see them.

Which of the following pieces of evidence supports Theory 1, Theory 2, both or neither. Discuss.

a. Light travels in straight lines
b. We can still see at night when there is no sun
c. Sunglasses are worn to protect our eyes
d. If there is no light we cannot see a thing
e. We 'stare at' people, 'look daggers' and 'catch people's eye'
f. You have to look at something to see it.

Fig 1: An example of materials developed by one teacher using framework 6.

The reasons for choosing to develop generic frameworks was essentially twofold. One was pragmatic in that the topics being taught by the teachers varied from school to school and lesson to lesson. Demanding that specific exemplar lessons be taught would have placed too restrictive a burden on the teachers of science and made the project unworkable. More fundamentally, providing a framework on which the substance of a lesson could be 'hung' provided teachers with a vital element of independence. This enabled them both to make a contribution in developing and trialling their own ideas, and to take ownership of the work – an element which is vital for successful curriculum innovation (Ogborn, 2002). Whilst a detailed analysis of all the lessons has not been conducted, the majority of the materials developed have made use of framework 6 (Fig 1).

In addition, we sought to develop methods by which student argumentation could be facilitated and scaffolded. For teachers' discursive engagement in the classroom a set of argumentation prompts were developed. Essentially these were open-ended questions designed to elicit a justificatory argument from the student such as why do you think that? Can you think of another argument for your view? Can you think of an argument against your view? How do you know? And what is the evidence for your view?

Another strategy is drawn from the literature on teaching pupils to write (Bereiter & Scardamalia, 1987). Constructing a good argument is not a simple task and students need guidance and support which will help them to 'scaffold' and build their sense of what is an effective argument. Wray and Lewis (1997) have shown that when such genres of writing or expression are not familiar, 'writing frames' that support the process of writing can provide vital support and clues as to what is needed. These are essentially drafting documents for recording notes of their discussion which can then be used as the basis for a verbal presentation to the rest of the class, or alternatively, as a structure for producing a written argument. Several variants of these 'writing frames' were developed for supporting argumentation. Essentially, these contain a set of stems such as my argument is.., my reasons are that.., arguments against my idea might be that.., I would convince somebody that does not believe me by.., the evidence to support my argument is...These stems provide essential prompts necessary to initiate the construction of a written argument and to structure it in a coherent manner.

Finally, an important aspect of developing an understanding of argument and evidence for pupils is the need to present examples of argument and model good practice. This practice requires that pupils are offered examples of both weaker and stronger arguments enabling discussion of the features that make one better than another. Examples of arguments of different quality which could be offered to pupils, were developed and shown to the teachers engaged in the research. Examples of poor quality argumentation were written to illustrate that such arguments relied on assertion with minimal use of data or warrants to justify claims. Examples of stronger argumentation showed how such arguments drew on a wider range of evidence and included rebuttals of counter arguments. Such examples serve an important illustrative function of what constitutes good argumentation. Two examples are shown below for illustration.

Weak Argument

We must see because light enters the eye [claim]. You need light to see by [data]. After all, otherwise we would be able to see in the dark [warrant].

Stronger Argument

Seeing because light enters the eye makes more sense [claim]. We can't see when there is no light at all [data]. If something was coming out of our eyes, we should always be able to see even in the pitch dark [rebuttal]. Sunglasses stop something coming in, not something going out [data]. The only reason you have to look towards something to see it is because you need to catch the light coming from that direction [rebuttal]. The eye is rather like a camera with a light sensitive coating at the back which picks up light coming in, not something going out [warrant].

Teachers were also asked to explain to their students the importance of thinking of counter-arguments to an argument or rebuttals which challenge the justification of another's argument.

Data Sources

The data sources were verbal conversations of teachers and students audio-taped in classes of year 8 (age 12-13) students. In year 1, we videoed two lessons – one at the beginning of the year and one a year later. At this stage of our work the focus was on argumentation in socio-scientific context. Consequently, the main task within these lessons was an exploration of arguments for, and against, the funding of a new zoo. Each lesson had 3 sections. At the onset, the teacher distributed a letter outlining the task and there was a whole class discussion on the pros and cons of zoos. Then the students were placed into small groups of 3 or 4 and asked to come to discuss whether or not the zoo should be built. Finally, in the last phase of the lesson, the groups made presentations and shared their opinions with the rest of the class. For homework, students were typically asked to write a letter or compose a poster that would communicate their arguments. Needless to say, there was considerable variation between teachers in the detail of their implementation. Microphones were attached to the teachers so as to capture their verbal contribution to the lesson as well as their interactions with students during the group format. In addition, two groups of four pupils were selected and their conversations recorded.

In the second year of our work, a subset of six teachers was selected on the basis that they were individuals who were considered to have made more progress in their ability to facilitate and incorporate argumentation in their pedagogical practice (a judgement born out retrospectively by the data analysis – see Table 2). As well as recording the teachers' second attempt at teaching the zoo lesson to use as a comparison with their first attempt a year previously, this phase sought to examine pupils' ability to incorporate and use argumentation in two contexts:

1. a socio-scientific topic to compare the development of the experimental group with a control group using data drawn from the zoo lesson at the beginning of the year, and a lesson about the possible siting of a leisure centre in a nature reserve at the end of the year; and
2. a scientific context to compare the development of the experimental group using data drawn from a lesson at the beginning of the year and the end of the year, and to compare their development at arguing in a scientific context with arguing in a socio-scientific context.

Thus, in addition to the data collected from the lessons exploring arguments for and against the establishment of a new zoo at the beginning of year 2 (6 teacher tapes, 11 student videos[1]), data were also collected from the same teachers teaching the same lesson to a control group (5 teacher tapes, 9 student videos); and from the same teachers implementing argument in a scientific context (6 teacher tapes, 12 student videos). In each of the lessons, wherever possible, a tape was collected for each teacher and two selected groups of four pupils. In the intervening period, teachers taught a minimum of 8 lessons using argument *in a scientific context*.

At the end of the second year, another set of data was collected from the same group of 6 teachers teaching argumentation to the intervention class in a scientific context and in a socio-scientific context. Again, data were collected by audiotaping the teachers and

[1] Due to a set of factors such as changes in teachers timetable and occasional technical problems, a complete data set does not exist for all lessons.

videoing the same set of four pupils, wherever possible (11 teacher tapes, 22 student videos). In addition, a set of exactly similar data was collected from the control group for argumentation in a socio-scientific context for comparison purposes (5 teacher tapes, 10 student videos). Field notes were also collected of salient features of the lesson and the materials used by the teachers.

Finally, a semi-structured interview was also conducted with the teachers at the beginning of each year to ascertain their views on argumentation and to explore their reflections on the zoo lesson. These data sought to identify teachers' perceptions of the salience of teaching argumentation to pupils and their understanding of its significance. Such interviews were also used as a means of identifying any changes that had occurred over the year. Each interview was recorded and transcribed. The interviews included questions on how teachers felt about their zoo lesson and what they viewed as important for student participation and learning of argumentation. As the focus of interest in the second phase was the students development with argumenation, no final interview was conducted but a group discussion was held at the end of the project which was also recorded and transcribed.

Phase 1: Assessing Teachers Development with Argumentation

This section briefly summarises the work and analysis undertaken in phase 1 of the project with the group of 12 teachers. A fuller description of this aspect of the work and its findings can be found in Simon, Erduran and Osborne (submitted, 2002).

Analyses

All of the audiotapes were transcribed and analysed to determine the nature of argumentation in the whole class and the small group student discussion formats. The analysis of the teacher transcripts sought to answer our second question – that is what development had taken place in the teachers' use of argumentation in the classroom.

The approach taken to the analysis of the teachers' discourse was to use Toulmin's (1958) model of argument as an analytical framework to identify the salient features of argument in the speech. This required an extended process of defining and elaborating how this framework should be interpreted and used. The following section illustrates our method of coding the transcripts using TAP as a guiding framework. In the case of the following example of pupil discourse:

'Zoos are horrible, I am totally against zoos'

our focus would be on the substantive claim. In this case, the difficulty lies in the fact that both can be considered to be claims i.e.

'Zoos are horrible' *and* *'I am totally against zoos'*

The question for the analysis then becomes which of these is the substantive claim and which is a subsidiary claim. Our general view is that there is inevitably a process of interpretation to be made and that some of that process is reliant on listening to the tape and hearing the force of the various statements here. Part of this might be substantiated by Austin & Urmson's (1976) distinction between locutionary statements – ones which have an explicit meaning and perlocutionary statements – ones which have implicit meaning. The perlocutionary force with which these statements are made – something which can often only be determined by listening to the tape – is an aid to resolving which statement is intended as the substantive claim and the locutionary meaning.

Here our reading is that the emphasis lies on the second part of the statement because the task context demands a reference to a particular position (for or against zoos) and that this is therefore the substantive claim. In choosing to use TAP in this manner, we have developed a good reliability (more than 80 %) between the coders.

As an example, consider the following case between the student and the teacher.

S I've got a con. If the animals are always walking about in the same places they might get angry and be dangerous.

T Right, this is an anti, is it? So, being caged may alter their behaviour.

The position represented by the student is 'against zoos' expressed as a claim in the phrase: "I've got a con." The student further adds to this claim by saying that "if the animals are always walking about in the same places, they might get angry and be dangerous." This elaboration, we consider as data to support his claim. The teacher's subsequently interprets and justifies the choice for data by saying that "being caged may alter their behaviour." We regard the teacher's contribution as the warrant to the argument which is being constructed. Such a co-construction of arguments between students and teachers was typical in all the transcripts we have studied in our project. Thus, our approach to the work was always to seek to identify, through either a careful reading of the transcript, or alternatively, listening to the tape, what constituted the claim. Once, the claim was established, the next step was the resolution of data, warrants and backings. Our view here is that a necessary requirement of all arguments that transcend mere claims is that they are substantiated by data. Therefore, the next task is the identification of what constitutes the data for the argument which is often preceded by words such as 'because', 'since' or 'as'. The warrant, if present, is then the phrase or substance of the discourse which relates the data to the claim.

Nevertheless, in undertaking this task, we were conscious of the methodological difficulties in using TAP as a method of determining the structure and components of an argument (Kelly, Drucker, & Chen, 1998). Our view is that the task was made somewhat easier here by the highly framed nature of the task which foregrounded the substantive claim – that we should/should not build a new zoo coupled with a significant element of time required to develop a good understanding of the elements of the Toulmin framework and its important signifiers. Reducing these difficulties was, therefore, a significant methodological challenge for the second phase of our work.

Lesson structure and features of teacher talk

Lesson structures were determined by viewing video material of each Zoo lesson and noting the main lesson phases and time spent in whole class and small group formats. Viewing was accompanied by a study of the transcript of the audiotape. Extracts of teacher talk focusing on aims and organisation of argument activity or facilitation of the processes of argument were identified and summarised for each phase of the lesson. For example talk focusing on a lesson aim, such as the extract beneath was coded as ‘introduces aim of task, to produce good arguments’.

> ‘And we are trying to think this morning about what sorts of things will make a good argument. How are you going to persuade this agency that yes, the zoos should be opened? You need to put forward strong arguments, or if you don’t want it, strong arguments against the zoo.’

Such talk is an indicator of the ways in which the teachers view the nature and teaching of argument, and how they view the learning process. In essence it provides insights into teachers’ beliefs, practices (Fullan, 2001), value congruence and knowledge and skills (Harland & Kinder, 1997) and how these may have changed in one year. In addition, it has helped us to identify some of the features of the explicit practice of teachers and identify those aspects of the discourse of the teachers who showed a significant improvement in the quality of their argumentation during the course of the year. These teachers were identified as making more use of the language of argumentation using explicit calls to develop arguments that contained justifications composed of data and warrants. A fuller description is offered in Simon, Erduran and Osborne (submitted, 2002).

Interviews

In contrast, the teachers’ views about argumentation and its implementation in the classroom were analysed using a grounded approach where a coding schema was developed to capture the major themes emerging from the teacher interviews, with reliability checks undertaken by two members of the research team. These coded themes were examined and cross-referenced to the data from the TAP analysis of teacher talk in the lessons. A particular focus of analysis was any comments relevant to teachers’ classroom actions or talk about argumentation including the ways in which they conceptualised the teaching of argument, the decisions they made about teaching strategies, and their reflections on students’ progress and performance with argumentation. Again, since the principal focus of this paper is on the changes achieved by the pupils over the course of the second year, a fuller discussion of these data is presented elsewhere.

Results: The Changes in the Teachers

Each teacher implemented the same activity one year apart with comparable students. The lessons were similar in structure in that there was an introduction, group discussions, group presentations and, finally, a homework assignment in both years. The analysis of the data obtained from the transcripts of the lessons of two teachers for the two years are summarized in Figures 2 and 3. The x-axis indicates the features of Toulmin’s argument pattern (TAP) that were used in different combinations. For example, CD indicates those instances where a claim (C) was coupled with data (D). CDWB indicates that there was a claim, data, warrant and backing as part of one argument presented. The y-axis illustrates the frequency of instances that such permutations of TAP occurred within the transcript. In other words, we counted the number of times each sort of TAP occurred in the data across both years for each teacher.

Overall, the figures illustrate the nature of progression of teachers across two years. Going from left to right on the x-axis, there is an increasing complexity in the way that TAP is constructed, i.e. the inclusion of warrants, backings, rebuttals. Hence, counts on the right side of the charts indicate, in our view, an improvement in the nature of argumentation in that they contain more components of TAP. Therefore, a shift, for example, from CD (claim-data) to CDW (claim-data-warrant) across two years represents an improvement in the arguments constructed in the class format. Using this approach to analysis for all the teachers, we have produced a profile of the discourse of argumentation for all the teachers across the two years (Table 2).

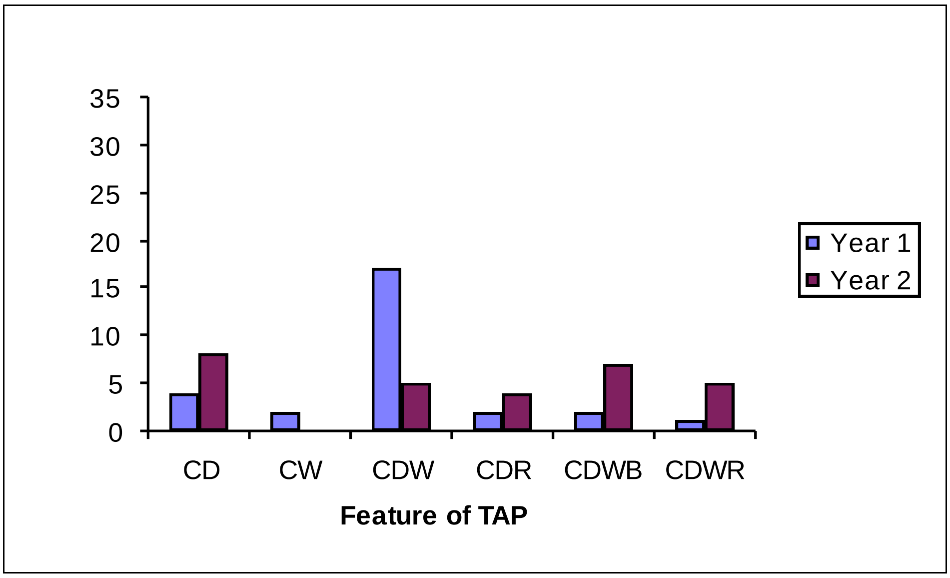

Figure 2: Sarah Year 1 vs 2 – No of Instances of each type of argument in the Zoo lessons

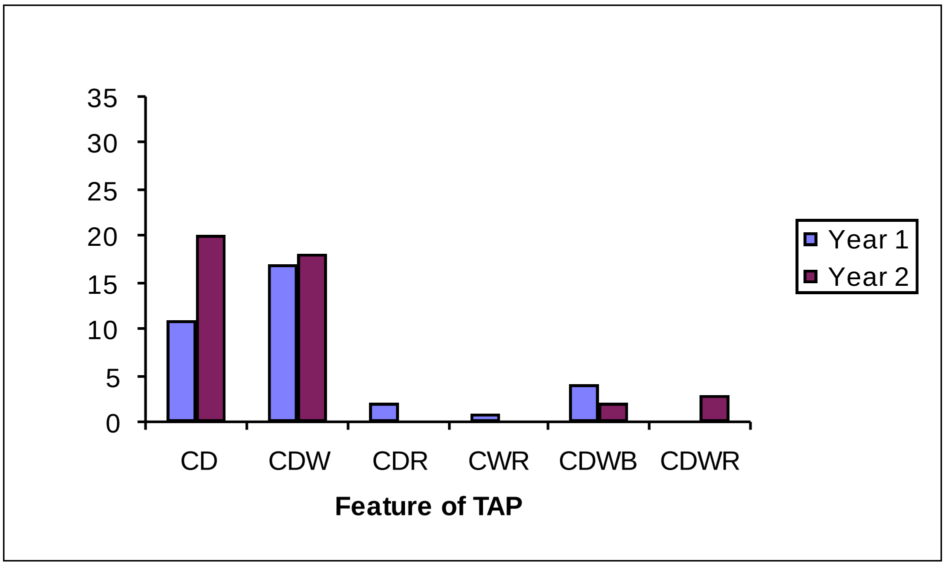

Figure 3: Matthew Year 1 vs 2 - No of Instances of each type of argument in the Zoo lessons

Although this analysis was undertaken for all teachers and the data for these two teachers have been chosen as they illustrate the major trends. First, there was argumentation discourse in the classroom across both years. In the figures we see specific examples of the extent to which each teacher's class is involved in the construction of which aspects of TAP. In other words, we can trace the nature of different permutations of TAP in their lessons. Using this approach reveals the second feature which is that, for each of the teachers, the pattern of use of the different permutations of TAP is *similar across the two years*. However, a comparison of Fig 2 and Fig 3, illustrates that, whilst the patterns are individually homogeneous, those for the teachers as a whole are heterogeneous and *different from each other*. This finding would suggest that there is no common pattern to the use of argumentation, rather, that the pattern is unique to each individual teacher. Thus, the use of argumentation is teacher dependent – in other word, that there are no universals. Third, the two profiles shown above are not only quite different in terms of the pattern of use the features of TAP, but also in terms of their change from one year to the next, Sarah's profile shows a marked shift to the right whereas Matthew's shows very little change.

A full analysis of the teachers' lesson transcripts, summarised in Table 2, shows how the discourse of the classroom was dominated by arguments that contain fewer elements of TAP which are less elaborated. The important detail, nevertheless, is that there was a significant ($p<0.01$) improvement in the overall pattern of discourse between year 1 and year 2 with more elaborated arguments being used by some or all of the teachers. Closer analysis shows that this change was a result of the changes made by 8 of the twelve teachers and that for 4 teachers there was no significant change.

Analysis of teachers' classroom talk and interview data suggests some possible explanations for the differences between their TAP profiles, and the variation in shift to the right from one year to the next between teachers. One critical difference in the classroom talk of these two teachers lay in the emphasis placed on counter-argument. Sarah introduced this aspect of argumentation and encouraged it strongly. Matthew, however, entirely omitted reference to opposition and counter-argument in his teaching. He did not encourage students to rebut claims or produce further evidence in the face of opposition. The absence of this feature of practice suggests a possible explanation for the difference between the two TAP profiles. For rebuttals force students to elaborate their arguments with backings or counter arguments to defend their view leading to an improved quality of argumentation and such elements of argumentation are more frequent in the argumentation of Sarah's lessons.

Table 2: Profile of argumentation discourse for the classrooms of all the teachers from year 1 to year 2.

Teacher	*Year*	*CD, CR*	*CDW, CDR*	*CDWR, CDWB*	*CDWBR*	*Sig*
Jeremy+	Year 1	48	47	5	0	
	Year 2	59	27	14	0	*
Peter	Year 1	41	47	10	2	
	Year 2	23	31	38	8	**
Maureen	Year 1	36	43	21	0	
	Year 2	43	43	14	0	
Frances+	Year 1	33	9	49	9	
	Year 2	52	3	42	3	*
Jules	Year 1	0	82	18	0	
	Year 2	8	44	44	4	**
Patrick	Year 1	48	38	14	0	
	Year 2	25	57	16	2	**
Mary+	Year 1	20	70	10	0	
	Year 2	0	50	50	0	**
Annie+	Year 1	48	32	16	4	
	Year 2	5	85	10	0	**
Sarah+	Year 1	21	68	11	0	
	Year 2	28	31	41	0	**
Katie	Year 1	32	47	16	5	
	Year 2	38	43	19	0	
Jason	Year 1	36	48	16	0	
	Year 2	41	41	14	4	
Matthew+	Year 1	31	57	12	0	
	Year 2	46	42	12	0	
Totals						
Year 1		394	588	198	20	
Year 2		368	497	314	21	**

* Significant at $p<0.05$
** Significant at $p<0.01$
+Teachers with whom we continued working in year 2 of the project.

Sarah's classroom talk demonstrated more shifts in emphasis across the two years than Matthew's, which remained similar from one year to the next. Sarah communicated the aim of the task in year 1 as 'thinking of ideas', whereas in year 2 she immediately focused on the process of producing a good 'strong' argument with evidence. A second change occurred in the way she encouraged students to focus on good argument, opposition and counter-argument – an aspect which she emphasised much more in year 2. A third change in Sarah's practice occurred in the way she set up the group task in year 2. Here, she introduced role-play, where students were to be different members of the community, and tried to encourage them to anticipate opposing arguments:

> First thing you need to do in your pair is decide whether that person will agree with the opening of the zoo or be against the zoo. Then what you need to do is to think of what that person's main argument will be and what the evidence they will have to support their idea, you then need to give another argument they might have and the justification they might have for that. And finally, and this is quite important, you need to think what someone opposing the argument might say. What their argument would be - the person that's going to disagree with you. What might their argument be? And how would you persuade them you were right? That's very important, that last bit.

Her attempt at role-play in year 2 shows she was willing to take risks and try new approaches , a development recognised as indicative of teacher change (Loucks-Horsely et al, 1998). Analysis of Sarah's interview provides further insights into her changing practice. Though she was aware at the beginning of the project of the value of what she termed 'saying the opposite', she developed this idea much more during the course of the year. In terms of her own professional development, Sarah thought that teaching argumentation had made her 'a lot more conscious' about what she was saying and what she was trying to achieve in her teaching. She also valued argumentation for the way it provided a challenge for the students.

The changes in Matthew's talk were less dramatic, demonstrating a shift from telling students about evidence to evoking more extended answers from the students themselves. Matthew's changes did not extend the processes of arguing, rather, they resulted in a more refined pedagogy emphasising the same processes. Throughout both years Matthew had a strong focus on the use of evidence to justify arguments, and in the second interview, the extract beneath shows that he judged his own progress in terms of how he valued the use evidence:

> I now look much more critically at, both in teaching and setting homework, for questions which require more reasoning and evidence….whereas in the past I might have thought - well, that's going to be too difficult for them… I think I appreciate the importance of trying to ensure that students see a difference between a statement and a reason for that.

Our view is that whilst Matthew had internalised the need to expose the epistemic justification for many more of the claims advanced by the science teacher, he had not appropriated the means to stimulate such thinking in his students. Hence, the lack of any significant change in the classroom discourse. A much fuller discussion of the issues raised for teachers and their development can be found in Simon, Erduran and Osborne (2002, submitted)

Phase 2: Assessing students development with argumentation

The analysis of the student group discussions sought to answer our third question – that is what development had occurred in the quality of the pupils' ability to argue and reason in a scientific and socio-scientific context. In evaluating student discourse for argumentation of good or better quality, our essential position is a commitment to the development of rational and analytic thought and discourse. In that we share with Toulmin a belief that:

> A [person] demonstrates his rationality, not by a commitment to fixed ideas, stereotyped procedures, or immutable concepts, but by the manner in which, and the occasions on which, he changes those ideas, procedures, and concepts. (Toulmin, 1972, p. v)

Changing one's thinking is not possible unless there are opportunities to externalise your arguments and to hold up one's beliefs and their justification for inspection by others. In that sense, we feel that one of the major achievements of our work has been to permit and encourage deliberative and dialogical interactions between pupils. Such opportunities are rarely a feature of school classrooms which, in contrast, are dominated by monological interactions, triadic dialogue (Lemke, 1990) and transactional interactions (Barnes, 1976). Hence, in our analysis of the data obtained from the small groups of pupils in the second year, we sought to see whether had been opportunities for pupils to engage in deliberative discussion of a dialogic nature. This was done by examining the transcripts and categorising the talk into one of four categories: teacher talk; student talk which advanced claims only; student talk which consisted of claims *and* grounds; and student talk which was non-argumentative which was of a procedural or off-task nature. A sample of such talk with its coding is given beneath:

Teacher:	OK. So you are saying that if the moon is light the light is fire and fire needs oxygen. All right. That's kind of added to the stuff......... you have actually talked about...... but what Mark was saying about the shadow and light, the moon passes through a shape, it is in the shadow from the earth, and you can't see it. So we know it doesn't give out light. B, the moon shrinks . Let's discuss this one. Michael..... why is the moon..	*Teacher talk*
Pupil:	The moon is solid and it can't expand.	*Student Claim*
Teacher:	It can't expand. What were you saying about water?	*Teacher talk*
Pupil:	It can't expand because it hasn't got water on it.	*Student Claim with Grounds*

Counts of the number of words uttered were then made and a sample of results for one for all the tapes analysed so far is shown in Table 3. Methodologically, the procedure that we have adopted here makes a distinction between what we see as first order elements of an argument – that is claims, grounds and rebuttals – and the second order elements which are the components of grounds, the data, warrants and backings. The advantage of using such a schema is that it circumvents the main methodological difficulty of Toulmin's framework – the resolution of the second order components. In this analysis, rebuttals were not separately identified either as we were simply interested in identify instances of argumentative discourse with or without grounds to obtain some measure of the space available for argumentation. The table shows the type of discourse in each type of lesson – the zoo lessons that took place at the beginning of the year as did the Science 1 lessons – and the Science 2 and Leisure Centre lessons which were at the end of the year.

Table 3: Table showing percentages of group discourse of an argumentative nature

	LESSON	*Zoo Lesson* *%*	*Zoo Control* *%*	*Science 1* *%*	*Science 2* *%*	*Leisure Centre* *%*	*Leisure Centre Control* *%*
Type of Discourse	Claims	4	3	4	5	4	3
	Grounds	28	24	11	12	22	26
	Non-Argument	8	8	15	9	13	10
	Teacher	59	64	71	75	61	61

The data illustrate several features of the nature of the discourse in these lessons. First, previous research has shown that deliberative discourse of a dialogic nature commonly occupies 2% or less of all classroom discourse (Newton, Driver & Osborne, 1999). These data, however, show that in these lessons argumentative discourse (claims, claims + grounds) now occupies 13% to 32% of the total discourse which represents a major shift away from the normative form of authoritarian dialogue that permeates science classrooms. The second notable feature of these data is that argumentative discourse is significantly less for argumentation in science lessons than it is for socio-scientific lessons suggesting that initiating argument in a scientific context is harder and more demanding both for pupils and their teachers. The data also show that there is little difference in the amount of discourse between the experimental groups and the control groups suggesting that the amount of argumentative discourse is a feature of the teachers' structuring and organisation of the lesson rather than any feature of the groups.

Assessing the quality of argumentation

In seeking to answer our third research objective we have focussed on the discussions between pupils. In each class, the teacher identified two groups of 3 to 4 pupils and their discussions were taped and transcribed. The transcripts were then searched to identify genuine episodes of oppositional analysis and dialogical argument. Opposition took many different forms and many arguments where co-constructed where students provided data or warrants for others' claims. Transcripts of group discussions (2 groups per teacher) were examined to determine the number of episodes of explicit opposition in student discourse. In other words, the instances where students were clearly opposed to each other were traced. Typically these instances were identified through the use of words such as "but", "I disagree with you", "I don't think so". Once these episodes were characterized in the group format, they were re-examined for the interactions among the students in terms of who was opposing whom, who was elaborating on what idea, or reinforcing and repeating an idea. In this fashion, the pattern of interaction for each oppositional episode was recorded for two groups from each teacher's classroom. The main processes identified in such episodes were opposing claims by other (O), elaboration (E) or reinforcement (R) of a claim with additional data, warrants, advancing claims (C) or adding qualifications (Q). Such analysis helps to identify the features of the interaction and the nature of the engagement between the students.

The Nature of Opposition

Each oppositional episode was analysed using TAP to identify the principal components of an argument being deployed by the individuals in the group. In these episodes, claims were not always clearly stated but implied or extracted through questioning. All episodes were read independently by two coders who then met to compare their analysis and resolve differences in interpretation. These oppositional episodes are characterised by a diverse range of arguments and some examples are provided later to illustrate the nature of our analysis and the results. The essential issue raised by these episodes is how to define their quality. What, for instance, makes one better than another? To answer this question, we have developed a framework for the analysis of quality which is outlined beneath. A fuller description of our methodological approaches can be found in (Erduran, Osborne and Simon, submitted, 2002).

In approaching this task, we drew on the fact that we had found little problem in the identification of claims or rebuttals but the distinction between data and warrants was often hard to make as it depended on contextual information which was either absent from the transcript or impossible to determine unambiguously from the video. Our schema for argumentation therefore transcends this problem by requiring only the resolution of first order components of argument, that is the claims, justifications and rebuttals, avoiding the necessity to resolve the second order components of data, warrants and backings required to use the full analytical framework of TAP.

In establishing this framework, we have drawn two major distinctions. The first is does an argument contain any reasons and grounds i.e. data, warrants or backing to substantiate its claim, as transcending mere opinion and developing rational thought is reliant on the ability to justify and defend one's beliefs. Hence, we see the simplest arguments as those consisting of a claim. Some such as Zohar & Nemet (2002) would not wish to recognise claims without justifications as meriting any significance. However, we feel that they are important as they are the first step to initiating the process of establishing difference. Whilst we recognise that the opposition may simply consist of counter-claim – essentially a discourse transaction which is incapable of any resolution, such moves permit the establishment of difference and higher quality argumentation. In addition, teachers need to be able to identify such discourse moves and expose their limitations – the lack of justification – to their students. Hence, our second level is arguments accompanied by grounds containing data or warrants[2] followed by arguments consisting of claims, data, warrants and rebuttals.

Episodes with rebuttals are, however, of better quality than those without for oppositional episodes without rebuttals have the potential to continue forever with no change of mind or evaluation of the quality of the substance of an argument. Moreover, as Kuhn (1991: 145) argues the ability to use rebuttals is 'the most complex skill' as an individual must 'integrate an original and alternative theory, arguing that the original theory is more correct.' Thus, rebuttals are an essential element of arguments of better quality and demonstrate a higher-level capability with argumentation. This analysis has led us to define quality in terms of a set of 5 levels of argumentation (Table 4) as follows:

[2] In analysing argument, our view has been that any argument that transcends a mere claim must contain an item of data. Arguments that contain only warrants without data are very difficult to construct and have rarely been observed.

Table 4: Analytical Framework used in for assessing the quality of argumentation

Level 1:	Level 1 argumentation consists of arguments that are a simple claim v a counter claim or a claim v claim
Level 2:	Level 2 argumentation has arguments consisting of claims with either data, warrants or backings but do not contain any rebuttals.
Level 3:	Level 3 argumentation has arguments with a series of claims or counter claims with either data, warrants or backings with the occasional weak rebuttal.
Level 4:	Level 4 argumentation shows arguments with a claim with a clearly identifiable rebuttal. Such an argument may have several claims and counter claims as well but this is not necessary.
Level 5:	Level 5 argumentation displays an extended argument with more than one rebuttal.

The following two examples are provided, then, to illustrate how our analysis has been applied to the data.

Episodes without rebuttals

Example 1

In this example, taken from the zoo lesson, a claim is advanced supported by some data.

S1: I don't think they would hurt them in a professional zoo.

S2: But they might scare the other animals by seeing some sedated animal being dragged off.

S1: Maybe stress.

S3: Not stress. Distress.

Here, what we have is a claim that professional zoos would not hurt animals which is countered by claim that animals in zoos might be scared (claim) as they would see other sedated animals being dragged off (data). Thus, our summary of this example is that it consists of:

claim v counter claim + data

Moreover, despite some embedded complexity, as an example of arguing we would contend that it is essentially weak as there is no attempt at a rebuttal (by either party) permitting the justification of belief by both parties to remain unexamined. Therefore, we would consider this to be a level 2 argumentation.

Episodes with rebuttals

Our essential distinction here is between episodes with weak rebuttals – that is counter arguments which are only tenuously related to the initial claim (level 3), episodes with a single rebuttal (level 4), and episodes with multiple rebuttals (level 5). Example 2 illustrates a case of a weak rebuttal and example 3 a clear, unambiguous rebuttal.

Example 2

The episode beneath begins with the implicit claim that zoos are beneficial. The data for this argument is that 'some animals wouldn't be able to breed in the wild' and there is a warrant supplied that this is because 'they may not have enough food'. This claim is further supported or elaborated by the claim that 'the animals need a safe place to live' and the data to support this claim is that otherwise 'they will be at risk from predators'. This second claim is weakly rebutted with a negation which is thinly supported by the data that the risk from predators is just 'nature'. However, as the rebuttal of the proponent's data does not make a clear, self-evident connection to the data supporting the original claim, we consider this to be an example of a weak rebuttal and a level 3 argumentation. A summary of this argument would be that it consists of :

claim (+ data + warrant) + claim (+data) v rebuttal (+ data)

O
1. Some animals wouldn't be able to breed in the wild, because they may not have enough food.
2. No, no, no, because an animal.........

E
3. Extinction.
4. The animal needs a place to live because they would be at risk from other predators.

E
5. What are you putting?
6. A place to live, or they would be at risk from other predators.

O
7. They might not have enough food to eat.

O
8. But I mean, that's nature, one has to....
9. But we are for it

Example 3

Our third example is an argument taken from a scientific context where pupils have been given alternative theories to explain the phases of the moons that are on numbered card, A, B, C, D, which are referred to in the dialogue.

E O E	M	.... A, the moon spins around, so the part of the moon that gives out light is not always facing us. Jamal, A?
	J	The moon doesn't give out light.
	M	Right, so that's why A is wrong. That's true. How do you know that?
	J	Because the light that comes from the moon is actually from the sun.
	M	He is saying the light that we see from the moon is actually a reflection from the sun. How do we know that? Mark?
	M	Because the moon is blocked by the…..

Here, the first pupil advances the claim that it is explanation A appealing to a datum that 'the moon does not give out light'. There is then a rebuttal supplied with supporting data that the 'light that comes from the moon is actually from the sun' and a warrant which is unfinished.

Our summary of this argument would be that it consists of:

Claim (+ data) v Rebuttal (+ data + warrant).

This schema of analysis enables us to make various comparisons of the performance of the different groups at argumentation. Fig 4 shows the distribution of arguments by level for all of the oppositional episodes currently analysed. In total, we have identified 183 oppositional episodes from 63 group discussions in 33 lessons for all the argument lessons. Thus, in summary, there were, on average, approximately 3 oppositional episodes per group per lesson. For the experimental groups, the chart beneath shows the distribution of the levels of arguments obtained from 43 discussion groups in 23 lessons

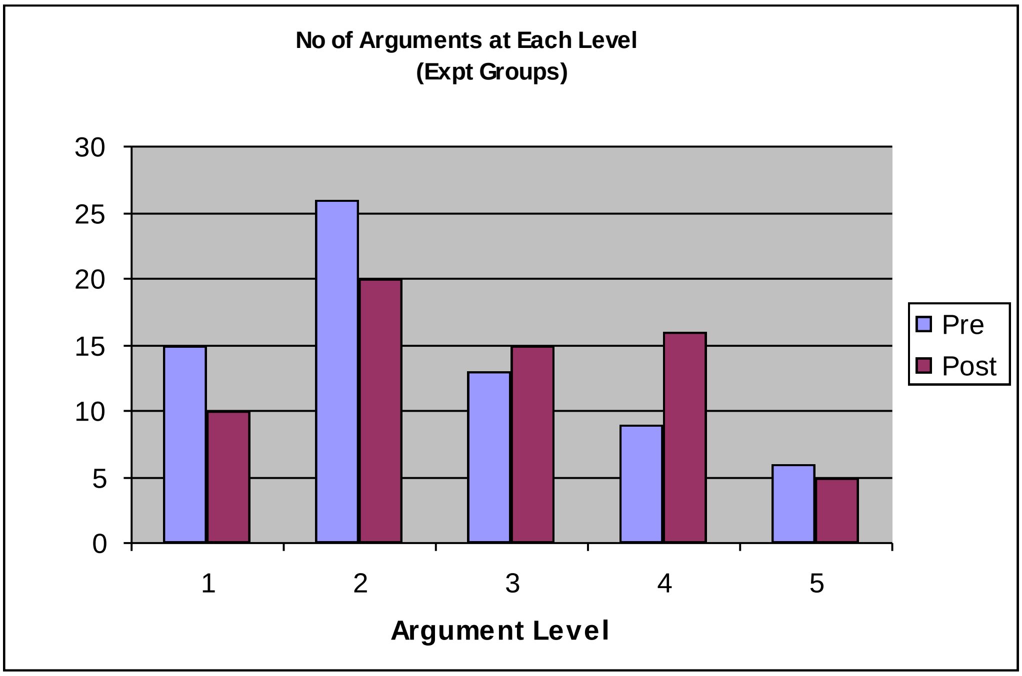

Fig 4: Chart showing numbers of each level of argumentation achieved in each oppositional episode (n=183)

This chart shows that the largest number of arguments emerging from the data both at the beginning and the end of the year was at level 2 (38% and 30% respectively). Encouragingly though, whereas at the beginning of the year only 40% of pupil arguments were at level 3 or above at the beginning of the year, by the end of the year, the corresponding figure was 55%. Whilst this change is not significant, it does show a positive development in the quality of argument. Moreover, the number of level 1 arguments has reduced from 22% to 15%. This finding is particularly encouraging as it suggests that only a small minority of arguments developed by pupils did not attempt to offer a rationale or some grounds for their claims and that the intervention has led to a diminishment in the number of such arguments. Pedagogically a preponderance of level 1 arguments would be problematic in that it is these types of argument that have the most potential for argumentation which is confrontational reinforcing the lay perception of 'argument as war' (Cohen, 1995). Rather, the metaphor of argument we chose to use in our work was of argument as a process of collaborative brainstorming towards the establishment of 'truth' or better understanding – the primary goal of science. A view succinctly summarised by Bachelard (1940)in his statement that 'two people must first contradict each other if they really wish to understand each other. Truth is the child of argument, not of fond affinity.'

This method of analysis permits a number of comparisons of the performance of the groups. Firstly, it is possible to compare the distribution of levels achieved by the experimental group, at the beginning of the year in the first zoo lesson and their first science lesson, with those achieved at the end of the year in the last science lesson and their final leisure centre lesson. Table 5 shows the data for this comparison.

Table 5: Levels of argumentation achieved by experimental groups, pre and post-intervention.

	Argument Level Achieved				
Lesson	**1**	**2**	**3**	**4**	**5**
Zoo & First Science Lesson (12 lessons & 23 groups)	22% (15)[1]	38% (26)	19% (13)	13% (9)	9% (6)
Leisure Centre (LC)& Second Science Lesson (11 lesson and 22 groups)	15% (10)	30% (20)	23% (15)	24% (16)	8% (5)

[1] Figures in brackets show raw data

This analysis for approximately two thirds of the data shows that there has been a shift towards the end of the intervention to more arguments of higher quality shown more clearly by Fig 5. However, this is not a significant shift.

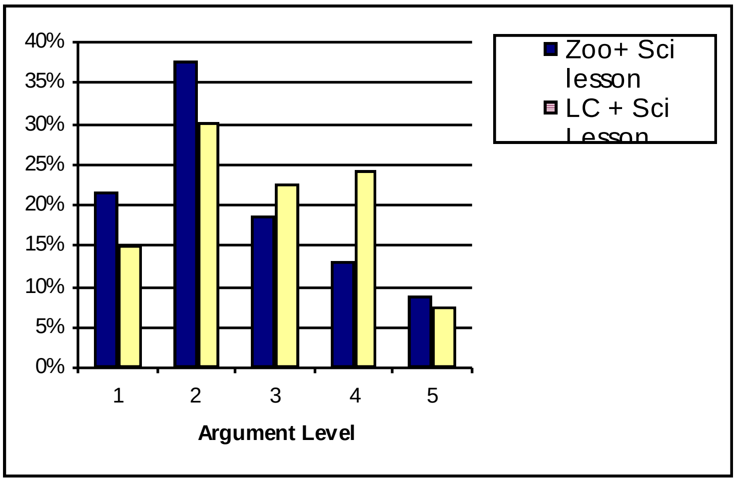

Fig 5: Levels of argumentation achieved by experimental groups, pre and post-intervention.

Likewise, table 6 shows a comparison of the levels of argument achieved by the groups in the discussion about the merits of zoos in the first zoo lesson with that 10 months later about whether a leisure centre should be placed in an area of well-established wildlife.

Table 6: Levels of argumentation (socio-scientific context) achieved at the beginning of the year (Zoo lesson) and at the end of the year (Leisure Centre lesson)

	Argument Level Achieved				
Lesson	**1**	**2**	**3**	**4**	**5**
Zoo Exp (6 lessons, 11 groups)	13% (5)	34% (13)	21% (8)	16% (6)	16% (6)
Leisure Centre Exp (6 lessons, 12 groups)	11% (4)	32% (12)	16% (6)	32% (12)	11% (4)

The difference between these two distributions is not significant although the pattern would again suggest that there were more high quality arguments at the end of the intervention than the beginning.

Likewise, it is possible to compare the levels of argument achieved at the beginning of the year in a scientific context with those achieved at the end of the year.

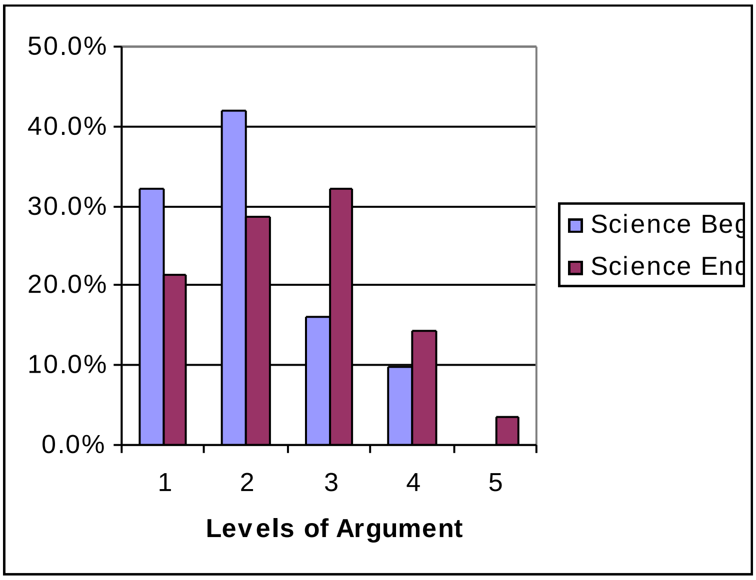

Fig 6: A comparison of levels of argumentation achieved by experimental groups in a scientific context at the beginning of the year (n=31) with that achieved at the end of the year (n=28).

Again this shows that there has been a positive improvement. Whereas at the beginning of the year 74% of the arguments were at level 2 or lower, at the end of the year, this figure had diminished to 50%.

One of the features of interest in this work was how the context of argument i.e. scientific or socio-scientific affected the quality of argument. A comparison of the levels of argumentation achieved in the socio-scientific lessons (zoo and leisure centre) with those achieved in the science lessons (Table 7) indicates that in general higher levels of argument are achieved in a socio-scientific context and that this difference is significant ($p<0.05$).

Table 7: Comparison of Levels of argumentation (socio-scientific context) with those achieved in a scientific context

	Argument Level Achieved				
Lesson	**1**	**2**	**3**	**4**	**5**
Zoo Exp & LC Expt (12 lessons, 23 groups)	12% (9)	33% (25)	18% (14)	24% (18)	13% (10)
Science Lessons (11 lessons, 22 groups)	27% (16)	36% (21)	24% (14)	12% (7)	2% (1)

Taken together, with our analysis (Table 2) of the discourse in lessons which showed that there was substantively less argumentative discourse in science lessons, these findings suggest that it is harder to initiate argumentation and argument in a scientific context than in a socio-scientific context. However, whether the quality of argumentation is dependent on the quantity of argumentative discourse remains an open question.

Another feature of the research design was the use of a set of control classes. As well as teaching the lessons to the treatment group, we asked each of the teachers to teach the same zoo lesson to a similar class at the beginning of the year, and the same leisure centre lesson to a class at the end of the year. This enabled a comparison to be made between the performance of the two groups at the beginning of the year and at the end of the year (Table 8).

The data would suggest that there was no significant difference between the groups at either the beginning or the end of the year. This finding would suggest that whilst the experimental group have shown an improvement in the quality of their argumentation, the control group also seem to have improved. However, a number of caveats must be placed on any interpretations of this data. First the sample size is very small; second, the similarity between the groups remains questionable as apart from asking for the teacher to select a class with pupils of similar ability, it was impossible to impose any other constraints or control other variables which might enable us to have made a more effective comparison between the two groups such as gender or ethnic mix; third, it would be unrealistic to place too much emphasis on a set of data which is collected from a very limited number of lessons at the end of the year.

Table 8: Comparison of Levels achieved by experimental groups with those achieved by controls.

	Argument Level Achieved

Lesson	1	2	3	4	5
Beginning of Year					
Zoo Exp (6 lessons, 11 groups)	13% (5)	34% (13)	21% (8)	16% (6)	6% (6)
Zoo Control (5 lessons, 9 groups)	18% (5)	46% (13)	25% (7)	11% (3)	0% (0)
End of Year					
Leisure Centre Exp (6 lessons, 11 groups)	11% (4)	32% (12)	16% (6)	32% (12)	11% (4)
Leisure Centre Control (5 lessons, 9 groups)	18.5% (5)	18.5% (5)	25.9% (7)	18.5% (5)	18.5% (5)

If, however, these comparisons are valid measures, then there are two hypotheses as to why both groups have improved and why no difference between the experimental and control group was found: a) that the improvement represents a natural developmental growth in individuals' reasoning and linguistic capability, or b) that the improvement is a reflection of the individual teacher's growing ability to structure and facility and argumentation. Our view is that there is insufficient data here to resolve this question and that, with hindsight, such methodological approaches have little value unless they are undertaken with considerably larger sample sizes. Rather, we would point to the fact that what we have attempted here is essentially a 'design experiment' (Brown, 1992) where the 'learning effects are not even simple interactions, but highly dependent outcomes of complex and social and cognitive intervention' – all of which have the potential to confound the data and their interpretation. More significant, we feel, is the fact that this intervention has achieved the positive effects we desired, albeit not as large as we might have hoped. In that sense, our intervention has demonstrated that such treatments are capable of improving young children's quality of argumentation.

Discussion & Conclusions

In this paper, we have presented the major findings emerging from our work on developing argumentation in school science classrooms, its analysis, and the assessment of its quality. Methodologically, we feel our work has made progress on several fronts. First the work has sought to develop with teachers sets of materials that can be used in a structured and focussed manner to facilitate argumentation in the classroom. As a result of this experience, we feel that we have gained some insights into the means of establishing a context which facilitates argumentation in the classroom both in terms of the materials and the pedagogic strategies required for its support. Hence, in the next phase of our work, we will now be attempting to develop and disseminate such materials through the ideas, evidence and argument in science (IDEAS)[3] project. This project is rooted in the belief that a major barrier to the uptake and dissemination of such work is the lack of good examples modelling the implementation of innovative practice (Joyce, 1990). Therefore, using the teachers we have worked with in both phases of this work, our intention is to video them implementing argumentation illustrating how such lessons are

3 Funded by the Nuffield Foundation

organised and the key features of practice. In addition, we will be developing packs of materials which provide both a teacher's handbook and classroom materials and providing materials that help to develop teachers' underlying theoretical understanding of the nature and function of argument in science. The latter we see as essential to developing the value congruence (Harland & Kinder, 1997) that argumentation is an important aspect of science and science education.

Second, our work with teachers has led to a change in the practice of the majority of this group leading us to believe that, despite the many obstacles and barriers posed by the demands to implement different and innovative practice, it is possible for science teachers to adapt, change and develop their practice to one where there is a fundamental change in the nature of classroom discourse. One of the biggest fears expressed initially about this kind of work by some of the teachers was that the presentation of plural explanatory theories would confuse the children or lead to the development or strengthening of a belief in scientifically incorrect idea. Such a reaction is very comprehensible for the rhetorical project of the science teacher is to present a carefully crafted and persuasive argument for the scientific world-view (Osborne, 2001). Presenting alternatives to the scientific explanation would, at first hand, seem to undermine that project and naturally generate hesitation and doubt if not resistance in teachers. Yet, it was notable in the interviews at the end of the project that this initial concern was much diminished, if not absent altogether. In short, the teachers had come to recognise that the opportunity for students to reflect, discuss and argue how the evidence did, or did not, support the theoretical explanation made debating the scientific case after the argumentation lesson, a much simpler task with which students were already engaged.

Third, one of the many problems that bedevils work in this field is a reliable systematic methodology for a) identifying argument and b) assessing quality. Our adoption and adaptation of Toulmin's Argumentation Pattern has also provided us with a method for discriminating the salient features of argumentation – the claims, rebuttals and justifications – which are critical for developing and evaluating practice with argumentation in the classroom (Erduran, Osborne & Simon, submitted). This is not to say that the full Toulmin framework is of no value. Currently, at least in the UK, the language used to describe the epistemic components of science is that of the 'ideas' of science and their supporting 'evidence' (Department for Education and Employment, 1999) 'Ideas', on the one hand, consist of hypotheses, theories and predictions that are essentially claims, whilst the data, warrants, backings, rebuttals and qualifiers are the components and conditions of 'evidence'. The use of these features of TAP offer teachers a richer meta-language for talking about science and for understanding the nature of their own discipline – and a language that we would urge the community, especially those engaged in teacher training or professional development to adopt.

.

More importantly, our work using TAP, and our focus on the *argumentation* rather than the content of arguments themselves has enabled the evolution of a workable framework for the analysis of the quality of the process in the classroom. To date most of those working in the field have focussed on the *content* of an argument and its logical coherence. Our preference, in contrast, has been to examine the *process* of argumentation as this is the foundation of rational thought and to examine whether that process can be facilitated and its quality assessed.

We have also illustrated how we can apply this schema to sets of data obtained from teachers implementing argumentation in the classroom. These data sets do show evidence of positive improvement in the quality of student argumentation, however, this change has not been significant. This would suggest to us that developing the skill and ability to argue effectively is a long-term process – something which only comes with recurrent opportunities to engage in argumentation throughout the curriculum rather than the limited period of 9 months of our intervention. Our findings, admittedly, stand in contrast to those of Zohar and Nemet (2002) who found significant improvements after a relatively short intervention for which we have no explanation. However, they are supported by the work of Zoller et al. (2000; 2002) who concluded from their work with 1st year college undergraduates that one semester is too short a period to develop higher order cognitive thinking and that, rather, systemic longitudinal persistence is necessary to achieve significant outcomes. The main messages, nevertheless, is that all of these studies, including our own, show that *improvement at argumentation is possible if it is explicitly addressed and taught.* Thus, it is possible for science education to make a significant contribution to improving the quality of students' reasoning redressing the weaknesses exposed by the work of Kuhn (Kuhn, 1991) and Hogan and Maglienti (2001)

Finally, our data give a clear indication that supporting and developing argumentation in a scientific context is significantly harder than enabling argumentation in a socio-scientific context. Our own view is that argumentation of quality is dependent on a body of appropriate knowledge that can form the data and warrants of an individual's arguments. In the context of socio-scientific issues, pupils can draw on ideas and knowledge developed informally through their own life-world experiences, and a sense of ethical values and economic considerations. In contrast, argument in a scientific context requires very specific knowledge of the phenomenon at hand. Without this resource, constructing arguments of quality will be severely restricted and hampered. Thus, supporting scientific argument in the classroom requires that relevant evidence must be provided to pupils if arguments of better quality are to be constructed and evaluated. Some will see that as an argument to defend the status quo – that students must acquire a knowledge of the major components of the scientific canon before they can engage in discourse activities that resemble or model those of the professional scientist. This is an argument we would refute for two reasons. First, because even the simplest scenarios can engage students in epistemic activities that closely model that of professional scientists. What is essential is that the process is scaffolded with a body of relevant evidence which students can then consider and martial to support one theory or another. So, for instance, students can consider whether day and night are caused by a spinning Earth and moving Sun. Data for consideration can be that the Sun appears to move; that when you jump up you land in the same spot; that it is night time in Australia when it is daylight in Europe; that the Earth is not an exact sphere but slightly wider at the Equator; that a long pendulum does not swing in the same plane all day and more. Dividing students into groups and asking them to argue the case for one view or the other, and to think how they would argue against any items of evidence that are not supportive of the theory they are defending requires thought and develops students' critical thinking. The only legitimate moral requirement of the teacher is that they ensure that all students have some knowledge of this data – none of which is excessively demanding. The work of Keogh and Naylor (1999) on concept cartoons has shown that there are many more natural phenomena which can also be a locus of argumentation from an early age. Second, it has been our experience, and that of others (Ogborn, Kress, Martins, & McGillicuddy, 1996), that opportunities to engage in argumentation generate student engagement - the sine qua non of significant learning.

Once engaged scientific arguments about whether plants get their 'food' from the soil or from the air have, we believe, a much greater chance of being seriously considered and assimilated.

Perhaps most significantly, however, we see our work not in isolation but as part of a growing body of work in this area (Herrenkohl & Guerra, 1995; Herrenkohl & Guerra, 1998; Kelly & Crawford, 1997; Kelly et al., 1998, Zohar and Nemet, 2000) that has begun to explore the difficulties and dilemmas of introducing argument to science classrooms – work which attempts to offer some insights into how practice can be developed. Contemporary research has guided many educational researchers to conceive of thinking and reasoning as acts that are socially driven (Brown, Cobb, 1994; Rogoff, 1990), language dependent (Wertsch, 1991), governed by context or situation (diSessa, 2000; Brown, Collins and Durgid, 1989); and involving a variety of tool-use and cognitive strategies (Edelson, Gordin, & Pea, 1999; Kuhn, 1999). Putnam and Borko (2000), in an article that examines the challenges these new ideas about knowledge and learning have for teacher education, summarize these newer conceptions of learning respectively as cognition as social (in that it requires interaction with other), cognition as situated (in that it is domain specific and not easily transferable), and cognition as distributed (in that the construction of knowledge is a communal rather than individual activity). Nevertheless, a missing crucial component of this body of research is any significant evidence demonstrating that engaging in discursive problem solving activities leads to enhanced cognition – one of the major goals of any education. Having established a modus operandi for argument in the classroom, and demonstrated that student skills at argumentation can be enhanced, the question we ask is whether regular engagement in such activities over an extended period would lead to enhanced cognitive development? It is this question which further research in the field needs to address.

Finally, if science is the epitome of rationality and as a corollary the commitment to evidence now permeates the discourse of contemporary life, then exposing the nature of the arguments and epistemic thinking that lies at the heart of science is a growing imperative of any contemporary science education that seeks to establish its broader cultural value and significance. For not only is such reasoning a major constitutive element of science itself, and of contemporary discourse, but only through engaging in argumentation and such reasoning can we hope to achieve some of the aspirations of the many who would seek to realise Schwab's vision that science education should be an 'enquiry into enquiry'.

Acknowledgements
This study was supported by the UK Economic and Social Science Research Council grant number R000237915. In addition, we would like to acknowledge the many teachers who have worked with us on this project and their efforts with this research.

References

Aikenhead, G. S. (1991). *Logical reasoning in science and technology*. Toronto, Ontario: John Wiley of Canada.
Alexopoulou, E., & Driver, R. (1997). Small group discussions in physics: peer interaction modes in pairs and fours. *Journal of Research in Science Teaching, 33*(10), 1099-1114.
Alverman, D. E., Qian, G., & Hynd, C. E. (1995). Effects of interactive discussion and text type on learning counterintuitive science concepts. *Journal of Educational Research, 88*, 146-154.
Andrews, R. (1995). *Teaching and Learning Argument*. London: Cassell.
Bachelard, G. (1940). *The Philosophy of No*. Paris: Paris University Press.
Barnes, D. (1976). *From communication to curriculum*. London: Penguin.
Beck, U. (1992). *Risk Society: Towards a new Modernity*. London: Sage.
Bell, P., & Linn, M. (2000). Scientific arguments as learning artefacts: designing for learning from the web with KIE. *International Journal of Science Education, 22*(8), 797-817.
Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale, New Jersey: Lawrence Erlbaum.
Billig, M. (1996). *Arguing and Thinking* (2cnd ed.). Cambridge: Cambridge University Press.
Boulter, C. J., & Gilbert, J. K. (1995). Argument and Science Education. In P. J. M. Costello & S. Mitchell (Eds.), *Competing and Consensual Voices: the theory and practice of argumentation* . Clevedon: Multilingual Matters.
Bruner, J. (1961) The Act of Discovery. *Harvard Educational Review*, *31*, 21-32
Cohen, D. (1995). Argument is War....and War is Hell: Philosophy, Education, and Metaphors for Argumentation. *Informal Logic*, *17*(2), 177-188.
Department for Education and Employment. (1999). *Science in the National Curriculum*. London: HMSO.
Dillon, J. T. (1994). *Using Discussion in Classrooms*. Buckingham: Open University Press.
Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young People's Images of Science*. Buckingham: Open University Press.
Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287-312.
Duschl, R. A. (1990). *Restructuring Science Education*. New York: Teachers College Press.
Eichinger, D. C., Anderson, C. W., Palinscar, A., & David, Y. M. (1991, April, 1991). *An Illustration of the Roles of Content Knowledge, Scientific Argument, and Social Norms in Collaborative Problem Solving.* Paper presented at the American Educational Research Association, Chicago.
Garratt, J., Overton, T., & Threlfall, T. (1999). *A Question of Chemistry: Creative Problems for Critical Thinkers*. Harlow: Pearson.
Geddis, A. (1991). Improving the quality of classroom discourse on controversial issues. *Science Education*, *75*, 169-183.
Gee, J. (1996). *Social Linguistics and Literacies* (2nd ed.). London: Taylor and Francis.
Giddens, A. (1990). *The Consequences of Modernity*. Cambridge: Polity Press.
Gilbert, J. K., & Watts, D. M. (1983). Concepts, Misconceptions and alternative conceptions: Changing perspective in Science Education. *Studies in Science Education, 10,* 61-98.
Goldsworthy, A., Watson, R., & Wood-Robinson, V. (2000). *Developing understanding in scientific enquiry*. Hatfield: Association for Science Education.
Hammer, D. (1995). Student inquiry in a physics class discussion. *Cognition and Instruction, 13,* 401-430.
Harland, J., & Kinder, K. (1997). Teachers' Continuing Professional Development: framing a model of outcomes. *British Journal of In-service Education, 23*(1), 71-84.

Herrenkohl, L., Palinscar, A., DeWater, L. S., & Kawasaki, K. (1999). Developing Scientific Communities in Classrooms: A Sociocognitive Approach. *The Journal of the Learning Sciences, 8*(3&4), 451-493.
Herrenkohl, L. R., & Guerra, M. R. (1995). *Where did you find your theory in your findings? Participant Structures, Scientific Discourse, and Student Engagement in Fourth Grade.* Paper presented at the AERA Annual Meeting.
Herrenkohl, L. R., & Guerra, M. R. (1998). Participant Structures, Scientific Discourse, and Student Engagement in Fourth Grade. *Cognition and Instruction, 16*(4), 431-473.
Hogan, K., & Maglienti, M. (2001). Comparing the Epistemological Underpinnings of Students' and Scientists' Reasoning about Conclusions. *Journal of Research in Science Teaching, 38*(6), 663-687.
Jiménex-Aleixandre, M. P., Bugallo-Rodríguez, A., & Duschl, R. A. (1997, March 1997). *Argument in High School Genetics.* Paper presented at the National Association for Research in Science Teaching, Chicago.
Jiménex-Aleixandre, M. P., Rodríguez, A. B., & Duschl, R. (2000). "Doing the Lesson" or "Doing Science": Argument in High School Genetics. *Science Education, 84*(6), 757-792.
Joyce, B. (Ed.). (1990). *Changing school culture through staff development: 1990 yearbook of the Association for supervision and curriculum development*.
Kelly, G. J., & Crawford, T. (1997). An ethnographic investigation of the discourse processes of school science. *Science Education, 81*(5), 533-560.
Kelly, G. J., Drucker, S., & Chen, K. (1998). Students' reasoning about electricity: combining performance assessment with argumentation analysis. *International Journal of Science Education, 20*(7), 849-871.
Keogh, B., & Naylor, S. (1999). Concept Cartoons, teaching and learning in science: an evaluation. *International Journal of Science Education, 21*(4), 431-446.
Koslowski, B. (1996). *Theory and Evidence: The Development of Scientific Reasoning*. Cambridge, MA: MIT Press.
Kuhn, D. (1991). *The Skills of Argument*. Cambridge: Cambridge University Press.
Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*(3), 287-315.
Lansdown, B., Blackwood, P., & Brandwein, P. (1971). *Teaching Elementary Science: Through Investigation and Colloquium.* New York: Harcourt Brace Jovanich, Inc.
Lemke, J. L. (1990). *Talking Science: Language, Learning and Values*. Norwood, New Jersey: Ablex Publishing.
Mason, L. (1996). An analysis of children's construction of new knowledge through their use of reasoning and arguing in classroom discussions. *Qualitative Studies in Education, 9*(4), 411-433.
Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction, 14*, 139-178.
Millar, R., & Osborne, J. F. (Eds.). (1998). *Beyond 2000: Science Education for the Future*. London: King's College London.
Minstrell, J., & Van Zee, E. (Eds.). (2000). *Teaching in the Inquiry-based science classroom*. Washington, DC: American Association for the Advancement of Science.
Monk, M., & Osborne, J. (1997). Placing the History and Philosophy of Science on the Curriculum: a model for the development of pedagogy. *Science Education., 81*(4), 405-424.
National Research Council. (2000). *Inquiry and the National Science Education Standards* . Washington D.C.: National Academy Press.
Naylor, S., & Keogh, B. (2000). *Concept Cartoons in Education*. Sandbach: Millgate House Publishers.

Newton, P., Driver, R., & Osborne, J. (1999). The Place of Argumentation in the Pedagogy of School Science. *International Journal of Science Education, 21*(5), 553-576.
Norris, S., & Phillips, L. (2001, March 26-28, 2001). *How Literacy in its Fundamental Sense is Central to Scientific Literacy.* Paper presented at the Annual Confernce of the National Association for Research in Science Teaching, St Louis.
Ogborn, J. (2002). Ownership and transformation: Teachers using Curriculum Innovation. *Physics Education, 37*(2), 142-146.
Ogborn, J., Kress, G., Martins, I., & McGillicuddy, K. (1996). *Explaining Science in the Classroom*. Buckingham: Open University Press.
Osborne, J. F. (1997). Practical Alternatives. *School Science Review, 78*(285), 61-66.
Osborne, J. F. (2001). Promoting Argument in the Science Classroom: A Rhetorical Perspective. *Canadian Journal of Science, Mathematics and Technology Education, 1*(3), 271-290.
Osborne, J. F. (2002). Science without Literacy: a ship without a sail? *Cambridge Journal of Education, 32*(2), 203-215.
Osborne, J. F., & Young, A. R. (1998). The biological effects of ultra-violet radiation: a model for contemporary science education. *Journal of Biological Education, 33*(1), 10-15.
Pontecorvo, C. (1987). Discussing and Reasoning: The Role of Argument in Knowledge Construction. In E. De Corte, H. Lodewïjks, R. Parmentier, & P. Span (Eds.), *Learning and Instruction: European Research in an International Context* (pp. 239-250). Oxford: Pergamon Press.
Russell, T. L. (1983). Analyzing arguments in science classroom discourse: can teachers' questions distort scientific authority? *Journal of Research in Science Teaching, 20*(1), 27-45.
Scheffler, I. (1965). *Conditions of knowledge: an introduction to epistemology and education*. Glenview, IL: Scott, Foresman.
Scott, P. (1998). Teacher Talk and Meaning Making in Science Classrooms: a Vygotskian Analysis and Review. *Studies in Science Education, 32*, 45-80.
Solomon, J. (1991). *Exploring the Nature of Science: Key Stage 3*. Glasgow: Blackie.
Solomon, J., Duveen, J., & Scott, L. (1992). *Exploring the Nature of Science: Key Stage 4*. Hatfield: Association for Science Education.
Suppe, F. (1998). The structure of a scientific paper. *Philosophy of Science, 65*(3), 381-405.
Sutton, C. (1992). *Words, Science and Learning*. Milton Keynes: Open University Press.
Thorley, N. R., & Treagust, D. F. (1986). Conflict within dyadic interactions as a stimulant for conceptual change in physics. *European Journal of Science Education*.
Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.
Walton, D. N. (1996). *Argumentation schemes for presumptive reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.
Welch, W. (Ed.). (1979). *Twenty-five years of science curriculum development.* (Vol. 7). Washington, DC: American Educational Research Association.
White, R., & Gunstone, R. (1992). *Probing Understanding*. London: Falmer Press.
Wray, D., & Lewis, M. (1997). *Extending Literacy: Children reading and writing non-fiction*. London: Routledge.
Zohar, A., & Nemet, F. (2002). Fostering Students' Knowledge and Argumentation Skills Through Dilemmas in Human Genetics. *Journal of Research in Science Teaching, 39*(1), 35-62.
Zoller, U., Ben-Chaim, D., Pentimalli, R., & Borsese, A. (2000). The disposition towards critical thinking of high school and university science students: an inter-intra Israeli-Italian study. *International Journal of Science Education, 22*(6), 571-582.

Zoller, U., Dori, Y. J., & Lubezky, A. (2002). Algorithmic, LOCS and HOCS (chemistry) exam questions: performance and attitudes of college students. *International Journal of Science Education, 24*(2), 185-203.