December 2009

Deliberation versus Dispute: The Impact of Argumentative Discourse Goals on Learning and Reasoning in the Science Classroom

Mark Felton
San Jose State University

Merce Garcia-Mila
University of Barcelona

Sandra Gilabert
University of Barcelona

Recommended Citation
Mark Felton, Merce Garcia-Mila, and Sandra Gilabert. "Deliberation versus Dispute: The Impact of Argumentative Discourse Goals on Learning and Reasoning in the Science Classroom" *Informal Logic* (2009): 417-446. https://doi.org/10.22329/il.v29i4.2907

Deliberation versus Dispute: The Impact of Argumentative Discourse Goals on Learning and Reasoning in the Science Classroom

MARK FELTON

College of Education
San Jose State University
One Washington Square
San Jose, CA 95192-0074
mfelton2@yahoo.com

MERCE GARCIA-MILA

SANDRA GILABERT

Universitat de Barcelona

Abstract: Researchers in science education have converged on the view that argumentation can be an effective intervention for promoting knowledge construction in science classrooms. However, the impact of such interventions may be mediated by individuals' task goals while arguing. In argumentative discourse, one can distinguish two overlapping but distinct kinds of activity: dispute and deliberation. In dispute the goal is to defend a conclusion by undermining alternatives, whereas in deliberation the goal is to arrive at a conclusion by contrasting alternatives. In this study, we examine the impact of these discourse goals on both content learning and argument quality in science.

Résumé: Les chercheurs dans les sciences d'éducation se convergent sur l'idée que l'enseignement de l'argumentation peut être une intervention efficace pour promouvoir la construction des connaissances dans l'enseignement des sciences. Cependant l'impact de telles interventions peut être influencé par les buts d'une tâche d'un individu pendant qu'il argumente. On peut différecier dans un discours d'arguments deux types d'activités qui se chevauchent mais qui se distinguent: le désaccord et la délibération. Le but dans un désaccord est d'arriver à une conclusion en sapant les alternatives, tandis que dans la délibération le but est d'arriver à une conclusion en comparant les alternatives. Dans cette étude nous examinons l'impact de ces buts sur l'apprentissage du contenu et sur la qualité des arguments dans les sciences.

Keywords: Argumentation, argumentative discourse, discourse goals, dialogue, scientific reasoning, science learning, deliberation

1. Introduction

Over the past decade, science educators have converged on the view that argumentation is essential to scientific thinking and knowledge construction and that it should be a central learning outcome in the science curriculum (Jimenez-Aleixandre & Erduran, 2008; Kelly & Chen, 1999; Newton, Driver & Osborne, 1999). This view rests on the idea that students forge a more robust and integrated conceptual understanding of science when they engage in the process of marshalling evidence in support of claims (Leitao, 2000); and that students improve their skills of scientific thinking by considering the merits of their claims and evidence within a framework of alternatives (Kuhn, 1993). Argumentative dialogue, by extension, provides an ideal context for learning in this regard because it establishes a dialectical exchange where peers prompt one another to produce claims and evidence and challenge one another with alternative perspectives.

However, argumentative dialogue also has the potential to interfere with scientific knowledge building and reasoning. Social psychologists have long held the view that arguing can have a polarizing effect on opinions. In a landmark study, Lord, Ross and Lepper (1979) found that participants who were exposed to mixed evidence on a controversial topic developed a bias in favor of evidence that supported their initial views and became more resistant to opposing arguments. In other words, exposure to alternative arguments polarized individuals' opinions by provoking confirmation bias in their thinking. If this is the case, then argumentative dialogue might actually undermine the goals of fostering knowledge construction and scientific reasoning in the classroom by making students resistant to examining and potentially revising their initial beliefs. Subsequent studies of polarization suggest that this effect may be limited, in most cases, to individuals' perceptions of their opinion, rather than their actual position as measured on a pretest-posttest opinion scale (Kuhn & Lao, 1996). Nonetheless, even changes in perceptions may have an impact on how individuals process opposing side claims and evidence. Because they believe that they have become more certain of their opinion, they may be less likely to examine and revise their beliefs.

To date, there have been few studies exploring the conditions under which argumentative dialogue has a positive effect on learning and why. Understanding when argumen-tation facilitates knowledge construction and alternative-based reasoning, and when it impedes them is critical to optimizing its use in science classrooms. In the present study, we address this question by taking a closer look at how students' discourse goals while arguing may affect how they argue, and ultimately how they build knowledge and arguments on a topic. By exploring the mediating effects of discourse goals on dialogue, we hope to shed light on how best to use argumentative dialogue to support knowledge building and reasoning in science classrooms.

2. The role of argumentative dialogue in knowledge building and reasoning

Social constructivists such as Lemke stress that to understand how scientists elaborate their view of the world, we must understand how they exchange ideas and how they change their opinions in response to evidence. As he puts it, to learn science is not to know what the last generation of scientists thinks of the world, but to find out how each new generation of scientists re-elaborates our view of the world (Lemke, 2002). Science is thus seen as a social construction that results from inquiry processes (planning and performing experiments), as well as from communication and public scrutiny processes among the scientific community that lead to resolution and consensus. Driver, Newton, and Osborne (2000) also emphasize the "socially constructed nature of scientific knowledge, and the consequent need to give a much higher priority to discursive practices in general and to argument in particular" (p. 297). Along this line but from the perspective of developmental psychology, Kuhn establishes the goal of science education as promoting a way of thinking in which inquiry and argument are two central skills (Kuhn, 2005). Scientific discursive practices such as assessing alternatives, weighing evidence, juxtaposing competing claims, and evaluating the potential validity of scientific claims are all essential to constructing scientific arguments, and ultimately to advancing scientific knowledge (Erduran, Simon, & Osborne, 2004; Garcia-Mila and Andersen, 2008; Latour & Woolgar, 1986).

For our present work, we take a similar view on science and argument, and see argumentative dialogue as a central feature to the process of constructing knowledge through a process of dialectical exchange. We define "argumentative dialogue" broadly as verbal discourse in which individuals elaborate, juxtapose and evaluate opposing viewpoints. There are two critical elements to this definition: First, two or more individuals recognize that they hold conflicting views on a topic; and second, they engage in dialogue to resolve that conflict. The first condition provides the impetus for engaging in dialogue and the second condition provides the impetus for elaborating, juxtaposing and evaluating the opposing viewpoints. For reasons that will become clear below, our use of the term "argumentative dialogue" includes, but is not limited to discourse that would typically be called an "argument" in lay terms.

From a social constructivist perspective (Vygotsky, 1978) argumentative dialogue provides an ideal context for knowledge building. When students explore their diverging views on a topic, they engage in a host of activities that socially scaffold knowledge construction by producing questions, statements and objections that prompt each other to clarify claims, provide evidence and rebut counterclaims (Felton & Kuhn, 2001). To respond to these interrogatives, individuals must elaborate their understanding of the content, situating simple declarative statements in a more complex explanatory and evidentiary framework. Felton and Kuhn (2001) found that both adults and adolescents employ argumentative strategies that aim at eliciting and addressing opposing viewpoints in conversational contexts. While the quality and sophistication of these strategies differed by age, both groups engaged in discourse moves that prompted partners to elaborate their arguments for the purposes of juxtaposing views. This process of social facilitation, or co-construction, in elaborating arguments through dialogue has been documented in a variety of contexts in ages ranging from childhood (Orsolini, 1993; Chinn & Anderson, 1998) to adolescence (Bell & Linn, 2000; Pontecorvo & Girardet, 1993; Resnick, Salmon, Zeitz, Wathan & Halowchak, 1993) and into adulthood (Felton & Kuhn, 2001). Over time, these co-constructive exchanges can leave individuals with more complex and robust knowledge structures, which have been vetted through argumentative dialogue (Dole & Sinatra, 1998). Studies have shown that when students argue with peers about the meaning and implication of conflicting data, they can prompt one another to substantiate their beliefs and assumptions with evidence and warrants (Bell & Linn, 2000), reconcile the discrepancies in their collective understanding (Nussbaum & Sinatra, 2003), and fortify their conceptual knowledge (Bell & Linn, 2000; Zohar & Nemet, 2002). Studies like these suggest that argumentation opens the door to conceptual change by making students' beliefs explicit and open to evaluation (Kuhn, 1991), providing an opportunity for individuals to examine their views and process claims more deeply. Stein and Miller (1993) have suggested that this kind of processing, in turn, may provide a more meaningful and elaborated structure in which to hold knowledge for retrieval and reconstruction at a later date. For example, Zohar and Nemet (2002) conducted a study in which ninth graders learning about human genetics were assigned to an argumentation group or a control. Students in the argumentation group were taught about how to argue effectively and were then given the opportunity to argue about ethical issues in human genetics. Control group students were given the same materials and time on task, but did not argue. Not surprisingly, students in the argumentation group developed better quality arguments at the posttest than their peers in the control group. But the argumentation group also performed better than the control on a test of students' content knowledge of genetics. Like Stein and Miller, they proposed that these gains may have been due to the fact that argumentation allowed students to actively construct new mental representations, connections and personal understandings in which to embed knowledge.

In addition to providing a context for knowledge building, argumentative dialogue also engages students in dialectical reasoning, exposing them to alternative perspectives and prompting them to address counter-claims, counter-evidence and objections (Felton & Kuhn, 2001). This social process, in which individuals learn to challenge and respond to alternative perspectives, may scaffold scientific reasoning (Kuhn, 2005). As Newton and colleagues (Newton, Driver, & Osborne, 1999) point out, by "taking part in activities that require them to argue the basis on which knowledge claims are made, students also begin to gain an insight into the epistemological foundations of science itself [p. 556]." When students engage in argumentative dialogue, they have the opportunity to discover that scientific reasoning involves thinking through alternative explanations and solutions to come to a conclusion. Similarly, when students discuss socio-scientific issues like genetic modification or stem cell research, the counter-claims, questions and prompts from partners, provide an opportunity to appreciate that evidence must be used to advocate a position while taking its alternatives into account. Whether discussing scientific principles or the use of science to shape policy, students must learn that knowledge is constructed through a dialectical process and that ideas become robust when they are subjected to scrutiny in a framework of alternatives.

In short, an analytical examination of argumentative dialogue suggests that it may offer two potential benefits to students in the science classroom. First, argumentative dia-logue offers students the opportunity to elaborate arguments as they respond to the questions, counter-claims and prompts of their conversational partners. Second, argumentative dia-logue offers students the chance to appreciate the role of alternative claims and evidence in the process of scientific knowledge construction as they try to reconcile their views with the views of others. But for argumentative dialogue to be effective in these ways, individuals must genuinely seek to respond to questions and challenges from their conversational partners with substantive answers and be open to re-evaluating their claims. Without this attempt to engage in an authentic exchange of views, they would not be exposed to the benefits of being prompted to elaborate their arguments or the benefits of weighing conclusions against their alternatives.

3. Dispute vs. deliberation in argumentation

To understand the conditions under which argumentative dialogue promotes scientific knowledge building and reasoning it is critical to consider people's goals while arguing. In argumentative dialogue, one can distinguish two overlapping but distinct kinds of activity: dispute and deliberation (Kroll, 2005). Both kinds of discourse involve two or more speakers who contrast alternative viewpoints by evaluating claims and the evidence used to support them. But dispute and deli-beration—or what Mercer would call *disputative* and *explor-atory talk*—can be distinguished by their goals (Makau & Marty, 2001; Mercer, 2000). In dispute the goal is to defend a viewpoint and undermine alternatives, whereas in deliberation the goal is to arrive at a viewpoint by comparing and evaluating alternatives. These diverging goals, in turn, create important differences in the social dynamic between conversational partners. In dispute, participants compete with the goal of persuading others to adopt their opinion. In deliberation, participants collaborate with the goal of working towards a consensus view.

These discourse activities, dispute and deliberation, in turn, may impact the ways in which individuals process opposing viewpoints. As Leitao (2000) points out, the process of negotiating viewpoints can prompt an array of responses from an individual. When speakers confront opposing claims and evidence in argumentative dialogue, they have at least four basic responses at their disposal: (1) to dismiss counter-arguments and maintain their position; (2) to agree with counterarguments locally, but deflect their impact by turning to other claims in support of their position; (3) to integrate counterarguments by qualifying or adjusting their position; or, (4) to accept counterarguments and abandon their position. When consensus is the goal of dialogue, individuals allow themselves the full range of these responses. In contrast, when persuasion is the goal of dialogue, individuals must dismiss or deflect counterarguments in order to convince others to adopt their conclusions. Thus, persuasive goals in discourse may limit the value of argumentative dialogue for scientific knowledge building and reasoning by constraining the options that individuals believe they have for responding to alternative viewpoints. As a result, the constraints of discourse goals while arguing may lead individuals to superficially process opposing side claims and evidence (Stein & Miller, 1993), negatively impacting their comprehension and memory and limiting the resources at their disposal for crafting a balanced argument.

4. The present study

For the present study, we set out to examine whether discourse goals mediate the effects of dialogic argument on learning and reasoning about socio-scientific issues. We established three conditions: a control group, a disputative group and a deliberative group. The control group was exposed to materials and asked to take a stand without dialogue, while the dispu-tative and deliberative conditions were given the opportunity to discuss the materials with peers who held opposing view-points. To distinguish the goals in the two dialogue conditions, we gave the disputative group the task of persuading their peers, and we gave the deliberative group the task of reaching consensus with their peers. Several studies have shown that task instructions can have an impact on the quality of written arguments that individuals produce (Ferretti, MacArthur & Dowdy, 2000; Nussbaum & Kardash, 2005). These studies have contrasted the effects of broad goals "to persuade" with specific goals "to produce claims, counterarguments and rebuttals." Both studies found that persuasion goals undermined the quality of argument, particularly in the area of citing and rebutting counterarguments. Nussbaum and Kardash (2005) have proposed that persuasion goals may lead individuals to suppress the use of alternative claims and evidence, that is, counterarguments, because they fear that it will undermine the persuasive strength of their essays.

We hope to extend this work on goals in two directions. First, we use a writing prompt at the pretest and posttest—across conditions—that asks students to produce alternative claims and evidence (similar to the specific goal conditions used by Ferretti and colleagues and by Nussbaum and Kardash). We then introduce discussion prompts in an intervention that either direct individuals to persuade (dispu-tative condition) or reach consensus (deliberative condition). Our thinking is that when dialogue instructions conflict with writing prompts, as in our dispute condition, persuasion goals in dialogue will trump the prompt to produce counter-arguments and rebuttals in writing. We measure the impact of this intervention in an analysis of pretest to posttest change. Second, we measure not only the quality of arguments at pretest and posttest, but also understanding of the information that could be used as claims and evidence in the essays. Our thinking here is that persuasion goals may not only interfere with what individuals choose to cite in their essays, but also what they choose to process for understanding. We measure students' understanding and memory for relevant information at the posttest to explore whether students have chosen not to cite opposing claims and evidence for rhetorical reasons, or because they simply haven't processed that information sufficiently to remember it.

In short, we hold that consensus seeking in the context of dialogic argument promotes the goal of deliberation rather than dispute. That is, when students try to reconcile their dif-ferences, they are more likely to examine alternative hypo-theses and weigh them against one another rather than jumping to dismiss alternative claims before carefully weighing their merit. We sought to test whether this goal of weighing alternative viewpoints, in turn, increases the likelihood that students appraise alternative views, process them for under-standing and integrate them into their own arguments. Students who have sought consensus with opposing-side peers may be more likely to have listened to and critically examined the alternatives. We believe that this elaborative processing will increase conceptual understanding and memory for oppos-ing arguments and will ultimately promote a more integrated representation of knowledge. In contrast when individuals seek to persuade, we believe they are more likely to pursue the goals of dispute, choosing to undermine or devalue opposing viewpoints rather than substantively address them, particu-larly when communicating to an audience of "undecided" readers, even when they are prompted to give an unbiased appraisal. As a result, we believe that they will be less likely to understand and remember information that conflicts with their views.

Research questions and hypotheses

Our research questions were as follows. First, do task instruc-tions that ask students to seek consensus and reconcile arguments with a peer (deliberative condition) have a greater impact on science learning and reasoning than task instruc-tions that ask them to persuade a peer (disputative condition) and do both of these conditions outperform a control group? Second, if the task instructions do have a differential impact, what do those differences look like? And third, is there evidence that participants in the deliberative condition spent more time attending to alternative perspectives?

With respect to the first question, we predicted that both dialogue groups would perform better than a control group that did not argue about the topic, and that the students in the deliberative condition would outperform students in the dispu-tative condition on tests of both content learning and argument quality. Second, we anticipated that students in the deliberative condition would be more successful at recalling and recon-structing knowledge from the intervention that conflicted with their own views and that they would produce better quality arguments, using more evidence and more references to alter-native perspectives. Finally, we expected that the key differ-ence between the deliberative and disputative conditions was that students in the deliberative group spent less time disa-greeing and interrupting one another, allowing for more time to elaborate arguments in their dialogues.

Method

Design

A pretest-posttest design was used to compare the effects of different types of dialogue instructions (argue to reach a con-sensus vs. argue to convince the partner) on learning. These two types of instructions were hypothesized to elicit different argumentation processes and thus different degrees of learning. The design involved three conditions: (1) deliberative argumentation, where students were instructed to argue to reach consensus; (2) disputative augmentation where students were instructed to argue to convince the partner; and (3) a control condition where students were asked to read a text on the topic object of debate and answer questions. In the two dialogue conditions (conditions 1 and 2) participants were organized in dyads. A key issue was to match pairs who disagreed on the topic they had to argue about. Therefore prior the each debate phase, all participants were presented a dilemma and were asked to write about their position so they could be matched with a disagreeing partner for the study.

Participants

One-hundred-one 1st year secondary school students (7th graders) attending a public high school in a small town near Tarragona, Spain, participated in the study. The participants' mean age was 12.2 (SD=0.4) (range 12.0-13.0). The develop-mental psychology literature indicates that this is the age at which students begin to spontaneously use arguments, coun-terarguments and rebuttals in their academic discourse, and they become more involved in socio-scientific issues (Felton, 2004; Golder, 1996, see also Kuhn and Franklin's 2006 review). Therefore, early adolescence seemed like an ideal age at which to test our interventions. Students were proportionally pooled from the five classes in grade level and randomly assigned to each of the three conditions. There were 31 students in the disputative condition, 34 in the deliberative condition, and 35 in the control condition. Dyads were formed according to two criteria. First, they had to be paired with a disagreeing partner in each of the three dilemmas they were presented; and second, they had to remain in the same condition throughout the entire intervention. To satisfy these criteria, the students were asked their position in the dilemmas, and were randomly assigned to the two experimental conditions. When these two conditions were full, the remaining students were assigned to the control group, making this a quasi-experimental design. Also we felt that it was essential for the treatment that the students always held genuinely opposing views rather than asking them to play the role of “pro” and “con” in a dilemma. Therefore, dyads were rearranged within group for each dilemma to ensure that students always met with an opposing-side peer. As a result, the unit of analysis when coding dialogue was the indivi-dual rather than the dyad.

Procedure

The intervention comprised 8 fifty-minute sessions on the topic of fuel sources and their role in climate change. The setting was a science class in which the experimenters worked closely with the teacher. In the first two sessions, the intervention was introduced, students took the pretest, and the teacher gave two presentations, one about the Greenhouse effect (session 1) and the other about sources of energy (session 2). In sessions 3, 5 and 7 students were presented with the dilemmas, one for each session. All students read through materials that provided background information, which could be used to create claims and evidence, and they wrote a short essay taking a position with respect to the dilemma (initial position). In sessions 4, 6 and 8 the text with the initial position was returned to all students and the groups for each condition and dyad were formed (15 minutes). Then the students in the two dialogue conditions were asked to argue about the dilemma according to the specific instructions for each condition and at the end (15 minutes), they wrote a text explaining their final position (15 minutes) (see Appendix A for the dialogue prompts used in each condition). To control for time on task, students in the control condition were told to review the text for each dilem-ma, outline the advantages and disadvantages of the options described in the dilemma and write a short essay explaining their final position (15 minutes). The final position text was not analyzed to assess learning; rather it was aimed at helping the students think about their conclusions. Finally, at the end of session 8, all the students took the posttest. All the dialogues in conditions 1 and 2 were audio taped and transcribed for analysis.

Instruments

Dilemmas

All three dilemmas were about fuel sources and climate change. The first was about an Energy Project for the city of Tarragona designed to accommodate that city's increased population and new energy needs. The Project required a choice among differ-ent sources of energy (e.g., nuclear, solar, biodiesel). The second dilemma centered on approving a project that involved developing windmill farms to generate energy. And the third dilemma was about research and development in bio-diesels.

Pretest and posttest

The pretest and posttest were identical. They were structured in two parts. The first part consisted of six open-ended content questions about energy sources, which tested students' understanding of material from the texts they had read and been exposed to in class. This section of the test was scored for number of correct responses according to the material developed in the class (max. score =10). Examples of questions in the first part of pre/post test are: *What types of fossil energy sources do you know? Why do we say they are nonrenewable? What types of nonrenewable energy sources do you know?* (See Appendix B).

The second part, an essay-prompt on the pretest and posttest, asked participants to propose an energy plan that argued in favor of using one or more energy sources. This essay-prompt asked students to identify the major advantages and disadvantages of each energy source, and to make a recommendation. Thus, students were prompted explicitly to consider the pros and cons of alternatives in their essays. The essays were scored using a rubric on argument quality (Appendix C) adapted from Kelly, Regev and Prothero (2008).

Results

The data analysis was performed on two different dependent variables: Content learning and argumentation. Content learning was measured from answers to questions 1 to 6 in Section I of the pre- and posttests. Argumentation was measured using participants' answers to the essay prompt in Section II at the pretest and posttest. Both statistical tests consisted of repeated measures with one within-subjects' variable and one between-subjects variable. The within-subjects variable consisted of the scores in the pre- and posttests while the between subjects factor was condition with three levels: disputative argumentation, deliberative argumentation, and the control condition.

Content learning

All students' pre and posttests were double-coded and reliabil-ity reached 98% agreement. Discrepancies in scoring were resolved by discussion. The scoring strategy is presented in Appendix B along with the content test. The means (and standard deviations) for the pretests and posttests appear in Table 1. The repeated measures analysis for the first dependent variable (content learning scores on questions 1, 2 and 3) yielded a significant interaction effect between pre-posttest and condition [(F (2,97)= 5.8, p=.004), effect size, η^2 =.107]

Given that the variances were not homogenous, and in order to control for experiment-wise error, T3-Dunnet con-trasts were performed. No differences between the means of three conditions in the pretest were found while the contrasts between the means in the post test yielded significant differ-ences between the deliberative group and the control group (p=.001). No significant differences were found between the deliberative and the disputative nor between the control and the disputative ones.

Argument quality

The student's texts for argumentation in the pre and posttest were coded according to an argumentation rubric (see Appendix C). This rubric consisted of 8 items with a dichotomous coding criterion (*yes* or *no*) for each item. For each criterion satisfied, the text was scored one point, and the final score was obtained by adding the scores of each item with a maximum score of 8. Eventually, this score was transformed into a score of 10[1]. Interrater reliability was calculated and reached 93% of agreement. Again disagreements were resol-ved by discussion. The dependent variable was obtained by adding all "yes" responses in the rubrics. The means (and standard deviations) for the pretests and the posttests are presented in Table 1. The repeated measures ANOVA also yielded a significant interaction effect between pre-posttest and condition [(F (2,97)= 7.37, p=.001), effect size η^2 =.132]. However, since the

---

[1] This transformation was done in order to equate the two variables—Content learning and Argumentation—on the maximum score. Therefore, the means for Argumentation were multiplied by 10 and divided by 8 (their former maximum score).

second dependent variable (argument quality) was not normally distributed, we used a nonparametric test to analyze the differences between means. None of the U Mann-Whitney comparisons between the means of the pretest were significant. In contrast, when the posttests were compared, the U Mann-Whitney yielded signifi-cant differences between the mean of the deliberative group vs. the control group [(*U (2)*=250, *p*=.001), effect size, *r*=.50], and between the two types of discussion groups, the deliberative vs. the disputative [(*U (2)*=343, *p*=.014), effect size, *r*=.30], while the differences between the disputative and the control groups were not significant (*U (2)*=423, *p*=.121).

Table 1. Means (and Standard Deviations) comparing the Scores in the Pretest and the Posttest, and the scores in the Two Types of Questions across Conditions

	Conditions	Pretest	Posttest
Content	Disputative (N=31)	1.5 (1.5)	5.3 (3.2)
	Deliberative (N=34)	1.66 (1.8)	6.7 (2.1)
	Individual (N=35)	0.96 (1.4)	3.7 (3.2)
Argument-ation	Disputative (N=31)	0.1 (0.2)	2.9 (2.5)
	Deliberative (N=34)	0.6 (0.2)	4.5 (2.1)
	Individual (N=35)	0.2 (0.1)	1.9 (2.2)

Categorical Analysis of Argument Quality.

In order to find out the cause of the significant differences between the means in the prior ANOVA, a categorical analysis was performed comparing the frequencies of students whose text satisfied each category of the rubrics used to code the answer to Section II in the pre/post tests. As we have mentioned, the rubrics consisted of 8 items with a dichotomous coding criterion (*yes* or *no*) for each item. Frequencies of students in each condition whose text satisfied each criterion were computed and an analysis of frequency distributions was performed. When the three conditions were pooled in the analysis, none of the items in the pretest chi-square compar-isons yielded significant differences between conditions while all of the chi-square comparisons across conditions for items in the posttest were significant. Since the frequencies in the control condition were much lower than in the other groups, a fact that could mask possible differences between the two dialogue groups: the deliberative and the disputative, the same analyses were performed including only these two conditions (disputative vs. deliberative). The chi-square analysis to test differences in the distribution of *yes* responses between the two dialogue groups yielded differences in items 1, 2, 3, and 7. The statistical tests were for item 1, *c2* (1)= 5.14 p=.023; for Item 2, *c2* (1)= 5.7 p=.01; for Item 3, *c2 (*1)= 4.4 p=.036; and finally for Item 7, *c2* (1)= 6.7 p=.009. For the rest of the items the analyses to compare the distribution of frequencies were not significant. Items 4 and 6 yielded equally distributed fre-quencies across the two groups and across the yes/no answers, while for items 5 and 8, the rubrics criterion was satisfied by very few students in each group, showing a ceiling effect (see Table 2 for the frequency distribution). The resulting signifi-cant differences between the deliberative and the disputative conditions show that the discussion to reach consensus help them build arguments for their views more clearly and provide a justification for them (items 1 and 2). Also, and very import-ant, another benefit from discussing to reach consensus was to become more open to take the limitations of one's own view into consideration (item 3). Also, as shown by the significant differences in item 7, there were more students in the deliberative condition compared to the disputative condition that were able to justify their position. The items for which no differences were found were also interesting. Items 5 and 8 showed very low frequencies, indicating a bottom effect for all groups. Item 5 would represent the presence of counterclaims in the texts and the low frequencies in item 8 shows the lack of extra information beyond the one provided in class. On items 4 and 6 both dialogue groups showed positive change on at the posttest yielding no significant differences. On item 4 both groups showed increases in their ability to cite the deficits of fuel sources they rejected. On item 6 both groups showed increases in the consistency of their position over the course of the entire essay. We believe that the nature of argumentative dialogue in both conditions were likely to produce these changes since even in the disputative condition, students would have been likely to clarify their position and indentify flaws in alternative positions over the course of the intervention.

Table 2. Distribution of Frequencies of *yes-answers* in each Item in the Rubrics Used to Code the Quality of the Argument in the Pre/Post Tests (Question 7) across Conditions

	Pretest			Posttest		
	Disput-ative	Deliber-ative	Indivi-dual	Dispu-tative	Deliber-ative	Indivi-dual
	n=31	n=34	n=35	n=31	n= 34	n= 35
Item 1	1	5	3	20	30	18
Item 2	1	5	2	17	28	16
Item 3	0	3	1	12	22	10
Item 4	0	2	0	12	18	4
Item 5	0	0	0	2	5	1
Item 6	0	0	0	10	17	7
Item 7	1	4	1	14	26	12
Item 8	0	0	0	2	6	0

Dialogues

In order to interpret differences found between the two dia-logue conditions, we looked at two parameters in the students' dialogues. The first parameter was the number of words per utterance, a measure of the length of each conversational turn. Our assumption was that students in the deliberative condition would have longer utterances than their peers in the disputative condition because they were trying to use dialogue to explore each others' position to find consensus. Conversely, we assumed that students in the disputative condition would have shorter utterances since they were competing to establish the persuasive force of their own position in the dialogue. All one-hundred-eleven dialogues were transcribed and the mean number of words per utterance was calculated. The mean number (and *SD)* of words per utterance in the deliberative condition was 23.1 (*SD*=12.1) whereas for the disputative condition it was 18.2 (*SD*=7.9). A *t*-test was performed and it yielded marginally significant differences, t (57.2)=-1.9, p=.06 (effect size: η^2 = .053).

Another assumption was that the students in the disputative condition would produce more simple negatives and affirmatives than those in the deliberative condition. A simple negative is defined as an utterance in which the speaker asserts that the partner's claim is wrong or incorrect without further elaboration to establish why (e.g., "You're wrong!" or, "That's not true!"). A simple affirmative is defined as an utterance in which the speaker asserts that his or her own claim is correct or true without further justification (e.g., "I'm right!" or, "Yes it is!"). Simple negatives and affirmatives were coded and tallied for the entire set of dialogues by one coder. To calculate reliability, 40% of the total dialogues were coded by a second coder. Interrater reliability was 90% agreement and all discrepancies were resolved by discussion. In the following piece of dialogue taken from the disputative condition, Alvaro and John[2] argue about the first dilemma, illustrating the use of simple affirmatives and simple negatives (indicated as SA and SN in parenthesis) and the shorter length of utterances:

Alvaro: I keep saying that I do not want a Nuclear Power station.
Javier: So… I say yes! If a nuclear station explodes it would not be that bad!
Alvaro: They are very harmful, and they should not build any mone, there are too many already.
Javier: I'm right and that's all (SA)
Alvaro: No way! (SN)
Javier: Is it clear? Building a nuclear power station is better! (SA)
Alvaro: No way, it is better to build a thermal station (SN)
Javier: Why?
Alvaro: Because I say it. (SA)
Javier: So what if you say it? The nuclear power station is better (SA)
Alvaro: No! I'll say it again. See the map? Here is Spain and this a thermal station. If it explodes, it would not reach us, but if it was a nuclear power station we would have radiactivity for years and years.

The number of simple negatives and affirmative were then added together for each dialogue to calculate the mean and standard deviation for each group. The mean for the deliberative group was 2.11 *(SD*=1.7) whereas the mean for the disputative group was 4.45 (*SD*=3.9). Since the variable was not normally distributed, a nonparametric test was performed, which yielded significant differences between groups [U =282.0, p=.001, Ranges for the disputative group were 40.9 and for the deliberative, 25.79, and the effect size was r=.40].

[2] All names are pseudonyms. All dialogues were held in Spanish and have been translated into English

5. Discussion

The impact of discourse goals on content learning and argument quality

The results of this study support our first two hypotheses, specifically, that although argumentative dialogue can improve content learning and argument quality on socio-scientific issues, the benefits are mediated by individuals' task goals while arguing. The pretest to posttest change in content learning suggested a trend in which the deliberative group outperformed the control and the disputative fell in between. Significant differences found between the deliberative and control groups suggest that deliberation is an effective inter-vention for promoting content learning. Students in the deliberative condition, instructed to reach consensus and reconcile differences, were more likely than their peers in the control group to understand and retain information that could be used as evidence and construct arguments that acknow-ledged the limitations of their proposals. These results held while controlling for both exposure to content and time on task across conditions. In many ways, our findings are in line with previous findings (Zohar & Nemet, 2002). Argumentation of any sort, whether disputative or deliberative, seems to have a positive impact on reasoning skills since they prompt students to engage in deep processing of information. However, when students engage in consensus building towards finding a position, rather than competitively defending a position, they are more likely to understand and recall information that can serve as evidence both in favor of and against their own position, and they construct arguments that show greater attention to claims and evidence on both sides of the issue. This outcome is observed in the argumentative text at the posttest and generated the significant differences found in the chi-square analyses.

The following excerpt from a dyad in the deliberative condition clearly illustrates the kind of collaborative dialogue described above. Discussing the first dilemma, Aaron defends the project that uses fossil energy produced in thermal stations while Paul claims that nuclear energy is the best solution. They keep arguing that radioactive leaks are dangerous on the one hand, and on the other, that burning fuel generates CO2, causing the Greenhouse effect. They weigh arguments on both sides of the issue, acknowledge counterarguments, and work towards crafting a solution that addresses the counter-arguments to their satisfaction.

Aaron- OK. It is clear that the two types of thermal stations are negative, and if we don't agree on which is best we will have to find a solution. I think that an alternative solution is the use *of a windmill farm. I think it is a good choice because the generators work based on wind energy. It is true that we need a lot of them but the energy is renewable it does not produce CO2 and it does not generate waste, it is easy to obtain and it contributes to technological development.*

Paul: OK, but I think that when there is no wind, the windmills consume energy and this is a problem because it takes energy to produce energy. Also, the windmill farms need land, and this land must be prepared, trees must be cut and this means the destruction of ecosystems in the area.

Aaron: So then why not put the windmill farms in an area where there are no trees or woods? This way the ecosystem would not be destroyed. There are many windy areas with no trees (…) Ok, I think the windmill energy is a little better than the other two sources of energy [nuclear and thermal.]

Paul: Fossil fuels, like gas […] are very important, they produce 79.6% of the energy, but at the same time they produce a lot of CO2 that increases the Greenhouse effect. This is very harmful for our health and for the atmosphere. The windmill farms could be a solution.

Aaron: Of course, just like thermal stations are harmful, nuclear stations can produce radioactive leaks. So, I think that the windmill farms would be an alternative to these two sources of energy [nuclear and thermal].

In contrast to the previous dialogue from the deliberative condition, we can see in the following excerpt from the disputative condition, that the two students ignore each other's counterarguments and ultimately show no signs of progress. Xavier defends fossil fuels while Norbert defends nuclear energy.

Norbert: Mine [the nuclear station] only has leaks [as a minor problem] and that's it. Why do you think it is better to go with fossil fuels?

Xavier: Because people prefer fossil combustion, that is, the one I chose.

Norbert: Why?

Xavier: Because they don't have leaks.

Norbert:But the CO2 will increase in the atmosphere.

Xavier: Yes, but there are also other things that increase CO2.

Norbert: OK, but the options that involve burning fuel are non-renewable, and we will run out of them someday. The stations based on burning fuel cause acid rain, they destroy ecosystems and can be the cause of serious health problems.

Xavier: But they are not as bad as radioactive leaks.

Norbert:Yes, because CO2 levels increase.

Xavier: OK, but everybody [causes] increases [in] CO2.
Norbert: Yes, but the acid rain and the health problems are very serious, and the ecosystems…
Xavier: And the leaks?
Norbert: And the acid rain?
Xavier: OK, my claim is the thermal station. I will not change my opinion
Norbert: So, I defend the nuclear station. I will not change my opinion either.

In short, both forms of discourse prompted students to retain information and develop more robust arguments for their views. Not surprisingly, they were also more likely than students in the control condition to produce arguments at the posttest that cited evidence in support of claims. However, students in the deliberative condition were more likely to retain information and craft arguments that acknowledged opposing viewpoints. They were also more likely to acknow-ledge the limitations of their own conclusions, suggesting that they were open to revising or refining their plans even after their dialogues. Finally, and this came as a surprise, students in the deliberative condition were also more likely to cite evidence for claims on their own side than their peers in the disputative condition, suggesting that the process of collabora-tively constructing arguments also may have helped them appreciate the need to substantiate their own opinions. Taken together these results suggest that consensus seeking may reduce the polarizing effects of argumentative dialogue and help students avoid the effects of confirmation bias, prompting them to process a wide array of arguments and evidence on either side of an issue.

To those of us working in the field of argumentation research, these results may seem somewhat trivial. Few researchers would argue that disputative argument is the sort of discourse they would encourage in science classrooms. However, our findings tease out what we see as a critical insight in organizing learning experiences around argument-ation. Students' discourse goals while arguing impact their learning and we cannot assume that they understand what those goals should be. Young adolescents have little or no exposure to argumentative discourse in their science class-rooms (Driver, Newton, & Osborne, 1998). And in most cases, the majority of their experience arguing comes from the context of interpersonal conflict (Stein & Miller, 1993) where the discourse goals are almost invariably arranged around winning a dispute. Indeed, dispute is such a powerful and pervasive model for argumentation in our daily discourse, that students need explicit directions when we want them to adopt a different goal. Science educators interested in using argu-mentative dialogue as a classroom intervention should be careful in crafting task instructions so that students are promp-ted to engage in more collaborative, deliberative forms of argumentative discourse.

The results also confirm our third hypothesis—that task instructions can prompt individuals to engage in distinctly different kinds of discourse, which in turn mediates the effects of argumentative discourse on learning. When they were told explicitly to reach consensus, the students in the deliberative condition spent less time disagreeing and more time elabor-ating opposing viewpoints than their peers in the disputative group. These results are promising given what we know about preponderance of experience adolescents have with conflictual discourse and the paucity of experience they have with more collaborative forms of arguing. This finding has important implications for future research in using argumentative dis-course as a tool for science learning. Deliberative discourse requires a host of skills in elaborating, juxtaposing and recon-ciling arguments, which might be explicitly fostered in the science classroom. It is possible that an intervention designed specifically to teach the skills of deliberative discourse would have had an even greater impact on learning and reasoning than we found in this study. Based on the results of the present study, we believe that this hypothesis merits further investigation.

Limitations and implications for further research

There are some limitations to this study worth noting. First, it may have been helpful to look at a control group that had the experience of non-argumentative collaborative dialogue as part of their intervention. Such a control would have allowed us to better isolate the effects of argumentation over and above the motivational effects that may accompany working with peers. For example, it could be that simply collaborating with a peer on processing the content had a positive effect on student learning, particularly on the content learning questions in the pre- and posttests. Similarly, peers in the two experimental conditions may have simply helped each other comprehend the readings through the course of their dialogues. We do not believe that the confounding of argumentative dialogue with collaboration, however, diminishes the importance of our findings. To begin, the effects of working with a peer have already been controlled for in other studies looking at the positive effects of argumentative dialogue (Kuhn, Shaw & Felton, 1996; Resnitzkaya, Anderson, McNurlen, Nguyen-Jahiel, Archidou & Kim, 2001). Also, the primary goal in this study was to investigate the differences between the deliberative and disputative task goals in learning. Since collaboration was held constant in those two groups, the confound in our control group does not limit our ability to address our central hypotheses in the study. Nonetheless, a more carefully con-structed control group would give an even better picture of the relative advantages of deliberation and dispute in science learning and warrants further study.

Another related concern involves the question of transfer appropriate processing (Morris, Bransford & Franks, 1977). Perhaps some groups in the study were more inclined than others to processes the strengths and weaknesses of each proposal due to the nature of the task they were instructed to engage in. In fact, one group did receive precisely these instructions—the control group. But as our results reveal, this group showed the lowest performance in both the test of content and the argumentation essay. For the reasons cited above, we believe that this effect can be attributed, at least in part, to the positive effects that argumentation has on building meaningful and elaborated knowledge structures. Transfer appropriate processing may also account for the differences found between the two dialogue groups. Both dialogue groups were given prompts that would expose students to the strengths and weaknesses of alternative viewpoints. However, students in the disputative group may have been less inclined to engage in the kind of processing that would support retention and reconstruction of the strengths and weaknesses of each position. But in a sense, this explanation of the findings is precisely our point—when students adopt persuasion goals, they are less likely to process and retain information that might conflict with their views even though they have exposure to those views. Conversely, students who hold opposing views, when asked to engage in consensus seeking may be more likely to process and retain information about both sides in the course of deliberating.

Third, although participants in the deliberative condition were more likely than their peers to acknowledge the limitations of their proposals in their posttest essays, they generally did not acknowledge the benefits of the discarded alternatives. Remember that the writing prompt asked students to explain the advantages and disadvantages of the alternatives. If deliberative dialogue really does involve a careful weighing of arguments and evidence on both sides of an issue, then we should expect that individuals should be able to cite the benefits of proposals they did not choose and speak to why the benefits of their proposal outweighed those of the alternatives. Without additional data on the planning process, the reasons for this shortcoming are a matter of conjecture. One possibility is that even in the deliberative condition, students struggled with the idea that their essays should address the strengths of alternatives. Another possibility is that they lacked a schema for structuring an essay that presents alternatives in their best light and explains why they were not chosen. Both of these possibilities are strengthened by our finding that students in the deliberative condition were generally able to cite the advantages and disadvantages of options in our test of content learning at the pretest and posttest. Again, further study would be needed to investigate this question.

Finally, for the present study we chose to examine between group differences in the context of socio-scientific controversies. This domain by no means represents the breadth of knowledge in science and limits the generalizability of our findings. It may be that when students discuss lab results or argue about forces in physics, consensus seeking may steer students away from the kinds of disagreement that foster cognitive conflict. However, we are inclined to believe that even in these contexts, tasks that prompt students with opposing views to suspend—rather than defend—judgment are most beneficial, and that seeking consensus fosters the kind of open dialogue that supports a careful consideration of alternative views, along the lines of Mercer's exploratory talk. Of course, demonstrating the benefits of consensus seeking in other science disciplines requires further study. Similarly, there are many species of argument types other than policy claims, which were not represented in this study. Our findings open the door to testing the suitability of consensus seeking as a means to foster deliberative argument for other domains of knowledge and for other argument types.

Despite these limitations, we believe that the current study offers important insights for researchers and practitioners who are interested in using dialogic argument in the science classroom. Educators must take care when using argument-ation to teach science. Although dispute-based activities like debates and "letters to the editor" are engaging for students and provide some benefits to learning, if they are not properly monitored, they may reinforce the mistaken assumption that the aim of such activities is to defend a view without carefully considering alternatives. We hope that the current study demonstrates that it is not the only the activity, but also the student's goals when engaging in the activity that matters. Students must understand that science advances through a careful process of testing claims against their alternatives. This requires an open and receptive approach to opposing viewpoints. To pursue this goal, teachers must develop learning activities that provide opportunities for students to disagree, while providing a process for suspending judgment and weighing alternatives. As the results of this study demonstrate, disagreement can be a powerful tool for teaching students about claims and evidence, but it must be situated in a context that fosters genuine attempts to understand opposing viewpoints. As Mercer (2000) recommends in his concept of exploratory talk, "partners engage critically, but constructively with each other's ideas…proposals may be challenged and counter-challenged, but if so reasons are given and solutions are offered (p. 153)." Striking this careful balance between contrasting competing claims and working collaboratively towards a solution is no mean task, but the science classroom is precisely the context in which to scaffold this sophisticated, and ultimately enriching form of disciplinary thinking.

Acknowledgments

The present work was made possible thanks to funding to the second author by the Ministerio de Educación y Ciencia, España (Projects SEJ2006-15639-C02-02 and PR2008-0155), and to the third author by the Generalitat de Catalunya.

References

Bell, P. & Linn, M.C. (2000). Scientific arguments as learning artifacts: designing for learning from the web with KIE. *International Journal of Science Education, 22*(8), 797-817.

Chinn, C.A., & Anderson, R.C. (1998). The structure of discussions that promote reasoning. *Teachers College Record*, 100, 315–368.

Dole, J.A. & Sinatra, G.M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, 33(2/3), 109-128.

Driver, R., Newton, P., & Osborne, J. (1998). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 287-312.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88*, 915-933.

Felton, M. & Kuhn, D. (2001). The development of argumentive discourse skills. *Discourse Processes, 29*(2&3), 135-153.

Felton, M. (2004). The development of discourse strategy in adolescent argumentation. *Cognitive Development, 19*(1), 39-58.

Ferretti, R.P., MacArthur, C.A., & Dowdy, N.S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92,* 694–702.

Garcia-Mila, M., & Andersen, C. (2008). Cognitive foundations of learning argumentation. In S. Erduran & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education. Recent developments and future directions* (*pp.* 29-47). Dordrecht, Netherlands: Springer.

Golder, C. (1992). Production of elaborated argumentative discourse: The role of cooperativeness. *European Journal of Psychology of Education, 1*, 51–59.

Golder, C. (1996). *Le développement des discours argumentatifs*. Lausanne (Switzerland): Delachaux et Niestlé.

Jimenez-Aleixandre, M., & Erduran, S. (2008). Argumentation in science education: An overview. In S. Erduran & M. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3-27). Dordrecht, Netherlands: Springer.

Kelly, J.G., & Chen, C. (1999). The Sound of Music: Constructing Science as Sociocultural Practices though Oral and Written Discourse. *Journal of Research Science Teaching, 36*(8), 883-915.

Kelly, G.J., Regev, J., & Prothero, W. (2008). Analysis of lines of reasoning in written argumentation. In S. Erduran & M. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 137-157). Dordrecht, The Netherlands: Springer

Kroll, B.M. (2005). Arguing differently. *Pedagogy, 5*(1), 37-60.

Kuhn, D., & Franklin, S. (2006). The second decade: What develops (and how)? In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology* (6th ed., Vol. 2). Hoboken, NJ: Wiley.

Kuhn, D. & Lao, J. (1996) Effect of evidence on attitudes: Is polarization the norm? *Psychological Science, 7*(2), 115-120.

Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press.

Kuhn, D. (1993). Connecting Scientific and Informal Reasoning. *Merrill-Palmer Quarterly, 39* (1), 74-103.

Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction, 15* (3), 287-315.

Latour, B., & Woolgar, S. (1986). An anthropologist visits the laboratory. In B. Latour & S. Woolgar (Eds.), *Laboratory life: The construction of scientific facts* (pp. 83-90). Princeton, NJ: Princeton University Press.

Leitao, S. (2000). The potential of argument in knowledge building. *Human Development (43),* 332-360.

Lemke, J.L. (2002). Enseñar todos los lenguajes de la ciencia: palabras, símbolos, imágenes y acciones. In M. Benlloch (Ed.), *La educación en ciencias: Ideas para mejorar su práctica.* (pp. 159-185). Barcelona: Paidós.

Lord, C.G., Ross, L., & Lepper, M.R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Social and Personality Psychology, 37*(11), 2098-2109.

Makau, J.M. & Marty, D.L. (2001). *Cooperative Argumentation: A Model for Deliberative Community*. Prospect Heights, IL: Waveland Press, Inc.

Mercer, N. (2000). *Words and minds: How we use language to think together.* London: Routledge.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533.

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

Nussbaum, E.M. & Sinatra, D. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology, 28*, 384-395.

Nussbaum, E.M. & Kardash, C.M. (2005). The effects of goal instructions and text on the generation of counter-arguments during writing, *Journal of Educational Psycho-logy,* 97(2), 157-169.

Orsolini, M. (1993). "Dwarfs don't shoot": An analysis of children's justifications. *Cognition and Instruction, 11,* (3 & 4), 281-297.

Pontecorvo, C., & Girardet, H. (1993). Arguing and reasoning in understanding historical topics. *Cognition and Instruction, 11,* 365-395.

Resnick, L.B., Salmon, M., Zeitz, C.M., Wathen, S.H., & Holowchak M. (1993). Reasoning in conversation. *Cognition and Instruction, 11,* 347-364.

Reznitskaya, A., Anderson, R., McNurlen, B., Nguyen-Jahiel, K., Archidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2&3), 155-176.

Stein, N.L., & Miller, C.A. (1993). The development of memory and reasoning skill in argumentative contexts: Evaluating,

explaining, and generating evidence. In R. Glaser (Ed.), *Advances in Instructional Psychology*. Hillsdale, NJ: Erlbaum.
Vygotsky, L.S. (1978) *Mind and society*. Cambridge, MA: Harvard University Press.
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.

Appendix A. Prompts for Each Condition of the Treatment

Prompt for the Disputative Condition

Your task is to discuss with your partner for 15 min. about the topic of the dilemma just presented to you. First you have to explain your partner your position in the dilemma and why. The goal of the task is to convince your partner of the choice you have made in the dilemma by means of a good justification. To accomplish the goal you should use any relevant knowledge that you think can help you convince your partner. To convince him or her you must identify his/her choice weak points and rebut them with possible counterarguments. Remember that you must maintain your position until the end and you must keep trying to convince your partner with your arguments. At the end you must have the perception that you have accomplished the goal of convincing your partner.

Prompt for the Deliberative Condition

Your task is to discuss with your partner for 15 min. about the topic of the dilemma just presented to you. First you have to explain your partner what is your position in the dilemma and why. The goal of the task is to reach an agreement with your partner and propose a consensus solution to the problem. To accomplish the goal you must identify the differences in your positions in the dilemma and the reasons that led you choose one position or another. You must analyze the different justifications that led to one or another position. Then you should try to get those positions closer until reaching a consensus. Remember that your goal is not to convince your partner of your position but rather to reach a consensus about the problem and propose a solution.

Prompt for the Individual Condition
Your task is to complete the demand as presented. You have 15 min. to complete it (see sample task for Dilemma 1).

Task: Although you have already made your choice in the dilemma, explain the advantages and disadvantages of the option(s) presented. *Advantages for choice A: The settlement of the Thermal station based on cool combustion* *Disadvantages for choice A: The settlement of the Thermal station based on cool combustion* *Advantages for choice B: The settlement of the Nuclear power station* *Disadvantages for choice B: The settlement of the Nuclear power station*

Appendix B. Pre/Post Test and Coding scheme for Section I

SECTION I: Content Learning

1. Cite the sources of fossil energy that you know?
(Score: 3 points: 1 points for citing cool, 1 point for citing gas and 1 point for citing oil)
2. Why do we say that they are sources of energy non renewable?
(Score: 1 point if the correct definition is provided)
3. What are the main sources of renewable energy?
(Score: 3.2 points: 0.8 points for citing each of the following: wind energy, sun power, hydraulic/wave energy, biomass/bio-fuel)
4. Why do we say they are renewable?
(Score: 0.8 points if the correct definition is provided)
5.What two types of thermal power stations that work with nonrenewable energy do you know?
(Score: 1 points: 0.5 point for citing nuclear power and 0.5 point for citing thermal energy)
6. What are the advantages and disadvantages of each one?
(Score 1: 0.5 for citing 7 advantages and 0.5 for citing 7 disadvantages: 0.07 points for each advantage and 0.07 for each disadvantage)

SECTION II. Argumentation

The Local Government is planning the energy sources of our region for the next decade. As an engineer of great prestige you have been asked to be part of the advisory team. Write a report of your energy project in the space below according to the following guidelines, first cite the main sources of energy that you know and explain the advantages and disadvantages of each. Second, make a proposal of one or more sources of energy explaining the reasons that make you think they are the most appropriate. You must think of environmental, economic, social and geographic factors. Try to always justify the reasons that make you propose some sources of energy and discard others. You have to be convincing with your report, and well organized and do not forget to use scientific reasons since the best-justified proposal will be chosen to be part of the Government Energy Report.

Appendix C. Rubrics for Coding Question 4 in the Pre/Post and Examples

Questions about the students' arguments	Examples
Item 1. Is there a clear proposal of the forms of obtaining energy? No (0 points): The participant does not make any claim about a convenient the type of energy. Yes (1 point): There is a claim about a type of energy that is convenient.	*i.e. Nuclear power stations are very energetic, they do not contaminate by generating CO.2* (1)
Item 2. Is the proposal justified by explaining the advantages of the choice? No (0 points): The participant justifies the choice of a given energy explaining its advantages. Yes (1 point): The participant does not justify the choice of a given energy explaining the advantages of the choice.	*i.e. I propose the windmill farms because they do not contaminate and wind energy is renewable energy* (1) *i.e. I think sun energy is the best!* (0)

Item 3. Although there is clear proposal of the forms of getting energy, is the student aware of the limitations of the proposal? No (0 points): The participant makes a clear proposal of a type of energy and also provides information of its limitations. Yes (1 point): The participant makes a clear proposal of a type of energy but does not provide any limitations.	i.e. *I recommend sun energy because it is clean and it does not contaminate, and is unlimited* (0) i.e. *Nuclear power stations because they are very energetic, and do not contaminate with CO2, but they are dangerous. They have the risk of leaks* (1)

Item 4. Are the discarded forms of getting energy justified by explaining their limitations? No (0 points): The participant discard a given type of energy without providing any justi-fication with its limitations. Yes (1 point): The participant discards a given type of energy by explaining its dis-advantages.	i.e. *I would never recommend nuclear energy* (0) i.e. *I discard the biodiesel because it would make the poor countries even poorer* (1)
Item 5. Although there is clear proposal of the forms of getting energy rejected, is the student aware of its advantages? No (0 points): Although the participant may make a claim against a given type of energy, he or she may not be aware of possible advantages Yes (1 point): The participant makes a claim against a given type of energy, and the text provides potential advantages.	i.e. *I'd never propose the wind energy because it destroys the ecosystems, birds get trapped in the helix and die and gener-ators make a lot of noise* (0) i.e.*I'd never suggest the nuclear energy, because although it does not contaminate with CO2, it could have leaks and this would kill the population* (1).

Item 6. Is the thesis about the forms of getting energy proposed and discarded kept until the end of the text in a coherent manner? No (0 points): The proposal is not maintained throughout the text coherently Yes (1 point): The proposal is maintained (or changed) throughout the text coherently	i.e. *I propose biodiesel energy because it does not cause acid rain and it is unlimited (…) To finish with, I would make propose wind energy* (0) i.e. *I think the best is sun energy because it does not contaminate and you can sell the extra energy produced. It does not destroy the environment (…).As I mentioned before, the best is the sun energy because in addition to al the advantages, it is unlimited* (1).
Item 7. Is the proposal justified by relevant information? No (0 points): The participant does not justify his/her own proposal with relevant information beyond that advantages and disadvantages Yes (1 point): The participant justifies his/her own proposal with relevant information beyond that advantages and disadvantages.	i.e. *I'd choose nuclear energy because in 2 years we could be the most developed country* (0) i.e. I*'d choose wind energy because it is renewable, it does not contaminate cause acid rain and it does not cause climate change. It does not cause the greenhouse effect, and it generates a lot of electric energy* (1)
Item 8. Does the student appeal to information other than the one provided by the teacher during the instructional session? No (0 points): The participant does not use any other information a part from the content provide in the class Yes (1 point): The participant uses any information a part from the content provide in the class	i.e. *Thermal power stations have the following advantages* (repeats what has been said in class) i.e. *I'd never propose fuel.It hurts to see how the big oil companies lie. I once read that in Mexico the Government funded 20 taxis that worked with hydrogen that didn't contaminate, and the oil companies covered* *it because they did not want to loose money.*