British Journal of Educational Psychology (2015)

The British Psychological Society

www.wileyonlinelibrary.com

Arguing collaboratively: Argumentative discourse types and their potential for knowledge building

Mark Felton[1]*, Merce Garcia-Mila[2], Constanza Villarroel[2] and Sandra Gilabert[3]

[1]Lurie College of Education, San José State University, San Jose, California, USA
[2]Department of Developmental and Educational Psychology, University of Barcelona, Spain
[3]School of Education, Rovira i Virgili University, Tarragona, Spain

Background. There is growing interest in using argumentative discourse in educational settings. However, in a previous study, we found that discourse goals (persuasion vs. consensus) while arguing can affect student outcomes in both content learning and reasoning.

Aims. In this study, we look at argumentative discourse data from a previous study to ask how differences in discourse might account for the differences we observed in learning and reasoning outcomes.

Sample. One hundred and five dialogues (57 disputative, 48 consensus) between 7th grade science students attending a public high school near Tarragona, Spain.

Methods. Participants were randomly assigned to conditions and paired with peers who disagreed with them on three topics related to renewable energy sources. After instruction on each topic, they were asked to either 'argue to convince' (persuasion condition) or 'argue to reach consensus' (consensus condition) on that topic. Conversations were audio-recorded and transcribed for analysis.

Results. Students in the persuasion condition engaged in shorter conversational exchanges around argumentative claims and were more likely to use moves that foreclosed discussion, whereas students in the consensus condition were more likely to use moves that elicited, elaborated on, and integrated their partners' ideas.

Conclusions. When arguing to reach – rather than defend – a conclusion, students are more likely to coconstruct knowledge by exchanging and integrating arguments. These findings are consistent with predictions about the potential of argumentation for knowledge building and suggest that teachers must attend to discourse goals when using argumentation to support learning and reasoning.

Across a range of academic disciplines, there is growing interest in using argumentative discourse to promote learning and reasoning in schools (Hess, 2009; Keefer, Zeitz, & Resnick, 2000; Osborne, 2010; Yackel & Cobb, 1996). This interest stems from a variety of advantages that argumentative discourse presents for learning, including increased engagement in conceptual learning (Nussbaum & Sinatra, 2003), the opportunity to

*Correspondence should be addressed to Mark K. Felton, Lurie College of Education, San Jose State University, 1 Washington Square, San Jose, CA 95192-0077, USA (email: mark.felton@sjsu.edu).

DOI:10.1111/bjep.12078

develop disciplinary thinking skills (Garcia-Mila & Andersen, 2007), and the active production of knowledge (Andriessen, 2007; Driver, Asoko, Leach, Scott, & Mortimer, 1994; Howe, 2014; Leitão, 2000). But to harness its pedagogical power, it is important to understand how argumentative discourse promotes learning and reasoning and the conditions under which it is most effective.

According to Mercier and Sperber (2011), argumentative discourse serves two fundamental communicative functions: To convince others to accept an argument (persuade) and to assess arguments posed by others (deliberate). They propose that the value of argumentative reasoning is best understood not as a form of independent thinking, given the well-documented flaws and biases in our decision-making, but instead as a context for the social exchange of views. In the social context, argumentative discourse offers the potential dual benefits of exposure to alternative arguments (increasing our access to ideas and information) and the opportunity to critically examine opposing arguments to optimize decision-making. But argumentative discourse can provide more than the opportunity to advance or reflect on arguments in groups and potentially extends to a third, related function. As we advance and examine conflicting claims and evidence through discourse, we have the opportunity to collaborate with others to integrate our knowledge and refine our collective understanding (Asterhan & Schwarz, 2007; Mercer, 2000; Resnick, Salmon, Zeitz, Wathen, & Holowchak, 1993). Discussing alternative arguments and counterarguments establishes a dialectical exchange, which under certain conditions opens the door to subtle shifts in our thinking (Billig, 1996) from qualifications of our original position to withdrawal of an argument. Working in tandem with a friendly critic, we have the opportunity to coconstruct new arguments that integrate valid points from multiple perspectives; this process in turn allows us to develop more reasoned, refined, and robust conclusions (Leitão, 2000). Understood in this way, argumentative discourse becomes a context for the social construction of knowledge and a catalyst for collaborative learning.

Researchers have referred to this kind of coconstructive argumentative discourse as collaborative argument (Golanics & Nussbaum, 2008; Nussbaum, 2008), collective argumentation (Brown & Renshaw, 2000), critical discussion (Keefer *et al.*, 2000), and exploratory talk (Mercer, 2000). And while these constructs differ in many ways, particularly in context and the methods by which they were studied, they do share several defining features. All involve the exchange of competing ideas, the critical examination of these ideas, and the social negotiation of meaning. They have also all provided evidence of the positive effects of argumentative discourse to promote more rigorous thinking and more effective learning in laboratory and classroom settings. And yet, despite evidence of the benefits of argumentative discourse for learning and reasoning, it is not difficult to cite instances where public dialogue sinks well below any standards of effective persuasion and reasoned deliberation. In fact, some of the same flaws and biases found in individual reasoning can be provoked or even amplified by argumentative discourse (Ferretti, MacArthur, & Dowdy, 2000; Nussbaum & Kardash, 2005; Wolfe & Britt, 2008).

Argumentative discourse goals

One way to resolve the apparent contradiction in the empirical literature is to recognize the natural tension between the persuasive and deliberative functions of argumentative discourse. According to Walton (2010), arguments must be understood not only in terms of their logical structure, but also by the purposes to which they are put. In persuasion dialogue, we argue to convince, whereas in deliberation dialogue, we argue to decide.

These distinct goals, in turn, impact both the norms and the nature of engagement. Persuasion dialogue is an adversarial exchange in which speakers advance incompatible claims with the goal of convincing others to accept their claim. In some situations, speakers may be motivated to dismiss or even suppress effective opposing arguments and critiques in the interest of advancing their own position. There is evidence from the psychological literature to support this claim. Wolfe and Britt (2008) have found that although individuals will search through text on both sides of an issue when preparing a persuasive essay, they withhold opposing-side claims and evidence when subsequently asked to craft an argument. Importantly, these effects hold even when participants are asked to play the devil's advocate and write a persuasive essay that goes against their own true position on an issue. In a similar vein, Nussbaum and Kardash (2005) have found that individuals will fail to present and rebut opposing-side arguments when asked to persuade, but will include these elements of argument when explicitly prompted to do so (see also Ferretti *et al.*, 2000). These findings suggest that the omission of opposing-side arguments stems less from a biased search or a lack of empathy for opposing-side arguments and more from beliefs about how to craft an effective persuasive argument.

In contrast to persuasion, deliberation dialogue is a collaborative exchange in which speakers hold incompatible claims and seek to resolve these differences to arrive at a consensual decision. As with persuasion dialogue, deliberation dialogue includes a stage in which speakers introduce and critically examine opposing arguments (Walton & Krabbe, 1995). However, deliberation dialogue requires an additional activity in the argumentation stage, which Walton labels 'revision', whereby proposals and views are adapted in the light of incoming arguments and evidence (Walton, 2010). This collaborative activity creates the motivation and opportunity for argument–counterargument integration in argumentative discourse. Again there is evidence to support this proposal. Crowell, Felton, and Liu (2014) have found that college students are more likely to cite arguments that originated from their peers and are more likely to integrate arguments with counterarguments when they have been asked to argue to reach consensus in a chat with a peer than when they have been asked to argue to persuade a peer. In this case, goal instructions, manipulated to contrast directly the effects of persuasion dialogue with deliberation dialogue, had a clear impact on the quality of arguments written immediately afterwards. Thus, the extant literature suggests that discourse goals have a direct influence on the integration of arguments and counterargument when individuals write after arguing.

There is also evidence that discourse goals affect the degree to which individuals learn the content of the arguments they discuss. In a previous study, Felton, Garcia-Mila, and Gilabert (2009) found that argumentative discourse goals (again, arguing to persuade vs. arguing to reach consensus) had an impact not only on the participants' quality of written arguments from pretest to post-test, but also on their ability to learn the content that formed the reasons and evidence supporting conclusions on either side of an issue. The study was run in a middle-school science curriculum unit on renewable energy sources and their relative advantages with respect to energy production, cost, efficiency, and environmental impact. Adolescents in the two discourse conditions received instruction on content related to each topic of discussion and engaged in three different dialogues with peers. Students in the consensus condition produced more complex, two-sided essays in support of their positions and also learned content related to the pros and cons of the energy sources better than their peers in the persuasion-goal condition even when exposure to content was controlled for through instruction. These findings suggest that the discourse goals in the two conditions affected both argument quality and content

learning at the post-test. These findings are of particular interest because students were blind to their conditions and the nature of the research, suggesting that a simple manipulation of goals ('Try to convince your partner if you can' vs. 'Try to reach consensus with your partner if you can') produces different responses. A subsequent study of the data (Garcia-Mila, Gilabert, Erduran & Felton 2013) revealed that more complex argument structures were coconstructed in these dialogues. However, dialogue data from the study were not analysed to search for the mechanisms by which the learning outcomes may have unfolded.

The present study

In this study, we return to the data from this previous investigation (Felton *et al.*, 2009) to look more closely at the characteristics of the students' argumentative discourse. Here, we focus on the nature of dialogue in each condition to look for evidence of persuasion and deliberation dialogue. As an intervention validation, we examine whether the task goals in that study indeed produced these two forms of discourse. If we cannot find discernible differences in discourse, then other variables associated with task goals would have to be explored to explain student learning outcomes. However, if differences are found in the kinds of discourse produced, it would be useful to examine discourse patterns to get a better understanding of the mechanisms by which task goals (to persuade vs. to reach consensus) lead to between-group differences in learning.

Research questions:

1. Did our two discourse goal conditions (persuasion vs. consensus) produce significantly different types of discourse during our intervention? If so, how did they differ?
2. What do between-group differences in discourse tell us about the relative advantages that arguing to reach consensus provides for learning?

Methods

Design

A between-subjects design was used to compare the effects of task instructions on argumentative discourse. Task instructions were presented in one of two ways: (1) a consensus condition, where students were prompted to argue to reach consensus; (2) a persuasion condition, where students were prompted to argue to convince the partner.

Participants

Participants were seventy 7th grade science students attending a public high school near Tarragona, Spain. The participants' mean age was 12.2 (SD = 0.4) (range 12.0–13.0). Students were proportionally pooled from five classes to ensure roughly equal representation from each class and randomly assigned to two conditions. Thirty-eight students were assigned to the persuasion condition (19 female and 19 male), and 32 students (16 female and 16 male) were assigned to the consensus condition. For each of the three topics of discussion in the study, participants were paired with a new partner from the same condition who genuinely disagreed with them on the topic of discussion. Because dyads changed from one topic to the next and because the interactive nature of dialogue data does not allow us to observe individuals in dyads

independently, the unit of analysis for this study is the dyad. The complete sample is made up of 105 dialogues (57 in the persuasion condition and 48 in the consensus condition). In the end, both groups had a combination of same-sex and mixed-sex dyads in roughly equal proportions. In the persuasion condition, 20 dyads were same-sex (male), 20 dyads were same-sex (female), and 17 dyads were mixed groupings. In the consensus condition, 15 dyads were same-sex (male), 15 dyads were same-sex (female), and 18 dyads were mixed groupings.

Procedure

The dialogues were embedded in a unit on the topic of climate change and renewable energies, which took place across eight 50-min sessions presented over a 3-week period. The first session of the unit was dedicated to pretesting students on content knowledge for the curriculum. For the next six sessions, students went through three two-day mini-units, each dedicated to a dilemma (thermal vs. nuclear energy, wind or not, and biodiesel or not). On the first day of each mini-unit, students learned content related to the week's dilemma and then wrote a short essay taking a position on the dilemma. On the second day of each mini-unit (sessions 3, 5 and 7), students in the two discourse conditions were paired with a disagreeing partner and the two were asked to discuss the dilemma with the goal of either persuading their partner (persuasion condition) or reaching an agreement (consensus condition). Students were blind to their condition as well as the nature of the study and all dyads were given equal time to engage in discussion, although actual discussion times varied. Finally, at the end of session 8, all the students took a post-test in which they answered content questions from each mini-unit and wrote an essay supporting their energy plan for the city.

Instruments

Analytic scheme for argumentative discourse

To answer the research questions, we have adapted a coding scheme developed for argumentative discourse (Felton, 2004; Felton & Kuhn, 2001). That scheme was originally developed to analyse argument in persuasive contexts, so we have adapted the scheme to make it more sensitive to the more collaborative forms of argument that one might expect to encounter when individuals argue to reach consensus (see Appendix S1). For each conversational turn in a dialogue, raters assigned one or more codes to represent the communicative function of that utterance. When raters disagreed on the number of codes in a given turn, the higher number was added to the total number of codes and the disagreement was added to the tally. Raters then resolved any disagreements by discussion before moving on to the next conversational turn. Inter-rater reliability, tallied using 25% of the dialogues in the data set, was .821 (Cohen's κ).

The first wave of data analysis focused on the proportional use of argumentative moves in each condition. Simple frequencies for each move in a dialogue were tallied and divided by the total number of coded moves in each dialogue to calculate proportions. Next, data were analysed for the relative length of argumentative exchanges in each dialogue. An argumentative exchange is defined as a sequence of transactive turns in a dialogue, which ends when one of the two speakers introduces an unconnected question or statement. An exchange can be understood as the total number of turns that conversational participants take before moving on to another claim for discussion. The length of exchanges

(corrected for repetitions of argumentative moves and meta-conversational talk) can be used as a rough indicator of the depth to which a pair has discussed a claim before moving on.

Analytic scheme for meta-conversational talk

Meta-conversational talk (*Meta-talk*) refers to dialogue that departs from the topic of discussion to address the discussion itself or the speakers in the discussion. Because meta-talk comprised a large proportion of the coded dialogue with significant between-group differences, a new scheme was developed to categorize these moves for further analysis. All instances of meta-talk were recoded using a refined coding scheme by researchers blind to the dialogue conditions (see Appendix S2). Three broad categories were developed to capture whether meta-talk addressed the goals of the dialogue (meta-goal), reasoning and the exchange of claims in the dialogue (meta-reasoning), or the procedural elements of the dialogue (meta-procedural). Each of these categories was divided into subcategories developed from the data and finally, each subcategory was divided into productive talk, which focused on positively engaging the partner or promoting the exchange of ideas, and counterproductive talk, which focused on negatively engaging the partner or foreclosing the exchange of ideas. Again, two scorers, blind to the condition and identity of participants, coded the complete data set of meta-talk. Inter-rater reliability for the meta-talk coding scheme, calculated using 25% of the dialogues in the data set, was .817 (Cohen's κ).

Results

Length of dialogues

A between-group comparison of the mean length of dialogues (measured through a word count) has been reported elsewhere (Garcia-Mila *et al.*, 2013) and was non-significant.

Analysis of discourse data

The only move that was normally distributed was *meta*, and we used a Student's *t*-test for means comparison. For the rest of the statistical comparisons, we used the Wilcoxon tests. To reduce the risk of Type I error accumulation, we excluded from the statistical analysis those moves with means that were below the 50th percentile (i.e., with means lower than .009). Table 1 shows the means and standard deviations of the 22 argumentative moves that were analysed. Correcting for Bonferroni on these 22 tests, we considered *p*-values that were <.0023 (or .05/22) as significant. In the right column, we indicate the condition that had a higher mean. *Counter-UC* (*counter unjustified critique*), *Disagree*, and *Meta* were the only three moves whose mean was significantly higher in the persuasion condition, while *Accept*, *Accommodate-PQ*, *Advance*, and *Recap* had significantly higher means in the consensus condition. Means (and *SD*) for *Counter-UC* for the consensus and the persuasion condition were 0.0504 (0.039) and 0.0155 (0.019), respectively, $W = 1{,}552$, $p = .001$ (effect size $r = .492$). For *Meta*, means were 0.1497 (0.077) and 0.0863 (0.070), $t(96) = 4.2$, $p = .001$ (effect size $d = .969$) for the persuasion and consensus conditions, respectively. For those moves whose means were significantly lower in the persuasion condition than the consensus condition, means (*SD*) for *Accept* were 0.0034 (0.010) and 0.0251 (0.021), respectively ($W = 1{,}812$, $p = .001$, $r = .644$); for

Table 1. Means and standard deviations for discourse moves with tests of significance

Discourse moves	Condition	*M*	*SD*	Statistics	Higher mean	Effect size
Clarify?	Persuasion	0.0131	0.015	$W = 2{,}406$		
	Consensus	0.0121	0.015	ns		
Justify?	Persuasion	0.0133	0.020	$W = 2{,}913$		
	Consensus	0.0140	0.020	ns		
Position?	Persuasion	0.0051	0.011	$W = 231$		
	Consensus	0.0100	0.012	ns		
Stance?	Persuasion	0.0144	0.021	$W = 2{,}942$		
	Consensus	0.0149	0.021	ns		
Accept	Persuasion	0.0034	0.010	$W = 1{,}812$		
	Consensus	0.0251	0.021	$p = .001$*	Consensus	$r = .644$
Accomodate + PQ	Persuasion	1.2632	1.575	$W = 2{,}523$		
	Consensus	2.8333	2.823	$p = .001$*	Consensus	$r = -.333$
Advance	Persuasion	0.0012	0.006	$W = 2{,}327$		
	Consensus	0.0090	0.017	$p = .001$*	Consensus	$r = .341$
Agree	Persuasion	0.0035	0.007	$W = 2{,}337$		
	Consensus	0.0976	0.013	ns		
Add	Persuasion	0.0037	0.009	$W = 2{,}709$		
	Consensus	0.0112	0.021	ns		
Anticipate	Persuasion	0.0178	0.030	$W = 2{,}812$		
	Consensus	0.0223	0.027	ns		
Argument	Persuasion	0.1362	0.086	$W = 2{,}712$		
	Consensus	0.1594	0.080	ns		
Clarify	Persuasion	0.0138	0.016	$W = 2{,}892$		
	Consensus	0.0146	0.017	ns		
Concede	Persuasion	0.0273	0.023	$W = 2{,}744$		
	Consensus	0.0347	0.029	ns		
Continue	Persuasion	0.0394	0.036	$W = 2{,}331$		
	Consensus	0.0321	0.269	ns		
CA	Persuasion	0.0524	0.035	$W = 2{,}818$		
	Consensus	0.0594	0.043	ns		
CC	Persuasion	0.0524	0.035	$W = 2{,}772$		
	Consensus	0.0594	0.043	ns		
Counter-UC	Persuasion	0.0504	0.039	$W = 1{,}552$		
	Consensus	0.0155	0.019	$p = .001$*	Persuasion	$r = .492$
Disagree	Persuasion	0.0494	0.052	$W = 1{,}851$		
	Consensus	0.0232	0.029	$p = .006$		
Dismiss	Persuasion	0.0147	0.017	$W = 2{,}244$		
	Consensus	0.0111	0.017	ns		
Meta	Persuasion	0.1497	0.077	$t(96) = 4.2$		
	Consensus	0.0863	0.070	$p = .001$*	Persuasion	$d = .9.69$
Rebut	Persuasion	0.0293	0.038	$W = 2{,}447$		
	Consensus	0.0244	0.027	ns		
Recap	Persuasion	0.0087	0.027	$W = 1941.5$		
	Consensus	0.0306	0.021	$p = .001$*	Consensus	$r = .412$

*Statistically significant after Bonferroni correction.

Accommodate + PQ, they were 0.0203 (*SD* = 0.026) and 0.0432 (*SD* = 0.046), respectively (*W* = 2,523, *p* = .001, *r* = .333); for *Advance*, they were 0.0012 (0.006) and 0.0089 (0.017), respectively (*W* = 2,327, *p* = .001, *r* = .341); and for *Recap*, they were 0.0087 (0.027) and 0.0306 (0.021), respectively (*W* = 1,941, *p* = .001, *r* = .412). Between-group differences for the remaining moves were non-significant (see Table 1).

Analysis of conversational exchanges

The length of argumentative exchanges was calculated and the means for each condition compared. The mean (*SD*) for the persuasion condition was 2.56 (1.05), and the mean for the consensus condition was 3.4 (1.22). The Student's *t*-test yielded significant differences, $t(96) = -3.67$, $p = .001$, effect size $d = .738$.

Analysis of meta-talk (post-hoc)

Given that the means for meta-talk were exceptionally high and between-group differences were significant, we performed *post-hoc* tests to analyse between-group differences in subcategories of *Meta*. In the interest of reducing the likelihood of Type I error, we limited analyses to meta-moves with means above .008 (i.e., in the top 40th percentile). The resulting moves were productive (+) or counterproductive talk (−) for *Goal Setting* (*Meta-G GS*), *Goal Checking* (*Meta-G GC*), *Goal Met* (*Meta-G GM*), and *Managing Stalls* (*Meta-P ST*). Thus, eight between-group comparisons of means were made in all. After correcting for Bonferroni ($p < .05/8 = .0062$), all tests for counterproductive talk (−) yielded significant between-group differences except in the case of Meta-P ST. Means for *Meta-G GC* were 0.0116 (0.021) and 0.0014 (0.006) for the persuasive and consensus conditions, respectively, *W* = 2,052, *p* = .006, *r* = .113. Means for *Meta-G GC* were 0.0044 (0.012) and 0.0002 (0.001) for the persuasive and consensus conditions, respectively, *W* = 2,234, *p* = .006, *r* = .240. Means for the *Meta-G GM* were 0.0109 (0.022) and 0.0005 (0.003) for the persuasive and consensus conditions, respectively, *W* = 2,115, *p* = .001, *r* = .314. In addition, significant differences were found in the means (*SD*) for *Meta-P ST+*. These were 0.0142 (0.026) and 0.0015 (0.005) for the persuasion and consensus conditions, respectively, *W* = 2,128, *p* = .003, *r* = .217. Differences in the rest of the means for productive talk moves and for *Meta-PST* were non-significant across conditions (see Table 2).

Analysis of dialogue

Here, we present two typical exchanges that illustrate the discourse patterns in each of the two conditions. In the first excerpt (Table 3), we see two students from the persuasion condition arguing about nuclear and thermal (i.e., petroleum based) power. Alex and Jesse both make claims about the relative strengths or weaknesses of the options (lines 1, 4, 5, and 13), but they spend relatively little time talking about each of these claims. Instead, each speaker competes for 'airtime' by interrupting the partner (lines 6–7, 10–11) and advancing claims that do not address the partner's previous utterance (lines 1, 4, and 7). In the one case where Alex invites Jesse to justify the claim that nuclear power stations generate more energy (line 2), Alex ignores Jesse's answer and simply moves on to a different point (line 4). Spending relatively little time on each claim or simply denying the truth of these claims (lines 5 and 14) offered little opportunity for these students to make sense of alternative claims or integrate knowledge with their positions on the topic.

Table 2. Means and standard deviations for meta-conversational moves with tests of significance

Meta-conversational move	Condition	*M*	*SD*	Statistics	Higher mean	Effect size
GGS+	Persuasion	0.0045	0.013	$W = 2{,}694$		
	Consensus	0.0078	0.012	ns		
GGS−	Persuasion	0.0116	0.021	$W = 2{,}052$		
	Consensus	0.0014	0.006	$p = .006$	Persuasion	$r = .113$
GGC+	Persuasion	0.0165	0.033	$W = 2{,}888$		
	Consensus	0.0120	0.017	ns		
GGC−	Persuasion	0.0044	0.012	$W = 2{,}234$		
	Consensus	0.0002	0.001	$p = .006$	Persuasion	$r = .240$
GGM+	Persuasion	0.0141	0.021	$W = 2{,}304$		
	Consensus	0.0083	0.013	ns		
GGM−	Persuasion	0.0109	0.022	$W = 2{,}115$		
	Consensus	0.0005	0.003	$p = .001$	Persuasion	$r = .314$
PST+	Persuasion	0.0142	0.026	$W = 2{,}128$		
	Consensus	0.0015	0.005	$p = .003$	Persuasion	$r = .321$
PST−	Persuasion	0.0059	0.015	$W = 2{,}304$		
	Consensus	0.0011	0.003	ns		

Table 3. Sample coded dialogue from persuasion condition

Speaker	Dialogue	Discourse move(s)
Jesse	Nuclear power plants generate more energy	Counter-A
Alex	Why?	Justify-?
Jesse	Because the atoms generate more energy than your coal	Justify
Alex	But there is danger of the radioactivity	Counter-A
Jesse	It's not that dangerous./Besides, think about climate change. That's really damaging	Dismiss/Counter-A
Alex	Look, the thermal power plant is better. . .	Counter-UC
Jesse	No, nuclear power plants don't cause the greenhouse effect	Counter-A
Alex	So what?	Clarify-?
Jesse	So then we could die	Clarify
Alex	No./What are you talking about? Look, I will explain you again, this is the third time: This is Spain. . .	Disagree/Meta-R CL-
Jesse	I don't care where Spain is!	Dismiss
Alex	Well, you haven't convinced me	Meta-G GC-
Jesse	What? And what happens when the thermal power plant causes acid rain, then what?	Counter-A
Alex	It doesn't cause acid rain	Disagree
Jesse	Yes, it does	Disagree
Alex	No	Disagree

The second excerpt (Table 4) is taken from the consensus condition where the two students discuss their views on biodiesels. Chris is in favour of using biodiesels, while Pat is against them. In line 1, Pat introduces the critique that biodiesels can lead to the exploitation of poorer, biomass-producing countries, but proposes a way to address this critique as well. Chris accepts this proposal and extends it by introducing ways in which

Table 4. Sample coded dialogue from consensus condition

Speaker	Dialogue	Discourse move(s)
Pat	Like I said, you need a large amount of land and that means that South America and Africa would be left without land to grow their own food. They need to eat, too./Plus, it'll also destroy a lot of forests and their ecosystems./Anyway, I say we use biofuels, and grow biomass, but we make sure we leave enough land for producing countries to grow their own food	Counter-C/ Argument/ Accommodate
Chris	That's a good suggestion./Look, what I would do is pass a law to make sure biofuels are introduced gradually and become the energy of the future. What I would try to do is have each country keep some land for growing and work hard to grow their own biomass. If a country needs more biomass, other countries could sell them their extra or even exchange it for food	Accept/Advance
Pat	Good idea./If one country has to buy biomass from another country, the government of that country will try not to consume so much biofuel because buying it outside would be too expensive	Accept/Advance
Chris	Right,/so they wouldn't be interested in buying biomass from other countries. The government would try to use only the biomass produced in their own country	Agree/Add
Pat	And maybe that's the key to getting people to adopt other renewable energies like solar	Advance

to balance the power dynamic (line 2). This solution leads both students to recognize ways in which laws can be used as incentives for energy efficiency.

In the end, Pat agrees to adopt Chris' position but does so in a way that has allowed each speaker to consider the other's arguments. Because the two feel free to take up and build upon one another's ideas, they have the opportunity to make sense of opposing-side claims and integrate knowledge from both sides of the issue. The search for consensus has created opportunities for the coconstruction of knowledge.

Discussion

As our analyses indicate, students in the two conditions engaged in different kinds of discourse along several dimensions. Students in the persuasion condition engaged in shorter exchanges around each claim, produced more unjustified critiques (Counter-UC), and were more likely to engage in counterproductive meta-conversation with peers (Meta-). When engaged in meta-conversation, they were more likely to assert that they had no intention of being persuaded (Meta-G GS-), to disingenuously announce that they had persuaded their partner when they hadn't (Meta-G GM-), and to note that they had nothing more to say (Meta-P ST+). In contrast, students in the consensus condition engaged in longer exchanges around each claim to extend their partners' claims with additional argument (Advance), to adapt their own arguments in response to valid critiques from their partners (Accommodate-PQ), to summarize major points made in the dialogue (Recap), and to assent to these recaps and accommodations (Accept). There were also several discourse moves with relatively high means in both groups that did not show significant between-group differences. Most notably, students in both conditions

produced relatively high and comparable proportions of claims to justify their positions (Argument, Counter-A) and counterarguments and rebuttals (Counter-C and Rebut) to challenge each other's claims.

Taken together, the data support the view that students pursued the argumentative goals assigned to them in each condition. Students in both groups exchanged claims in support of their opinions and critiqued those claims, as would be expected in both persuasive and deliberation dialogue (Walton & Krabbe, 1995), but their discourse differed in areas that one might expect as well. Whereas students in the persuasion condition spent more time entreating their partners to change sides and resisting the open exchange of ideas, students in the consensus condition spent more time elaborating each other's ideas and working out ways to integrate opposing views.

So what do these data tell us about the potential for each discourse type to promote knowledge construction and integration? The results suggest that the competitive discourse goals of the persuasion condition elicited discourse moves that constrained and, in some cases, shut down the exchange of ideas in students' dialogues. Shorter conversational exchanges, the use of unjustified critiques ('my plan is better!'), and more frequent stalls in conversation ('I don't know what else to say') signal a limited investigation of ideas and breakdowns in dialogue about arguments. Explicit refusals to be persuaded and false claims to having won the debate also indicate an obstinacy unconducive to the thoughtful exchange and examination of arguments. In short, while students in the persuasion condition posed arguments to one another in equal measure to their peers in the consensus condition, the competitive goals of the condition led students to spend less time responding to their partner's claims and offered them a limited range of responses to those claims (primarily counterarguments and rebuttals). These constraints in the time and effort spent on processing each other's arguments certainly help explain the lower student outcomes in learning and reasoning in this group that were reported in our previous study (Felton *et al.*, 2009).

In contrast, the data suggest that the collaborative goals of the consensus condition produced more opportunities for the coconstruction and integration of knowledge, which in turn may help explain why students in this condition learned more content and wrote more advanced arguments at the post-test (as reported in Felton *et al.*, 2009). Students in this condition engaged in longer conversational exchanges and made use of a wider palette of responses to opposing arguments. They extended their partners' thinking, recapped key ideas, and adapted their arguments to accommodate valid critiques. In trying to accommodate their partners' critiques and offering qualifications to their position, students integrated opposing views and refined their thinking. But these conversational moves, though consensus-driven, were not shallow attempts at compromise, which themselves can be a barrier to learning (Damsa, Ludvigsen, & Andriessen, 2013; Mercer, 2000). Students in this condition did not start by agreeing with each other and were just as likely as their peers in the persuasion condition to advance opposing arguments and critically examine ideas. However, in this case, critical dialogue became the springboard for the integration of arguments. In essence, students in the consensus condition pursued three norms that facilitated open dialogue and deeper processing of the content:

1. A commitment to invite and understand opposing arguments
2. A commitment to critically examine the reasons and evidence supporting arguments on both sides of the issue.

3. A commitment to resolve and reconcile conflicting claims and evidence where possible

Pursuing mutually exclusive outcomes, students in the persuasion condition found themselves in conflict with one other and frustrated in their attempts to control the direction and outcome of dialogue. As a result, they were more resistant to understanding alternative views, less committed to validating legitimate critiques of their own views, and less inclined to reconcile conflicting claims in any way other than to rebut them. Because students in the consensus condition had a full range of responses to opposing arguments open to them, from qualifying their original arguments, to adapting their arguments to accommodate valid critiques, to the accepting opposing-side arguments and changing position, they were more likely to make use of the content advanced by a partner in ways that promoted learning.

Of course, this study was based on the dialogues of young adolescents and may not apply to older, more expert, arguers. More advanced arguers may be more adept at navigating the cognitive and social conflict evoked by persuasion goals and may also appreciate the value of integrating conflicting claims and evidence to be more persuasive. As a result, they may not demonstrate the same differential effects of discourse goals on learning and reasoning we have seen here with adolescents or elsewhere with young adults (Crowell *et al.*, 2014; Nussbaum & Kardash, 2005). In addition, these data were collected in peer dyads where students did not benefit from the interventions and guidance of a teacher who could help guide thinking, mediate conflict, and keep talk productive (Forman, Larreamendy-Joerns, Stein, & Brown, 1998; Mercer & Littleton, 2007; Yackel, 2002), but this was the very point of our study. We wanted to see what kinds of argumentative dialogue students at this age group would default to when given persuasive and consensus goals. Certainly teachers can influence the course and conduct of dialogue if there to intervene, but such intervention must operate in concert with students' own goals and intentions while arguing. Furthermore, we believe that it is telling that students so clearly fell into distinct kinds of discourse patterns with nothing more than a simple manipulation in instructions. Without any specific training in the features of persuasive and deliberation dialogue, they naturally produced each, indicating that by early adolescence, students are capable of adjusting argumentative discourse to fit their purposes. Finally, discourse analysis inevitably introduces issues with non-independence of observation as each individual's contributions to the dialogue are determined in part by the contributions of his or her partner. For this reason, we have made the dyad the unit of analysis in this paper and focused our inferences on between-group differences in the quality and characteristics of dialogue in each condition. In addition, because each student participated in three dialogues, there is a recurrence of individuals in the data set. We have tried to mitigate the potential influence of this repetition on our findings by pairing individuals with new partners in each dialogue to create unique dyads for each dialogue and by maintaining parity between conditions in the assignment of individuals to dyads.

As a communicative tool, language does more than express thought; it shapes thought as well (Cazden, 2001; Wells, 1999; Wertsch, 1991). When we engage in argumentative discourse, we contrast competing perspectives, question and challenge claims, and build a stronger foundation in support of our views (Kuhn, 2005). However, depending on our intentions, we may also argue to invite new ideas or to probe, explore, and elaborate each other's thinking. Together, these communicative acts offer rich opportunities to go beyond our present understanding, allowing us to examine, refine, and even revise our thinking. These aims foster the educational potential of argument to support knowledge

construction (Leitão, 2000) and conceptual change (Dole & Sinatra, 1998). But we need fertile ground for this kind of generative discourse to occur and the social context for the arguments in which we engage must facilitate the open exchange of ideas and motivate a genuine search for a richer, deeper and more nuanced understanding of the world. In the short run, contextual factors like argumentative discourse goals may drive the impact of social context on learning and reasoning, as we have seen in the present study. With prolonged exposure to collaborative argument over the time, however, these contextual factors may have an even greater cumulative effect on the skills and dispositions that we develop for leveraging talk to engage in collective reasoning (Driver, Newton, & Osborne, 2000; Kuhn, Wang, & Li, 2010; Lemke, 1990; Michaels, O'Connor, & Resnick, 2008). Ultimately, successful collaborative experiences in argumentative discourse may make us more inclined to see discourse as a useful context for extending our thinking, challenging our perspective, and exploring new ideas.

Acknowledgements

The present work was made possible thanks, in part, to funding for the first author by the Ministerio de Educacion, Spain (SAB2010-0124), the second author by the the Ministerio de Ciencia y Innovacion, Spain (EDU2013-47593-C2-2-P), the third author by the Universitat de Barcelona (APIF2013) and the fourth by the Generalitat de Catalunya, Spain, and the European Social Funds (EDU/3600/2006).

References

Andriessen, J. (2007). Arguing to learn. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 443–460). Cambridge, UK: Cambridge University Press.

Asterhan, C. S., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, *99*, 626–639. doi:10.1037/0022-0663.99.3.626

Billig, M. (1996). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge, UK: Cambridge University Press.

Brown, R., & Renshaw, P. (2000). Collective argumentation: A sociocultural approach to reframing classroom teaching and learning. In H. Cowie & G. Van der Aalsvoort (Eds.), *Social interaction in learning and instruction: The meaning of discourse for the construction of knowledge* (pp. 52–66). New York, NY: Pergamon.

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning*. (2nd ed.) Portsmouth, NH: Heinemann.

Crowell, A., Felton, M., & Liu, T. (2014). *Arguing to agree: Mitigating the effects of my-side bias through consensus-seeking dialogue*. Manuscript submitted for publication.

Damsa, C., Ludvigsen, S., & Andriessen, J. (2013). Knowledge co-construction: Epistemic consensus or relational assent? In M. Baker, S. Jarvela & J. Andriessen (Eds.), *Affective learning together: Social and emotional dimensions of collaborative learning* (pp. 97–119). London, UK: Routledge.

Driver, R., Asoko, H., Leach, J., Scott, P., & Mortimer, E. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, *23*, 5–12. doi:10.3102/0013189X023007005

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*, 287–312. doi:10.1002/(SICI)1098-237X(200005)84:3<287:AID-SCE1>3.0.CO;2-A

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33*, 109–128. doi:10.1080/00461520.1998.96.53294

Felton, M. K. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development*, *19*, 35–52. doi:10.1016/j.cogdev.2003.09.001

Felton, M., Garcia-Mila, M., & Gilabert, S. (2009). Deliberation versus dispute: The impact of argumentative discourse goals on learning and reasoning in the science classroom. *Informal Logic*, *29*(4), 417–446.

Felton, M., & Kuhn, D. (2001). The development of argumentive discourse skill. *Discourse Processes*, *32*, 135–153. doi:10.1080/0163853X.2001.9651595

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. doi:10.1037/0022-0663.92.4.694

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction*, *8*, 527–548. doi:10.1016/S0959-4752(98)00033-4

Garcia-Mila, M., & Andersen, C. (2007). Cognitive foundations of learning argumentation. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (Vol. *35*, pp. 29–45). Dordecht, the Netherlands: Springer.

Garcia-Mila, M., Gilabert, S., Erduran, S., & Felton, M. (2013). The effect of argumentative task goal on the quality of argumentative discourse. *Science Education*, *97*(4), 497–523. doi:10.1002/sce.21057

Golanics, J. D., & Nussbaum, E. M. (2008). Enhancing online collaborative argumentation through question elaboration and goal instructions. *Journal of Computer Assisted Learning*, *24*, 167–180. doi:10.1111/j.1365-2729.2007.00251.x

Hess, D. E. (2009). *Controversy in the classroom: The democratic power of discussion*. London, UK: Routledge.

Howe, C. (2014). Optimizing small group discourse in classrooms: Effective practices and theoretical constraints. *International Journal of Educational Research*, *63*, 107–115. doi:10.1016/j.ijer.2013.03.011

Keefer, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction*, *18*, 53–81. doi:10.1207/S1532690XCI1801_03

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

Kuhn, D., Wang, Y., & Li, H. (2010). Why argue? Developing understanding of the purposes and values of argumentive discourse. *Discourse Processes*, *48*, 26–49. doi:10.1080/01638531003653344

Leitão, S. (2000). The potential of argument in knowledge building. *Human Development*, *43*, 332–360. doi:10.1159/000022695

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing Corporation.

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

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. London, UK: Routledge.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57–74. doi:10.1017/S0140525X10000968

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*, 283–297. doi:10.1007/s11217-007-9071-1

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology*, *33*, 345–359. doi:10.1016/j.cedpsych.2008.06.001

Nussbaum, E. M., & Kardash, C. M. (2005). The effects of goal Instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology*, *97*, 157. doi:10.1037/0022-0663.97.2.157

Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology*, *28*, 384–395. doi:10.1016/S0361-476X(02)00038-3

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, *328*, 463–466. doi:10.1126/science.1183944

Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction*, *11*, 347–364. doi:10.1080/07370008.1993.9649029

Walton, D. (2010). Burden of proof in deliberation dialogs. In P. McBurney, I. Rahwan, S. Parsons & N. Maudet (Eds.), *Argumentation in multi-agent systems* (pp. 1–22). Berlin, Germany: Springer.

Walton, D. N., & Krabbe, E. C. (1995). *Commitment in dialogue: Basic concepts of Interpersonal reasoning*. Albany, NY: State University of New York Press.

Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. Cambridge, UK: Cambridge University Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking & Reasoning*, *14*, 1–27. doi:10.1080/13546780701527674

Yackel, E. (2002). What we can learn from analyzing the teacher's role in collective argumentation. *The Journal of Mathematical Behavior*, *21*, 423–440. doi:10.1016/S0732-3123(02)00143-8

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458–477. doi:10.2307/749877

Received 17 November 2014; revised version received 6 March 2015

Supporting Information

The following supporting information may be found in the online edition of the article:

Appendix S1. Argumentative discourse coding scheme.
Appendix S2. Meta-conversational moves.