EFFECT OF COLLABORATIVE LEARNING AND DIRECT INSTRUCTION ON MYSIDE BIAS

BY

JOSHUA ANDREW MORRIS

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Educational Psychology in the Graduate College of the University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee
Professor Richard C. Anderson, Chair
Assistant Professor Emma Mercier
Professor Michael Nussbaum, University of Nevada, Las Vegas
Professor Michelle Perry
This study investigates the effect on instructional discourse on myside bias; the tendency of people to argue without considering different sides of an issue. 766 Fifth-grade students from two urban districts completed a unit on wolves integrating science, social studies, and the English language arts. The study design was quasi-experimental with triples of classroom matched on demographics, and randomly assigned to one of three conditions: Direct Instruction (DI), Collaborative Group Work (CG), and Wait-Listed Control (CTL). Transcripts of a representative sample of classroom videos were coded for low-inference discourse markers of connected and elaborated talk. The analysis indicated that CG students used connective words much more frequently than DI students. CG students also made more challenges and requests for explanation than students or teachers in the DI condition. Three outcome measures designed to elicit student argumentation were analyzed for aspects of reasoning including the consideration or reasons on both sides of the controversy. When writing about the issue of wolves, which students had studied for six weeks, students in DI and CG were equally likely to consider reasons on both sides of an issue. In two transfer tasks - with topics that had varying degrees of overlap with the wolf unit- students in CG outperformed students in DI, who were no more likely to consider both sides than students in the uninstructed control group. Students in CG who considered reasons on both sides of the wolf issue, were twice as likely as DI and CTL students to consider a reason on one or both of the transfer measures. This study provides distinctive new evidence about how instructional frameworks can impact students’ propensity to consider both sides when reasoning about various issues. We theorize that CG students’ increased propensity to consider
both sides was based on their improved sensitivity to potential audiences for the arguments they were constructing, and that this sensitivity was fostered during students’ participation in dialogic argumentation where the challenging of positions and the requesting and providing of explanations was a common occurrence.
ACKNOWLEDGEMENTS

I’d like to start by expressing my profound gratitude to my advisor, Richard C. Anderson. Dick has been a kind and dedicated mentor and I am most thankful for how he always made himself accessible to me and all his students. I don’t think we ever shared a meeting where we weren’t laughing and joy is something that isn’t always easy to come by in graduate school.

I’d also like to thank the other members of my committee. Dr. Michelle Perry helped me keep my eye on the prize and taught me that it’s ok to ask for help. Her editing and support during the writing of the dissertation were indispensable. Dr. Emma Mercier focused my thinking at every stage of the process and challenged me to shift my focus to theories of transfer. Dr. Michael Nussbaum provided crucial feedback on argumentation theory and made sure my conclusions matched my data. They are all peerless scholars in their areas and I appreciate the time and support they willingly gave.

I’d like to give a special thanks to Dr. David Zola. He is a scholar and a gentleman and after my time here I am lucky enough to call him my friend. I would also like to briefly acknowledge Illinois faculty that had a hand in shaping who I am as a scholar. Dr. Kiel Christianson and Dr. Liz Stine-Morrow both taught me how to think critically about methodology and how to shape a research question. Dr. Carolyn Anderson’s classes and office hours are the only reason I understand statistics. Dr. Paul Prior showed me the danger of reductive methods and challenged me to think about the impact of sociocultural forces on learning. Finally, I’d like to thank Dr. William F. Brewer whose expansive knowledge and overhead transparencies challenged my thinking about everything.

I am grateful to my friends with whom I shared this journey: Dr. Brennan Payne, Dr. Brian Miller, and Dr. Jingjing Sun. All provided support in their unique ways and I am better for having known them.
Above all I am thankful to my partner, Theresa Scott. It feels inadequate to try to sum up her support and love with a pithy anecdote or phrase, but that’s the nature of an acknowledgements section. All in all, I wouldn’t be here without her, I wouldn’t have stayed without her, and there is no part of this journey that hasn’t been made better by her.

The research reported in this paper was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A080347 to the University of Illinois at Urbana-Champaign, Richard C. Anderson, Principal Investigator. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.
## TABLE OF CONTENTS

List of Tables ........................................................................................................... viii
List of Figures ............................................................................................................. ix

Chapter 1 Introduction ................................................................................................. 1

Chapter 2 Literature Review and Theoretical Context ................................................ 4
  2.1 Classroom Discourse and Argumentation ............................................................. 11
  2.2 Classroom Learning and Norms of Discourse ...................................................... 11
  2.3 Argumentation in Classroom Discourse ............................................................... 12
  2.4 Transfer Across Contexts .................................................................................... 15
  2.5 Current Study and Predictions ........................................................................... 18

Chapter 3 Methods ....................................................................................................... 21
  3.1 Participants .......................................................................................................... 21
  3.2 Participating Teachers ......................................................................................... 22
  3.3 Wolf Management Unit ...................................................................................... 23
    3.3.1 Teaching Training ....................................................................................... 23
    3.3.2 The Curriculum ............................................................................................ 24
  3.4 Assessments ......................................................................................................... 27
    3.4.1 Wolf Decision Letter .................................................................................. 27
    3.4.2 Whale Controversy Oral Interview ............................................................. 27
    3.4.3 Pinewood Transfer Essay ............................................................................ 28
  3.5 Analysis of Outcome Measures .......................................................................... 29
    3.5.1 Coding ........................................................................................................ 30
    3.5.2 Elements of Argument .............................................................................. 30
    3.5.3 Additional Coding for Oral Interview Task ............................................... 33
    3.5.4 Content of Reasoning ................................................................................ 35
  3.6 Transcription and Coding of Classroom Talk ..................................................... 36

Chapter 4 Results ......................................................................................................... 40
  4.1 Analysis of Classroom Talk ................................................................................ 40
  4.2 Analysis of Outcome Measures ......................................................................... 42
    4.2.1 Students with All 3 Outcomes .................................................................. 46
    4.2.2 Elaboration of Reasons ............................................................................ 48

Chapter 5 Discussion .................................................................................................... 50
  5.1 Overview ............................................................................................................ 50
  5.2 Summary of important findings ....................................................................... 50
  5.3 Evidence for consideration of both sides as a habit of mind .............................. 52
  5.4 Internalizing norms of discourse ..................................................................... 52
  5.5 Collaborative group work and attention to others’ viewpoints ....................... 53
    5.5.1 Transfer of ‘consideration of both sides’ heuristic .................................. 55
5.6 Direct instruction and a different habit of mind .................................................. 56
  5.6.1 Reasoning about the wolf controversy .......................................................... 57
  5.6.2 Transfer of a different habit of mind ............................................................ 58
5.7 Limitations and future directions ......................................................................... 60
5.8 Conclusions ........................................................................................................... 62

Works cited .................................................................................................................. 64
LIST OF TABLES

Table 2.1 Taxonomy of Near and Far Transfer .................................................................18
Table 3.1 Protocol and Prompts .......................................................................................34
Table 3.2 Global Scores for Quantity and Quality of Prompting ..............................35
Table 3.3 Challenges and Their Forms .........................................................................38
Table 4.1 Model Building for Reasons on Both Sides in the Three Outcomes ..........43
Table 4.2 Proportion of Students in Prompting Categories ..........................................45
Table 4.3 Number of Outcomes in which Students Considered Both Sides ..........46
Table 4.4 Average Number of Elaborated Reasons in Each Outcome .....................49
Table 4.5 Model Building for Elaboration of Reasons in Each Outcome ...............49
LIST OF FIGURES

Figure 4.1 Average Use of Connectives by Students and Teachers .......................41

Figure 4.2 Percentage of Students by Condition who Articulated Reasons on Both Sides of the Issue ...........................................................................................................42

Figure 4.3 Percentage of Students who Articulated Reasons on Both Sides, in one or two Outcomes ..................................................................................................................47
Chapter 1

INTRODUCTION

At the heart of a successful democracy is the citizens’ willingness to temper personal liberty with the needs of others in exchange for the safety and security provided by a government. The power of the governing body comes from the citizens’ belief that the common good is worth a degree of personal sacrifice. According to enlightenment thinking, the citizen’s ability to acknowledge the views of others is the cornerstone of a modern democracy.

Yet, when it comes to reasoning about topics that require balancing the needs of different stakeholders, people are notoriously bad at considering positions other than their own. John Dewey famously tackled this issue nearly a century ago arguing that uncritical privileging of one’s position is due to people’s tendency to get stuck in habits of thought that preclude the exploration of ideas necessary for reasoning about policy issues. Dewey did not use habits in a pejorative sense. Instead he thought of them as “the blood and sinews of learning” that develop naturally as a result of individual reasoning and societal influence on reasoning (Dewey, 1922, p. 334). And because habits are not so automatic as to be beyond introspection and therefore changeable, it is job of the educator to replace poor habits with habits of mind consistent with productive citizenship.

The current study is an investigation into the influence that education can have on individuals’ habits of mind involved in reasoning. I analyze how different instructional frameworks affect students’ reasoning across several tasks. Considering several tasks enables me to look for evidence of trans-situational reasoning heuristics—an operationalization of “habits of mind.” I will focus on whether students can avoid a
common fault in reasoning — *myside bias* — that is particularly detrimental to thinking about ill-structured and complex controversies like those that commonly arise in a democratic society.

Myside bias is the pernicious tendency to uncritically accept information and arguments consistent with one’s beliefs and ignore, distort, or reject information and arguments inconsistent with one’s beliefs (Perkins, 1985; Wolfe & Britt, 2008). Considering alternative positions that may be inconsistent with one’s beliefs is essential to scientific inquiry. The practice of simultaneously holding a position while examining counterevidence or different explanations is at the heart of many conceptions of scientific processes, such as falsification (Popper, 1959/2002) and statistical hypothesis testing (R.A. Fisher, 1922). Outside of science, considering both sides of issues will have a salutary effect on reasoning about practical affairs and public policy. In both scientific and everyday reasoning, the process of generating and elaborating contrary positions can minimize bias and improve the truth-value of conclusions.

Myside bias is a fault in reasoning, but it is not the same as a fallacy in formal logic where an improper inference will lead to a potentially false conclusion. However, failing to examine all sides of an issue may yield conclusions that would prove problematic for decision-making in a variety of domains. Because myside bias is so common, persistent, and fundamentally damaging to reasoning in multiple domains, (Baron, 1995; Kuhn, 1991; Perkins, 1985), classroom interventions have been designed to minimize the bias. Interventions may target a certain aspect of the reasoning process that is the likely origin of the bias, and for that reason it is helpful to review two theoretical accounts of the psychological basis of myside bias. This will be followed by
reviewing studies that have had success in targeting problematic aspects of the reasoning process. First, I discuss attempts to reduce myside bias by increasing knowledge about the topic. Next, I discuss attempts to reduce the bias by teaching students about the structure of a complete and sound argument and through experience with argumentation.
Chapter 2

LITERATURE REVIEW AND THEORETICAL CONTEXT

One potential source for myside bias is lack of knowledge about the different sides of an issue, which arises because people have a tendency to seek out information in support of beliefs, while eschewing information regarding alternative beliefs. As a result people start with a longer list of reasons in support of beliefs they hold than of alternative beliefs (Nickerson, 1998; Taber & Lodge, 2006; Wolfe & Britt, 2008). From this perspective, myside bias has its origins in a deficit in knowledge more than a difficulty in reasoning per se. Myside bias arises because available information is incomplete. Available information is incomplete because people favor information that confirms already held beliefs.

To counteract incomplete information, students are exposed either to information in support of an opposing position or inconsistent with their own position (Johnson & Johnson, 1985; Schwind & Buder, 2012). Confronting people with discrepant information may result in a better rounded picture of issues that avoids the “one-sided case building process” that otherwise leads people to reject incoming information that supports opposing views (Chinn & Brewer, 1993; Nickerson, 1998, p. 175).

More knowledge enables a more complete mental model which in turn may enable previously absent steps in the reasoning process such as integrating a counter position (Johnson-Laird, 2010). However, there is little empirical evidence that students will be able to integrate counter positions when they consider other issues. Helping students build a more complete representation of a particular problem is not without
merit, but whether doing so will produce a general reduction in myside bias remains to be seen.

Another potential source of myside bias is a reliance on a non-optimal argument schema. Counterarguments either are not considered or are suppressed in favor of other, overly simple reasoning heuristics such as “more reasons are better” (Baron, 1995; Wolfe & Britt, 2008). This explanation for myside bias suggests that people tend to eschew complex argument strategies such as rebutting counterclaims, integrating alternative positions, or proposing compromises when such strategies would be appropriate (Baron, 1991; Nussbaum & Schraw, 2007). According to the underdeveloped-schema explanation, the failure to include opposing views in an argument is due to the naïve conception that acknowledging other positions and counterclaims undermines the force of an argument. Indeed, think-aloud protocols with adults as well as children have shown that people cite the potential weakening of an argument as their reason for not mentioning counterclaims (Baron, 1995; Leitao, 2003).

There have been two general approaches to helping people elaborate and strengthen their argument schemas and we refer to these as rules-based and practice-based approaches. A rules-based approach attempts to build up understanding of what good reasoning entails through directly teaching the rules of reasoning or argumentation using some normative model such as the well-known model of the philosopher, Stephen Toulmin (1958). The rationale for a didactic approach is that having students organize their thinking using frames that force heuristics of good reasoning (e.g., attention to quality of evidence, proactive consideration of counter arguments) will facilitate positive
psychological processes such as self-explanation and metacognitive evaluation. That will enhance understanding of the content as well as the structure of an argument.

Direct instruction about forms of argument can positively influence science content learning (Berland & Reiser, 2009; McNeill, Harris, Hetziman, Lizotte, Sutherland & Krajcik, 2004), richness of scientific discussions (Osborne, Erduran & Simon, 2004), and reduce myside bias (Ferretti, Lewis, & Andrews-Weckerly, 2009; Nussbaum, 2005; Nussbaum & Kardash, 2007). However, we do not know whether students transfer what they have learned about the structure of an argument to subsequent reasoning about other topics.

The other approach to enabling students to elaborate their argument schemas is through the practice of argumentation. Students have repeated experience with argumentation with the expectation that this will enable them to learn how to reason through an inductive process. The idea is that reasoning heuristics can become internalized when their utility and flexibility is demonstrated through repeated use in varied contexts. Dialogic argumentation, whether the goal is consensus building or persuasion, usually involves confronting opinions different from one’s own, and the individual process of entertaining and integrating different opinions becomes a strategy that can then be used in other similar scenarios. With regard to minimizing myside bias, the hypothesis is that repeated experience with the discourse of disagreement and resolution leads to a generalized heuristic about the importance of considering both sides of an argument.

Evidence indicates positive transfer from interventions that provide the chance to participate in dialogic argumentation. In a series of studies, the Collaborative Reasoning
project (Reznitskya et al., 2009; Kim et al, 2011) has examined how myside bias is impacted through participation in peer-led groups that are encouraged to challenge each other’s ideas about ethical dilemmas faced by story characters. Experience defending viewpoints and challenging others leads to a greater likelihood that counterarguments will be considered in an argumentative writing task compared to a control condition (Dong et al., 2009; Kim et al, 2011; Reznitskaya et al., 2001). Students who engaged in argumentative discussions were likely to acquire and transfer argumentation strategies, such as considering counterarguments, to writing an argument about a topic they had not discussed. These studies provide compelling evidence that argumentative discussions can have an impact on reducing myside bias. Because the discussions all centered on ethical dilemmas, we can still question whether transfer effects would be seen if students were discussing other kinds of controversies that entailed different kinds of reasoning and evidence.

Because both rules-based and practice-based approaches have shown promise in reducing myside bias, some studies have combined the two methods. The expectation was that the two approaches would have an additive or even a synergistic effect in encouraging sound reasoning. The results, however, suggest that explicit instruction and argumentative practice together is less efficacious than just argumentative practice alone for reducing myside bias. For example, Kuhn and Udell (2003) compared two conditions. Students in both the treatment and control conditions engaged in a series of lessons in which they received explicit instruction about the form and function of arguments, and planned and discussed their own argument for or against capital punishment. Students in the treatment condition additionally engaged in discussion and other activities on how to
evaluate and rebut opposing sides as well took part in a debate with their classmates. Outcome measures came from an individual argument interview where students were asked their opinion on capital punishment. Students in the treatment condition showed more growth from pre- to posttest than students in the control condition. The primary improvement came from students entertaining and rebutting the opposing side. Kuhn and Udell claimed “exercise in argumentative discourse provided in the experimental condition…appear[s] necessary for greatest skill development” (p. 1256). This finding suggests that participation in argumentative discussions acts to reduce bias over and above the effect of being taught explicitly about good reasoning.

One limitation of the Kuhn and Udell (2003) study is that students were arguing about the same topic, capital punishment, throughout the intervention as well as on the written posttest. This makes it difficult to determine whether the reduction in myside bias was due to the experience with argumentative discussion or learning more about capital punishment. This is a methodological concern that the current study avoids by including transfer tasks that assess reasoning on issues the participants have not studied before.

Similar to Kuhn and Udell (2003), Reznitskaya, Anderson, and Kuo (2007) examined the combined effects of explicit instruction and practice in dialogic argumentation. An important difference between the studies was that Reznitskaya et al. compared students who received explicit instruction and engaged in peer-led argumentation to students who only participated in the argumentation. Recall that Kuhn and Udell provided all students with instruction, but only the students in one condition had the opportunity to practice argumentation. In Reznitskaya et al. all students had practice with argumentation while only some received explicit instruction about argument
structure. Surprisingly, the discussion-only condition outperformed the students who received the explicit instruction in argument structure suggesting that the explicit instruction had a deleterious effect on student reasoning and that practice may be the critical component for improving arguments.

Although not an investigation of myside bias, another study by Kuhn highlights the positive effect of practicing a reasoning heuristic and the null impact of explicit instruction. Investigating the acquisition of the control of variables strategy, Dean and Kuhn (2006) compared students who received instruction on the strategy and practiced the strategy with students who only practiced the strategy without instruction. Initially the group that received instruction and practiced the strategy outperformed the practice-only group, as well as a third group that only received explicit instruction, but after a week the practice-only group was performing as well as the instruction-and-practice group both on tasks they had seen before and on novel tasks. Both groups outperformed students who received only direct instruction.

Although Dean and Kuhn (2006) was not an investigation of myside bias, their finding combined with those of Kuhn and Udell (2003) and Reznitskaya et al. (2007) at the very least suggest that direct instruction is not a necessary component in helping students gain proficiency with a reasoning strategy. At best there may be an immediate advantage for including explicit instruction but this effect appears to drop off (Dean & Kuhn, 2006), and at worst the instruction can inhibit advantages otherwise gained from practicing reasoning with peers (Reznitskaya et al., 2007). Taken together, these findings support the conclusion that the driving force behind continued and flexible use of reasoning heuristics is engaging in reasoning processes in situ.
In summary, attempts to overcome myside bias fall into three categories (1) expanding knowledge about the topic, (2) explicit instruction in strategies of reasoning, (3) experience in arguing back and forth about a topic of mutual concern. Evidence reviewed here suggests that providing opportunities to confront, debate, and persuade others reduces myside bias in subsequent argumentation, both spoken (Felton, 2004) and written (Dong et al., 2008; Reznitskaya et al., 2007, Kim et al., 2011). Enabling argumentation among students minimizes myside bias more effectively than explicit instruction about argument forms or reminders about incorporating counterarguments. There is less data on the pervasive effects of the knowledge-based approach.

From this summary we arrive at two noticeable gaps in research on myside bias. First, few studies have pursued an examination of transfer to ascertain the impact that interventions have on robust and widely-applied habits of mind. Even studies that have included transfer have used transfer from one to another ethical dilemma (e.g., Reznitskaya et al., 2009) rather than transfer to controversies in a different domain. The current study addresses this gap by evaluating transfer across domains. Second, no studies have compared a knowledge-enrichment approach to a practice-based approach. The current study addresses this gap by examining the comparative merits of a practice-based approach versus a knowledge-enrichment approach. Moreover, with the inclusion of three open-ended reasoning tasks, the present study can test transfer of reasoning strategies across multiple domains. Note the absence of any explicit instruction about argument elements or argument form; this will allow for a test of whether practice without instruction can lead to successful reasoning and transfer.
To recap: the current study is an investigation of two different approaches for minimizing myside bias within the context of learning about a socio-scientific controversy. There are two treatment conditions, students in the first condition learn about the controversy through teacher-directed instruction and students in the second condition worked in peer-led groups to learn about the same controversy. In the following section, I will elaborate on the theoretical grounding for my conception of how an individual’s reasoning is likely shaped by this disparate experience with argumentative practice in the classroom and, in particular, how these experiences will likely impact myside bias on transfer tasks.

2.1 Classroom discourse and argumentation

This study is concerned with altering students’ reasoning processes. I am interested in whether providing students the opportunity to reason jointly with others changes the likelihood that they will consider positions and arguments different from their own. I am also interested in whether participation in discourse that involves argumentation in many forms (e.g. persuasion, consensus building) will have a more lasting effect on reducing myside bias than teaching students information that supports arguments on different sides of an issue. To inform research into these questions, I draw from sociocultural views on acculturating students into ways of talking and thinking in the classroom as well as theories of transfer to account for why students might apply these strategies of reasoning inter- and intra-personally.

2.2 Classroom learning and norms of discourse

Learning in a classroom is a predominantly social act as the acquisition of knowledge occurs within and through a community of other learners. A student must
learn to move between the different contexts to become a functioning part of the classroom community (Gee, 1996). Becoming accustomed to these contexts is facilitated by the fact that communication within different contexts tends to coalesce around certain norms, and people are generally adept at inducing the norms and enacting expected processes in order to establish their role in the community (Dyson, 1993; Lave & Wenger, 1991). Research on appropriation of classroom practice supports my hypothesis that students in both treatment conditions will tend towards communicative norms that are consistent with their respective instructional frameworks. The norms of discourse in certain instructional settings can support – or undermine – knowledge-enrichment and argumentative practice as means of reducing myside bias.

Like other discourse processes in the classroom, argumentation is not a singular process but a frame for communication that can be shaped by broader communicative norms and different contextual variables in a classroom (Berland & Hammer, 2012). Because argumentation is not a dominant discourse in the classroom (Applebee, Langer, Nystrand, & Gamora, 2003; Cazden, 1988; Mehan, 1979), the understanding that students in this study develop about argumentative discourse will be shaped by how they perceive argumentation in relation to other frames for interaction that are established in the classroom. Because novel or infrequent communicative frames, such as argumentation, can be altered by pre-existing norms of interaction, I need to specifically consider what sort of argumentative practice will likely take place in the two treatment conditions in the current study.

2.3 Argumentation in classroom discourse
I must briefly characterize what is meant by argumentation. Zarefksy (1995) offers a modern definition of argumentation as the process of “justifying decisions under conditions of uncertainty.” Argumentation is an embedded practice where the goal is to justify, as opposed to prove, one’s position on issues for which there is no determinate “right” answer (van Eemeren, 2001; Voss & Van Dyke, 2001; Zarefksy, 1995). The question then becomes, how does one justify a position? According to the New Rhetoric, criteria for justification are determined by the audience or someone else who has a stake in the outcome of the argument (Perleman & Olbrechts-Tyteca, 1969). The function of the audience varies within different argumentative schemes and so successful argumentation relies on correctly framing the argument and understanding what strategies will work best for the audience at hand (Walton, Reed, & Macagno, 2008). This modern conception of argumentation helps to highlight the role of audience and personal agency in argumentative practice, and these concepts help to illuminate differences in the role of argumentation in collaborative learning and direct instruction.

In teacher-led discussions it is atypical for students to be offered the opportunity, either explicitly or implicitly, to engage in authentic dialogue where topic, talking turns, and accountability are negotiated and managed equally by all parties. Researchers have shown that this is the case for teacher-led argumentation as well. Moves the teacher makes to enact the role of moral and content authority prevent students from engaging in discourse moves that are critical to dialogic argumentation.

With the teacher as leader, argumentation as an activity frame is subsumed by established discourse and power structures in the classroom. Structures in teacher-led instruction tend to limit the agency of the students and impact the students’ conception of
audience. When a discussion is continuously tethered to a teacher, as is still almost universal in US classrooms (Cazden & Beck, 2003), students are usually addressing their teacher, rather than each other. And so the teacher becomes the default audience for everything said in a discussion (Hand et al, 2007).

Two virtues of direct teacher-led instruction are that the concepts can be thoroughly covered and that students can receive expert explanations of the concepts. This implies that direct instruction could be an effective means of increasing student knowledge about the topic and, according to hypothesis, thereby reducing myside bias. It seems unlikely, however, that teacher-led discussion will result in authentic argumentative practice for the students.

In peer-led discussions, the students are sharing control of the topic and managing talking turns themselves (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009). A democratic interactional structure has been shown to increase integral aspects of argumentative practice, such as analogical reasoning (Lin et al., 2012), explanation of thinking (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, De Leeuw, Chiu, & LaVancher, 1994), requests for positions and reasons in support of the positions (Anderson et al., 2001), and acknowledging confusion (Mercer & Littleton, 2007).

Because of the social accountability involved in peer management of discussions, individuals have the opportunity to calibrate their positions in relation to other group members. Calibration can involve social roles, including turn and topic management and such functions as encouraging less vocal participants, as well as clarification of concepts and evaluation of the credibility of arguments (Li et al., 2007). In contrast to teacher-led instruction, the role of the individual in relation to the goals of the groups are likely to be
more explicitly negotiated and should be more flexible because the norms of discussion without a teacher are open to interpretation. There is less precedent for how students should discuss concepts when the teacher is not assuming the usual role of authority figure. Flexibility of roles is helpful for fostering argumentation because, without a default authority, the veracity and soundness of claims must be negotiated instead of determined through validation by the teacher.

This characterization of the two instructional frameworks in the current study leads to the prediction that collaborative groups of students will engage in more argumentative practice than students in teacher-led instruction, although students in the teacher-led condition did receive direct instruction about the various sides of the socio-scientific controversy studied about by all students. Both teacher-led instruction to increase knowledge of the topic and the peer-directed discussions to practice argumentation could reduce myside bias.

The study included several reasoning tasks to investigate whether the reasoning processes that reduced bias about one issue transferred to reasoning about different issues. A finding of transfer of reasoning would suggest a general disposition towards reasoning and decision-making consistent with ideals outlined by Dewey and other educational philosophers. Because this study is looking at the transfer of reasoning across multiple tasks, it is useful to consider how theories of transfer can help motivate predictions for the study.

2.4 Transfer across contexts
The educational ideal of thoughtful ‘habits of mind’ requires the flexible application of some set of reasoning strategies across different contexts. The nature of such a schema for reasoning is under investigation in this study.

One area of study that contributes to understanding transfer is analogical reasoning. Analogical reasoning can be seen as a general case of transfer, where transfer is the application of something learned, the base, to better understand something less known, the target, through recognizing the similarities between the target and the base. Some examples include children using their understanding of goldfish care to draw inferences about the care of other animals (Hatano & Inagaki, 1992) or using a story about splitting a force of soldiers to solve the problem of how to radiate a malignant tumor without using a single beam of radiation that could be toxic (Duncker, 1945; Gick & Holyoak, 1983). To transfer appropriate knowledge from the base to the target, individuals must frame the task in a way that selects pertinent features of prior knowledge. Inconsequential similarities between base and target must be ignored in favor of deeper structural features (Gentner, 1983). Perspicuous framing of a task increases the chance that relevant knowledge will be transferred. From this standpoint, building up background knowledge is necessary but not sufficient for transfer. Instead the emphasis should be on helping students effectively frame novel tasks.

In the case of reasoning, it is helpful if students have relevant background knowledge about argument structure or situations but it is more important that they effectively frame the task based on features of the context including their personal goals and the goals of others. One of the important aspects of the context is the audience for whom the argument is meant. Framing an argument begins with identifying the audience
and then transferring in past knowledge and reasoning strategies that will help achieve the goals in light of the audience.

The importance of framing is discussed by philosophers with a dispositional view of thinking who suggest that errors in reasoning are not due to lack of ability or know-how but instead a failure to use available knowledge (Perkins & Ritchhart, 2004). Improving reasoning is seen as a matter of increasing sensitivity to the expectations of the people who are invested in the reasoning process and the outcome. The best way to improve sensitivity to context may be to provide students with the freedom to explore ideas and the opportunity to “contribute to larger conversations that extend across time, people, places and activities” (p. 605, Engle, Nguyen, & Mendelson, 2011; Hatano & Inagaki, 1992; 1996; Mercier & Higgins, 2013).

The current study is testing the idea that providing students with a certain degree of freedom to argue and explore their own thinking and opinions, will increase their sensitivity to each other’s ideas and that this elaborated conception of an audience will minimize myside bias.

The concepts of near and far transfer are useful to characterize the outcome measures of this study. Near and far are conceptualized as lying on a continuum. Whether a transfer task is near or far depends on the overlap between the initial learning context and the transfer context. Barnett and Ceci (2002) outlined the varying ways that the two learning events could overlap. Their continuum is included below in Table 2.1.
Table 2.1  
*Taxonomy of near and far transfer*

<table>
<thead>
<tr>
<th>Knowledge domain</th>
<th>Mouse vs. rat</th>
<th>Biology vs. botany</th>
<th>Biology vs. economics</th>
<th>Science vs. history</th>
<th>Science vs. art</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical context</td>
<td>Same room at school</td>
<td>Different room at school</td>
<td>School vs. research lab</td>
<td>School vs. home</td>
<td>School vs. the beach</td>
</tr>
<tr>
<td>Temporal context</td>
<td>Same session</td>
<td>Next day</td>
<td>Weeks later</td>
<td>Months later</td>
<td>Years later</td>
</tr>
<tr>
<td>Functional context</td>
<td>Both clearly academic</td>
<td>Both academic but one non-evaluative</td>
<td>Academic vs. filling in tax forms</td>
<td>Academic vs. informal questionnaire</td>
<td>Academic vs. at play</td>
</tr>
<tr>
<td>Social context</td>
<td>Both individual</td>
<td>Individual vs. pair</td>
<td>Individual vs. small group</td>
<td>Individual vs. large group</td>
<td>Individual vs. society</td>
</tr>
<tr>
<td>Modality</td>
<td>Both written, same format</td>
<td>Both written, multiple choice vs. essay</td>
<td>Book learning vs. oral exam</td>
<td>Lecture vs. wine tasting</td>
<td>Lecture vs. wood carving</td>
</tr>
</tbody>
</table>

The tasks in the current study can be categorized as relatively nearer and further from the initial learning context. From the history of educational research near transfer is easier to obtain and much more likely than far transfer. Thus, conceptual “distance” can provide a metric for describing the robustness of the reasoning strategies acquired during the unit. A reduction in myside bias on a far transfer reasoning task can be held up as good evidence that myside bias was substantially reduced.

2.5 Current study and predictions

In the current study, students were asked to play the role of officials at a wolf management agency to decide whether or not a town should be allowed to eradicate a wolf pack. Classrooms were assigned to either a direct instruction (DI) condition or
collaborative group (CG) condition. In the DI condition, teachers led students in readings and activities that delineated the environmental and economic impact wolves—or the absence of wolves—would have on the town. In the CG condition, students worked in peer-led groups and completed the readings and activities with minimal assistance from the teacher. Importantly, neither the teachers, activities, nor readings provided any explicit instruction about argument function or form.

At the conclusion of the unit, students were asked to write a letter detailing their position on the wolf controversy (hereafter, called the wolf letter). As a near transfer task, they were interviewed about their opinion on whaling after hearing a brief essay about the controversy (hereafter, whale interview). Finally, as a far transfer task, they wrote about how they would handle an ethical dilemma introduced in a story *The Pinewood Derby* about whether a boy should tell on a classmate who cheated in a competition building and racing model cars (hereafter, pinewood essay).

The study design allows for the comparison of a knowledge-enrichment approach with an argument practice approach. Over the course of the unit, students in the DI condition should develop an elaborated understanding of the wolf controversy and all of the various stakeholders. This knowledge should bolster their reasoning pertaining to the decision of whether or not to kill the wolves. In particular we would expect a lower percentage of students to display myside bias when compared to students who did not study the wolf unit (control condition). On two subsequent reasoning tasks, however, we are less confident that expanded knowledge relevant to wolf management will influence reasoning about whaling or a model car competition. If the student is able to frame the whaling issue as similar to the wolf issue then there may be a reduction of myside bias.
due to the student characterizing both issues as complex and therefore requiring a consideration of multiple points. For the pinewood essay, there is minimal conceptual overlap with the wolf unit, and so we would expect very little transfer of reasoning across the two tasks.

Students in the CG condition will teach themselves about the wolf controversy through peer-led discussions. Because of the nature of the wolf controversy and design of the materials, as well as past research (Waggoner et al., 1995; Anderson et al., 2001), we expect that students will engage in dialogic argumentation during these discussions. Because there is not dominant discourse structure with which children are familiar for small-group discussions, argumentation will involve a more authentic exchange of positions and reasoning than in classroom discussions led by the teacher. The result should be frequent experience arguing with different group members who support various positions on the issue. For some students, these encounters with multiple addressees will potentially coalesce around a heuristic that involves comparing arguments to their own, or weighing both sides of the issue. Because this heuristic is undergirded by multiple experiences it is likely to be transferred to independent tasks because cues in the target scenario are likely to resemble previous argumentative practice, which in turn helps the student frame the current task as argumentative and requiring skills learned in their groups. For relatively far transfer tasks, we would expect the absence of any content or knowledge-based cues to reduce the number of students consider both sides of an issue, but we would assume that there will be a greater percentage of students in CG who do weigh both sides than in the DI condition.
Chapter 3

METHODS

3.1 Participants

Two Midwestern districts and 36 fifth-grade teachers agreed to implement the Wolf Management Unit in their classrooms. The study design was quasi-experimental with triples of classroom matched on demographics, and then randomly assigned to one of three conditions: Direct Instruction (DI), Collaborative Group Work (CG), and Wait-Listed Control (CTL). Half of the participating classes (6 CG, 6 DI, 6 CTL) were from three schools serving low-income African-American families. The percentage of students eligible for free and reduced-price lunch in these schools ranged from 80%-98%. The other half of the participating classes were from five schools in a district located in a major metropolitan area serving predominantly low-income Hispanic-American families. The percentage of students eligible for free and reduced-price lunch ranged from 80%-86%. Overall, 45.7% of the participants were Hispanic American, 41.5% African American, 9.4% European American, and 3.5% Other.

Prior to any instruction all students with both parental consent and individual assent completed a handful of assessments including a reading comprehension assessment (MacGinitie, MacGinitie, Maria, & Dreyer, 2000), a measure of basic English language proficiency (timed naming of common objects, Snodgrass & Vanderwart, 1980), a vocabulary assessment, a questionnaire about peer relationships, and a questionnaire about attitude towards school subjects.

Initially, 904 students were recruited to participate. Consent and assent was obtained for 790 students. Beside the students for whom we did not receive consent and
assent, another 24 students were not included in any analyses for several reasons. The most common of these was they were absent on the day of the pretests and/or outcome measures. There was no statistically significant difference between the reading scores ($t(788) = 0.38, p > .10$) or oral fluency measures ($t(788) = 1.81, p > .05$) of the 24 students whose data we did not include and the remaining 766 students.

There are three outcomes measures reported on here. For reasons that are explained below, not every student completed every outcome measure. There were 340 students who received all three outcome measures, and 766 who completed at least one of the outcome measures. A two-way ANOVA showed that neither the condition a student was assigned to, nor whether they received all three measures was associated with higher reading comprehension scores, $F(5, 761) = 0.96, p > .10$) or speeded naming, $F(5, 761) = 0.73, p > .10$). In other words, there were no reliable differences in pretest measures between the students who completed all three outcomes and those who completed one or two.

### 3.2 Participating Teachers

Before the intervention, the teachers completed a survey about teaching experience and instructional practices. Most teachers had taught for at least three years but three teachers had taught two or fewer years at the fifth-grade level. The one brand new teacher had two support teachers who co-taught and helped manage the classroom. The other teachers had either no support teacher or a rotating support teacher shared among all teachers at that grade level. Three teachers assigned to DI and two assigned to CG had a master’s degree while four of the teachers reported that they were working towards a master’s degree.
The survey included a section on allocation of instructional time. Teachers reported spending most minutes per day on language arts and math (M=82.0, SD=33.4), with social studies and science receiving less than half an hour a day combined. There were no significant differences between the two conditions in reported pre-intervention time allocations in any subject area. Teachers were also asked to divide total time into time spent using five different instructional strategies: Whole class instruction, group projects, group discussions, individual seatwork. In all subjects, teachers reported spending most time (56% of total time) in whole class instruction. There were no significant differences between conditions for any of the reported strategies of instruction.

3.3 Wolf Management Unit

3.3.1 Teacher training. Teachers assigned to the DI and CG conditions participated in a two-day workshop pertaining to their assigned instructional condition. The first day of the workshop was dedicated to the specifics of the instructional method. Those participating in the collaborative group work (CG) condition were provided with: (a) general theory behind the importance of student-led group work, (b) instructional moves that would facilitate group work, (c) a primer on the discussion method Collaborative Reasoning, and (d) videos of exemplary student-managed discussions and teacher scaffolding moves to support these discussions. Teachers participating in the Direct Instruction (DI) condition were provided with: (a) general theory and history of DI, (b) questioning strategies that facilitate comprehension and participation in whole-class instruction, (c) strategies for classroom management through engagement, (d) videos of exemplary whole-class direct instruction. The second day focused on the
particulars of the curriculum for all teachers. Teachers remained segregated for the second day in order to best explain aspects of the research design that were specific to each instructional condition (e.g. jigsaw design in CG). Only students who were assigned to classrooms in the DI and CG condition received the wolf unit prior to completing the outcome measures; students in the wait-listed control completed the wolf unit following data collection. Control teachers were also provided with professional development similar to that given to the CG teachers.

3.3.2 The curriculum. The curriculum was *The Wolf Management Unit*. The DI and CG classes worked through the unit for six weeks. Unless noted, the following description of the unit pertains to the materials and sequencing for both conditions. In the scenario, students were asked to role-play being officials in a fictitious Wolf Management Agency. The agency had been contacted by the town of Winona, asking for permission to hire professional hunters to kill a pack of wolves that many citizens regarded as a problem. The students were told that at the end of the unit they would be asked to write the town with their decision about whether or not they, as officials at the Wolf Management Agency, would allow Winona to hire hunters to kill the wolves.

The unit was divided into three sections which introduced the students to Winona’s ecosystem, Winona’s economy, and Winona as a community of citizens. The sections all introduced how wolves would impact each aspect of the town. The three sections were comprised of an information booklet and an activity booklet. The information booklet contained readings specific to the topic and the activity booklet contained a variety of tasks that were to accompany portions of the information booklet. A brief example is illustrative: In the Ecosystem packet students read about a food web,
read about how the presence of wolves impacted that food web, and they then completed an activity where they had to illustrate a food web and draw the connections between predators and prey. Activities were designed to reinforce concepts learned in the reading as well as challenge the students with extension questions. Many of these challenge questions required students to explain their reasoning. Students in the DI condition completed all three pairs of information and activity booklets. Students in the CG condition were split into heterogeneous groups, each designed to be a cross section of the class in terms of reading ability, talkativeness, ethnicity, gender, and attitude towards learning. Each CG group was then charged with becoming experts in one of the unit’s three sections. Each CG classroom had at least one ecosystem expert group, economy expert group and public policy expert group. Each of these expert groups read the information booklet and completed the activity booklet on its topic. Then, each group prepared a poster on its topic — ecosystem, economy, public policy — and made a presentation to the whole class. Students were provided with material for their poster presentations in order to standardize the concepts that were presented in the different CG classrooms.

Before the Wolf Management Unit, students in the CG condition had three literature discussions in the Collaborative Reasoning format (see Waggoner et al., 1995, for details). The students read a story and discussed a big question that related to an ethical and practical dilemma the protagonist was facing. Students were expected to discuss their opinions with one another. When students disagreed with the reasoning of another group member they were encouraged to respectfully challenge each other’s ideas without offending the individual. Collaborative Reasoning involves open participation,
meaning that students speak freely without being nominated by the teacher or a student moderator. Students in the DI condition read the stories, but did not hold collaborative discussions.

Students in the CG condition also held two Collaborative Reasoning discussions during the Wolf Management Unit about whether Winona should be allowed to hire hunters to kill the wolf pack. The first discussion was held at the beginning of the unit when the students were first presented with Winona’s request. The second discussion was held near the end of the unit after the students had spent a number of weeks gathering evidence and shaping their opinions. The teachers in the DI condition also held initial and final whole class discussions, in order to introduce the topic and discuss final thoughts, respectively.

Throughout the project a research assistant (RA) was a participant observer in each classroom shooting video as well as supporting the implementation of the unit. Each RA was assigned two treatment classrooms such that they could attend every lesson in the Wolf Management Unit for each condition. The RA was in daily contact with the teacher through phone, email, and pre- and post-lesson conferences. This level of involvement allowed the RA to facilitate the intervention in a number of ways, including but not limited to: highlighting key concepts for instructional focus, facilitating group work (CG condition), providing one on one help (DI and CG condition), providing reminders about supplies needed for activities, answering questions a teacher had about the unit or procedures for the study. The RA also helped maintain experimental validity by ensuring that teachers in the DI condition were not utilizing any group-based collaborative frameworks. Weekly meeting between RAs and the study’s principal
investigator further ensured a consistent implementation within and between treatment conditions.

3.4 Assessments

At the end of the unit students were administered various outcome measures, three of which are discussed here, and were designed to test the transfer of reasoning strategies. All outcomes listed below were completed on different days.

3.4.1 Wolf decision letter. At the end of the unit, students individually wrote a decision letter about whether they would allow Winona to hire professional hunters to kill the wolves. Students were given 45 minutes to write, with most finishing well before the time was up. Students in the control class during the first wave of data collection did not write the decision letter due to an oversight in the research design. Control students wrote the letter the next year based on a brief summary of the wolf issue. Because the control students had no context for the questions, each RA provided a short summary of the controversy before handing out the prompts. All student letters were transcribed by professional transcribers and research assistants and reviewed by research assistants. There were minimal discrepancies found through review. Transcriptions were corrected for spelling but maintained the student’s initial syntax.

3.4.2 Whale controversy oral interview. The interview was administered to a fraction of the participating students; around sixty percent. Each classroom had a target sample that was a heterogeneous group representative of the class on a number of characteristics including: Race, Gender, general cognitive abilities, talkativeness, and leadership. The groups were created collaboratively through discussion between the RA and the classroom teacher, in order to limit bias and ensure the group was representative.
of the class. A target sample was created for each classroom in CG, DI and control conditions. If there was extra time after the RA’s interviewed the target sample then the RA’s interviewed other students from the class that were selected from a randomly generated list of the remaining students. The random numbers were generated prior to the onsite interview using SAS.

The interviews consisted of the RA reading the students a 383 word essay about the pros and cons of whaling, and then asking the student whether or not whaling should be allowed. Students were given their own copy of the essay with which they could follow along. The interviews were done outside the classroom with only the RA and student present. All interviewers followed a protocol which prescribed the order and exact wording of prompts to elicit particular elements of argument. The protocol was essentially divided into three prompts, a prompt intended to elicit the student’s position, a prompt for counterarguments and a prompt for rebuttals. Table 3.1 gives a brief summary of the protocol and prompts for individual arguments. All research assistants administering the interview had been introduced to the students at some point in the intervention and so there should have been little variation due to familiarity.

The interview was recorded using a Sony digital voice recorder.

3.4.4 Pinewood transfer essay

All students individually completed a reflective writing assignment on a story entitled, The Pinewood Derby. In the story, two main characters, Jack and Thomas, are participating in their schools pinewood derby race. All students are instructed to complete their pinewood racer without any help. Thomas, a poor and unpopular kid, receives help from his brother and eventually wins the race. Jack, Thomas’ only friend, finds out that
Thomas’ brother helped make Thomas’ car. Jack isn’t sure whether he should tell the teacher about the cheating because he is glad that Thomas has finally won something. The students were asked to make a decision about whether Jack should tell on Thomas. The students had 45 minutes to plan and write the essay. Most of them finished before the allotted time.

3.5. Analysis of outcome measures

All outcome measures were transcribed into Word, and subjected to the following coding protocol. The two handwritten essays were transcribed with spelling corrected to facilitate computer-assisted coding. The interviews were transcribed by professional transcribers and research assistants and reviewed by research assistants. There were minimal discrepancies found through review. The transcriptions included SALT conventions to capture the idiosyncrasies and errors in the student’s natural language production. For clarity I will use the term “text” to apply to all outcomes.

Coding proceeded in two major phases. First, the texts were divided into communication units (Loban, 1966). A communication unit (C-unit) is defined as a main clause and its subordinate clauses. We chose C-units as opposed to the similar T-unit (Crooks, 1990), because the rules of C-unit division are less syntactically driven and allow for rule-based division of colloquial and substandard writing. C-units enabled coding of sentence fragments that contained content relevant to an argument, but had no verb (e.g. when rebutting an opposing position, “They wrong”). The essays and interview transcripts were divided into C-units by three raters working independently with satisfactory reliability.
The second phase involved the actual coding of texts for aspects of argumentation. This was a multi-step process involving various codes.

3.5.1 Coding. Three categories of codes were created to measure different aspects of argumentation. No extant coding scheme was expansive enough for our current analyses and so we created one that accounted for common elements of argument, content of the reasoning, and semantic markers indicating rhetorical moves.

Any holistic measures of quality were established based on different combinations of the available codes. This is different from other studies that have established quality through preconceived rubrics (e.g., Osbourne, et al., 2004). Basing statements about overall quality on individual codes allowed for greater flexibility and circumvented the problem of establishing a priori assumptions about quality. This would have been particularly problematic given that the subjects are all from underserved school districts and, as such, possess below grade-level literacy skills.

3.5.2 Elements of argument. For all three outcome measures, the coding was completed by two coders working independently following a training session. I coded the majority of all texts with different reviewers providing reliability. There were two coders that helped establish interrater reliability, one was a fellow graduate student and the other was a classroom teacher of twenty years. Both of these coders worked closely to initially fine-tune the codes based on pilot data or data from previous studies.

For the wolf letters, two coders each coded half of the letters. A selection of each raters’ coded letters was also coded by the other rater resulting in 30% of the letters being coded twice and these were used for reliability measures. For the wolf letters interrater reliability was good (Cohen’s K=.89). For the whale interview and the pinewood derby I
coded the majority of the texts with a second coder coding 30% to establish reliability. For the pinewood essays reliability was satisfactory (Cohen’s K=.84). For the whale interview, reliability for the individual argument codes was satisfactory (Cohen’s K=.83) and reliability for the holistic codes for amount of prompting was high (Cohen’s K=.91).

First, the texts were coded for the student’s overall position as, in favor of the killing of the wolves/whales or telling on Thomas (pro) or against the killing or tattling (con). If the students weighed both sides of the issue without having stated a preference for one side the overall positions was coded as “both sides.” Overall position was only coded after the reviewer had read the essay at least once. Texts were set aside if neither of the two coders could understand what the writer was saying. This was the case for less than 3% for all outcome measures. A “third way” code was used for statements that held the main question in abeyance by proposing a different way to handle the problem, as in the following examples:

…because they can take them somewhere else and put a fence around them

…Jack could have someone else tell the teacher so he could still be friends with Thomas.

The vast majority of statements that were coded as “third way” came in essays with a stated position and additional reasons in support of that position.

As well as codes for overall position, there were four broad code categories: statement of position, support for that position, consideration of a counter position, and response to a counter position. These categories account for most of the codes in the texts. A student’s stated position was coded as such. C-units that were reasons in support of a stated position were coded as elaborated or unelaborated. Elaborations were
propositions that included information increasing the relevance or further justifying an explicitly stated reason. Elaborations most often took the form of (a) explanations that were propositions that provided causal chains and/or details for why a circumstance envisioned in a reason might come about and (b) evidence that pointed to a fact or authority in order to increase the credibility of the argument. Unelaborated reasons were discrete premises that were in support of the stance, but were not developed beyond the statement of the premise. Below are examples of two elaborated and an unelaborated reason. For the elaborated examples, the first elaboration further explains the idea that wolves are bad for the ecosystem; the second cites an authority as support.

Elaborated reasons:

Wolves should be hunted (position). / Wolves are bad for the ecosystem (reason) / wolves are bad because they eat elk and the elk disappear (elaboration).

Wolves should not be hunted (position). / The wolf is important to the ecosystem (reason) / The wolf biologist said that (evidence)

Unelaborated reason:

Wolves should be hunted (position). / They eat elk (reason) /

Counterarguments and rebuttals to counterarguments were also coded. Propositions that contained a reason that ran counter to the stated position were coded as counterarguments. Context of use was evaluated to ensure that the student was considering a counterargument (as opposed to entertaining two equally held positions). Reasons that ran counter to the stated position but could not be conservatively assessed as counterarguments were marked as indeterminate. A similar protocol was adopted for rebuttals.
Counterargument:

I think that they shouldn’t kill the wolves…but I know that they keep attacking other animals like sheep, moose, cows and different animals like in the passage it said that a wild wolf killed Elmo [dog reported as missing in newspaper].

Indeterminate:

No, I don’t think the wolves should be killed/ the wolves are just doing what they do best/ also the people of Winona are upset/ the wolves are killing elk.

Similar to how reasons were coded as elaborated or not, counterarguments were also categorized in terms of whether they contained an elaborated or unelaborated explanation of the counterargument. It was also noted when students were entertaining arguments counter to their own, and when they were weighing both sides equally. Likewise, rebuttals were coded as elaborated or unelaborated. Finally, instead of rebutting the counterargument to reinforce an already held position, a few students were coded as making a concession to the counterargument. It was also noted when students used a unique reason as a rebuttal to a counterargument and when they simply reiterated a previously stated reason as a rebuttal.

3.5.3 Additional coding for oral interview task

As mentioned above, the interview protocol called for the interviewer to prompt the student if they had finished speaking without giving a fully realized argument (e.g. reason, counterargument and rebuttal). The three key prompts one for, reason “What do you think, should we allow people to hunt whales?” counterargument “Do you think there could be people who disagree with you?”, and response to counterargument “What would you say to those people?” Each of those prompts had an ancillary prompt to be used if the child was not explicit in their reasoning (see Table 3.1 provided directly
below. The transcripts were also coded for quality and quantity of prompts. Each of the three main sections of the protocol received a code which corresponded to the amount of prompting the RA gave. Table 3.2 summarizes the level of prompting and the appropriate code.

Table 3.1
Protocol and prompts

<table>
<thead>
<tr>
<th>Initial prompt</th>
<th>If no reasons given</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For Position</strong></td>
<td></td>
</tr>
<tr>
<td>What do you think [child's name] should people be allowed to hunt whales?</td>
<td><strong>What are your reasons for thinking that?</strong></td>
</tr>
<tr>
<td><strong>For</strong></td>
<td></td>
</tr>
<tr>
<td>Counterargument</td>
<td></td>
</tr>
<tr>
<td>Do you think there might be people who disagree with you?</td>
<td><strong>What could be some reasons that they would disagree with you?</strong></td>
</tr>
<tr>
<td><strong>For Response</strong></td>
<td></td>
</tr>
<tr>
<td>What would you say to those people who disagree with you?</td>
<td><strong>What are your reasons for disagreeing with them?</strong></td>
</tr>
</tbody>
</table>
Table 3.2
Global scores for quantity and quality of prompting

<table>
<thead>
<tr>
<th>Protocol Section</th>
<th>Code</th>
<th>Level of prompting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompt for Reason</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>RA uses prompt from protocol</td>
<td></td>
</tr>
<tr>
<td>Elaborate</td>
<td>RA explicitly asks for reasons</td>
<td></td>
</tr>
<tr>
<td>Considering the counter argument</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneous</td>
<td>Child considers CA prior to prompt</td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>RA uses prompt from protocol</td>
<td></td>
</tr>
<tr>
<td>Elaborate</td>
<td>RA explicitly asks for reasons for CA</td>
<td></td>
</tr>
<tr>
<td>Both positions</td>
<td>Child spontaneously provides reason for both sides</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Child never considers CAs even after prompt</td>
<td></td>
</tr>
<tr>
<td>No prompt</td>
<td>RA doesn't prompt</td>
<td></td>
</tr>
<tr>
<td>Response to counter argument</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneous</td>
<td>Child responds to CAs prior to prompt</td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>RA uses prompt from protocol</td>
<td></td>
</tr>
<tr>
<td>Elaborate</td>
<td>RA explicitly how the child would respond to CA</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Child never rebuts CA even after prompt</td>
<td></td>
</tr>
<tr>
<td>No prompt</td>
<td>RA doesn't prompt</td>
<td></td>
</tr>
</tbody>
</table>

CA = Counter argument RA = Research Assistant

3.5.4 Content of Reasoning. Using a bottom-up process, the content of the student’s reasoning was catalogued. These codes could then be categorized based on conceptual overlap. For example, both of the following reasons included in two wolf letters could be grouped under the conceptual label, *Balance of the foodweb*:

We shouldn’t kill the wolves because *if they die then the elk will eat too much trees.*

…if the wolves go away then *animal like the eagle won’t have leftovers to eat for their meal*

Because of the conceptual similarity between the wolf letter and the whale interview, many of the reasoning categories from the wolf letter could be applied to the
whale interview. For example, this reason could also coded as pertaining to the balance of a foodweb.

The whales eat so many fish *that they’ll be none left for other predators like us.*

Reasoning in the pinewood essay most commonly dealt with fairness and empathy and was less overtly similar to reasons used in the other two measures.

Jack shouldn’t tell on Thomas because *he doesn’t have any friends* (empathy)

3.6 Transcription and coding of classroom talk

Given the sheer volume of video recording, a principled sampling of classroom talk was analyzed. Six 4-minute excerpts were selected for each DI and CG class, giving a total sample of 144 excerpts. Control classes were not video-taped. The goal of the sampling was to maximize the probability that we analyzed time on task. We also wanted to eliminate potential bias by allowing the selection to be partially random. One excerpt was taken from each week of the unit. The majority of the excerpts were taken from Tuesdays because it was the least disrupted day in the participating districts. In order to ensure that the excerpts represented different parts of the unit, we occasionally took the sample from a Monday or a Wednesday. This also happened in the rare event that Tuesday’s class was disrupted (e.g. school assembly, fire drill). For the starting point of the four-minute excerpt, we chose a random start time with the parameters that (a) half of the excerpts for each class were taken from the first half of the lesson, and half were taken from the second half of the lesson and (b) no excerpt would begin in the last four minutes of the lesson. This was to ensure that non-instructional activities were not sampled. Once selected, the excerpts were transcribed. Transcriptions were reviewed with any discrepancies resolved through discussion and re-watching the video.
Our goal for the discourse analysis was to identify aspects of argumentative talk that could potentially undermine myside bias and was also likely to show up at this grade level. To this end we examined the discourse excerpts for (1) discussion about the sides of the controversial issue and (2) challenging and requesting explanations.

To preface the results I need to explain our assumptions about operationalizing argumentative talk. The readings and the assignments in this unit focused on (a) profiling the stakeholders and explaining their position and supporting reasons or (b) explaining concepts relevant to understand the wolves’ impact on an ecosystem, economy, and public opinion. Concepts were explicitly linked back to the broader argument both in the readings and in the assignments. Because of this unit design, I assume that any talk in relation to the unit and its activities was contributing to a holistic understanding of the controversial wolf issue.

Given the view that argumentation is the practice of justifying positions, we began our search for argumentative talk by completing a word search for pertinent discourse markers. As a low-inference indicator of the connectedness of the talk, a word search was performed for the conjunction and, the contrastive but and the causal connectives because, so, and then. Use of the terms was then examined in context in order to ensure that the words were (a) used in discussing the wolf unit and (b) indicative of connected talk, and relational thinking. Terms were removed that were simple repetitions, used in speech with no propositional content, read verbatim from the text and used in speech not related to the unit (e.g. chatting about recess, requesting materials).

Challenges and requests for information were also coded. A list of challenges and examples is provided in Table 3.3. Challening and requesting information were noted
because they usually result in the challenged individual providing additional information in defense of themselves. In this way challenges may force discussants to further elaborate their own thinking, and it also creates the expectation that such requests are fair game. It also serves to allow the challenger to act as agent of change in the discussion. It is important to note that counts of challenges are not mutually exclusive with the counts of discourse markers. The challenges are a more qualitative attempt to capture a specific argumentative element we hypothesize to be related to a reduction in myside bias; encountering others’ viewpoint and positioning oneself in relation to their opinion.

Table 3.3
Challenges and their forms

<table>
<thead>
<tr>
<th>Challenge form</th>
<th>Speaker: Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Student: <em>Look, wolf kill people too.</em></td>
</tr>
<tr>
<td></td>
<td>Challenger: <em>Nuh unh</em></td>
</tr>
<tr>
<td>Requests reason</td>
<td>Challenger: <em>The- the people who wanta- who want to kill the wolves give me two reasons why they should kill the wolves and give me another two reasons why they shouldn’t kill wolves</em></td>
</tr>
<tr>
<td>But [opinion]</td>
<td>Student: <em>I think they should kill the wolves.</em></td>
</tr>
<tr>
<td></td>
<td>Challenger: <em>Yeah. But I don’t know ’cuz wolves is like dogs and my favorite animal is dogs.</em></td>
</tr>
<tr>
<td>Yes</td>
<td>Student: <em>And uh, and like Kerinn said wolves can be heard from really far away so they never know how far away the wolves are. Noboby seen ’em and it it doesn’t say anywhere that anyone seen ’em.</em></td>
</tr>
<tr>
<td></td>
<td>Challenger: <em>Yes it did.</em></td>
</tr>
<tr>
<td>Why</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Student: <em>Huh? I think they shouldn't.</em></td>
<td></td>
</tr>
<tr>
<td>Challenger: <em>Ok. Why?</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modifies statements with counterexample.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher: <em>Because of [1] [1] the wolves, [NV1] right? You know I mean there's more animals like this they should be thankful in a way to the wolves because thanks to them, there was more what for the scavengers?</em></td>
</tr>
<tr>
<td>Challenger: <em>Not always. Like in the movie um um the wolves um they ate half of the e- elk, and they just left it there, but th- he but there were still near the elk, and a coyote came, and it was trying to eat it, and the wolf attacked it.</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No, [relevant info]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student: <em>We kill more sheep than hunters though.</em></td>
</tr>
<tr>
<td>Challenger: <em>No we kill more sheep than wolves.</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How do you know?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student: <em>When wolves kill elk they sharpen their teeth on their horns.</em></td>
</tr>
<tr>
<td>Challenger: <em>OK. Uh ok, when they kill elk and sharpen teeth on h- horns? How’d you learn that?</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>But [hypothetical]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student: <em>Because they are dangerous.</em></td>
</tr>
<tr>
<td>Challenger: <em>But, what if s- somebody like, had a, like a wolf- like as a pet ever since it was like a baby or something?</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>But [relevant fact]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student: <em>I think they [pause 4 sec] they should kill the wolves because they are um they're mostly killing a lot of people. And . . . a lot of wolves are mean. Most of them mean most of the wolves are mean.</em></td>
</tr>
<tr>
<td>Challenger: <em>But it doesn't say the wolves kill a lot of people.</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requests explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student: <em>Um, I think yes because the wolves threaten them to show them who is the top dog.</em></td>
</tr>
<tr>
<td>Challenger: <em>Okay. And what does that mean?</em></td>
</tr>
</tbody>
</table>
Chapter 4

RESULTS

The following section will present the analysis and results for the discourse analysis and each of the outcome measures. First, I will provide an analysis of the transcribed excerpts from the video data. Next, I will detail the analysis of the outcome measures starting with a look at myside bias in each of the outcomes using all available data. Following this I present an analysis using only the students who completed all three outcomes. Finally, I will present an analysis of the elaborateness of the reasoning used in all three outcomes, again using all available data.

4.1 Analysis of classroom talk

Figure 4.1 shows the differences in mean use of conjunctions, contrastives, and causal connectives. Using a mixed model with observations nested within classrooms, an analysis indicated that students and teachers in CG engaged in more connected and argumentative talk (β = .84, S.E. = .23, p < .001). Removing teacher talk, a second analysis showed that the disparity between conditions is greater when looking exclusively at student talk (β = 1.48, S.E. = .18, p < .001). Ninety percent of the connected talk in CG condition was made by students. All models were fit with and without a random intercept and with random slopes. No model with a random slope converged due to a Hessian matrix not being positive definite. Both models used a class average of the reading test to control for between-class academic proficiency and the class average of the object naming task to control for between-class English proficiency. These variables were not significant in either model, but their inclusion provided better model fit. No interactions were significant. Both models assumed a negative binomial distribution because the data
were count data and over-dispersed. Assuming a negative binomial distribution resulted in a better fit than assuming a Poisson distribution for both models, according to likelihood ratio tests.

Figure 4.1
*Average use of connectives by students and teachers*

Beyond looking for markers of connected talk and relational thinking, we wanted to look at other indicators of argumentative talk. In order to do this we compiled counts of challenges students and teachers made to one another. The resulting count is a way to operationalize two important argumentative elements of establishing a point of controversy and requesting explanations. Using the model building approach for mixed models detailed earlier, the final model showed that students and teachers in CG challenged each other more than those in DI (B=1.19 S.E. = .19, p < .001). Interestingly, 94% of the challenges in CG condition were student-to-student. It is important to note that challenges included some of the conjunctions, contrastives, and causal connectives,
so this result is not mutually exclusive of the first analysis. It is valuable, however, in that it provides a more specific test of the hypothesis that collaborative group work will foster more argumentative talk than a direction instruction format.

4.2 Analysis of outcome measures

To assess the level of myside bias I looked primarily at whether students established a position and then considered a counter argument (CCA) or whether students weighed both sides of the issues without coming down on one side or another. The whale interview involved prompts for a counterargument, if one was not spontaneously provided by the student, and for that reason I only counted what was provided spontaneously for this analysis. This ensured a conservative estimate of the student’s reasoning. Figure 4.2 shows the percentage of students from each condition who either took a position and considered a counterargument or weighed the issue.

Figure 4.2
Percentage of students by condition who articulated reasons on both sides

<table>
<thead>
<tr>
<th>PERCENTAGE OF STUDENTS</th>
<th>CG</th>
<th>DI</th>
<th>CTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wolf letter</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td>Whale interview</td>
<td>35%</td>
<td>25%</td>
<td>15%</td>
</tr>
<tr>
<td>Jack and Thomas</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
</tr>
</tbody>
</table>
To examine treatment effects on the likelihood that students would articulate reasons on both sides of an issue, we fit three multilevel logistic regression models, one for each outcome, with students nested within classrooms. The outcome variable was dummy coded to indicate whether the individual had articulated reasons on both sides for a particular outcome measure (1=Yes, 0=No). Table 4.1 provides information about the three models.

Table 4.1
*Model building for reasons on both sides in the three outcomes*

<table>
<thead>
<tr>
<th></th>
<th>-2LL</th>
<th>Intercept</th>
<th>( \beta_R )</th>
<th>SE</th>
<th>( \beta_C )</th>
<th>SE</th>
<th>( \beta_{CG} )</th>
<th>SE</th>
<th>( \beta_{DI} )</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wolf Letter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>2712.11</td>
<td>0.5835</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>2705.92</td>
<td>-1.28</td>
<td>.02**</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>2648.75</td>
<td>-1.16</td>
<td>0.66</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R, E</td>
<td>2644.92</td>
<td>-1.3</td>
<td>.03**</td>
<td>0.01</td>
<td>0.04</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>2734.38</td>
<td>-2.04</td>
<td></td>
<td></td>
<td>1.51*</td>
<td>0.37</td>
<td>1.47*</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R, E, C</td>
<td>2672.87</td>
<td>-2.41</td>
<td>.03**</td>
<td>0.01</td>
<td>0.08</td>
<td>0.86</td>
<td>1.48*</td>
<td>0.34</td>
<td>1.37*</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Whale interview</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>2028.11</td>
<td>-1.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>2056.32</td>
<td>-2.13</td>
<td>.05**</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>1961.87</td>
<td>-2.36</td>
<td>2.44**</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R, E</td>
<td>1986.83</td>
<td>-2.50</td>
<td>.05**</td>
<td>0.01</td>
<td>1.01</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>2021.93</td>
<td>-2.00</td>
<td></td>
<td></td>
<td>1.68**</td>
<td>0.3</td>
<td>0.43</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R, E, C</td>
<td>2055.58</td>
<td>-2.31</td>
<td>.05**</td>
<td>0.01</td>
<td>0.74</td>
<td>1.08</td>
<td>1.63**</td>
<td>0.31</td>
<td>0.37</td>
<td>0.34</td>
</tr>
<tr>
<td><strong>Pinewood essay</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>3266.45</td>
<td>-1.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>3270.46</td>
<td>-1.96</td>
<td>.03**</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>3166.63</td>
<td>-2.34</td>
<td>1.74*</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R, E</td>
<td>3166.58</td>
<td>-2.41</td>
<td>.03**</td>
<td>0.01</td>
<td>1.06</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>3308.24</td>
<td>-1.76</td>
<td></td>
<td></td>
<td>.76*</td>
<td>0.25</td>
<td>-0.07</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R, E, C</td>
<td>3202.75</td>
<td>-2.52</td>
<td>.02*</td>
<td>0.01</td>
<td>1.01</td>
<td>0.91</td>
<td>.74*</td>
<td>0.24</td>
<td>-0.01</td>
<td>0.28</td>
</tr>
</tbody>
</table>

* p < .05  ** p < .001

Parameters in Table 4.1 were fit using GLIMMIX in SAS. The model building was iterative and I maintained covariates across different models, even when not significant, when including the covariate improved overall model fit. Decisions about model fit were based on a log likelihood ratio test. Pretest measures of reading
comprehension and language proficiency were entered into the model first as past work has shown them to be useful individual differences measures. I also included gender, ethnicity, intervention year, and length of texts, in the model building process. None of these fixed effects were significant predictors and none of them improved overall model fit when predicting any of the three outcome measures and so they were left out of the final model. For both the comprehension and literacy measures I created a variable for both the class mean and the individual’s score centered on that class mean. This allowed for us to test more complex interactions between the condition and the different classrooms comprehension and language scores. No interactions between the group-mean centered variables and condition were found. Moreover, model fit was significantly decreased by including group means and group-mean centered variables and so in the final model both variables were left as individual measures and assumed to be level-1 variables in the two-level hierarchical linear model. The final models for each of the three outcomes ended up being identical,

\[ Y_{ij} = \gamma_{00} + \gamma_{10}[\text{Reading}]_{ij} + \gamma_{11}[\text{English}]_{ij} + \gamma_{11}[\text{Condition}]_{ij} U_{0j} + R_{ij} \]

For the wolf letter, as compared to CTL, both CG (odds ratio CG/CTL= 4.43, 95% CI [1.17, 13.30]) and DI (odds ratio DI/CTL= 3.97, 95% CI [1.26, 12.46]) students were significantly more likely to articulate reasons on both sides of the issue, either by considering an argument counter to their own or by weighing both sides of the issue. Specifically, the odds of a student in CG and DI considering both sides of an issue was around four times the odds that a student in the control group would consider both sides. There was not a significant difference between the DI and CG.
The whale interview was a near transfer reasoning task which overlapped with the wolf unit at the conceptual level but not the surface level. For the whale interview, the odds of a student in CG (odds ratio CG/DI=3.50, CI [1.59, 7.73] ,CG/CTL= 5.10, CI [2.11, 12.35]) considering reasons on both sides of the issue, without prompting, was 3.5 times the odds of a student in DI and 5 times the odds of a student in CTL considering both sides. Table 4.2 provides additional information about the amount of prompting students in each condition needed at each stage in argument construction.

Table 4.2
Proportion of students in prompting categories

<table>
<thead>
<tr>
<th>Protocol Section</th>
<th>Code</th>
<th>Percent by Condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CG</td>
<td>DI</td>
</tr>
<tr>
<td>Prompt for Reason</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>96</td>
<td>90</td>
<td>93</td>
</tr>
<tr>
<td>Elaborate</td>
<td>4</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Considering Both Sides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position and CCA</td>
<td>30</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Both positions</td>
<td>12</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Simple</td>
<td>17</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Elaborate</td>
<td>33</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>None</td>
<td>3</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Response to counter argument</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneous</td>
<td>36</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>Simple</td>
<td>39</td>
<td>45</td>
<td>33</td>
</tr>
<tr>
<td>Elaborate</td>
<td>5</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>None</td>
<td>11</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

*CCA = consideration of counter argument

The pinewood essay was a far transfer reasoning task that had little conceptual overlap with the other two tasks. In writing the pinewood essay, students in CG had a
higher chance of considering both sides than students in DI (odds ratio CG/DI = 2.22, CI [1.18, 4.43]) or students in CTL (odds ratio CG/CTL = 2.10, CI [1.03, 4.30]). In both cases the odds of CG students articulating reasons on both sides of the issues was twice the odds of the DI and CTL students.

**4.2.1 Students with all three outcomes.** A reduced sample of 340 students completed all three outcome measures. Table 4.3 shows the percentage of students in each condition who considered reasons on both sides of the issue in none, one, two, or all three of the outcomes. Given my interest in looking at reasoning as something that is transferrable across topics, I divided individuals into two groups, those that had considered both sides in at least two of the outcomes and those that had considered both sides in zero or one of the outcomes. This categorization allowed me to look at conditional differences in students who avoided myside bias with different controversies.

Using model fitting strategies detailed earlier we arrived at a model that included both reading comprehension and language proficiency. Neither measure was a significant predictor in the model but their inclusion improved model fit substantially. According to the final model, students in CG were more likely to consider reasons on both sides of an
issue for two or three of the outcomes than students in the CTL classrooms (odds ratio CG/CTL=11.62, CI [3.43, 39.38]) or the DI classrooms (CG/DI= 4.12, CI [1.5, 11.67]).

Figure 4.3 gives a breakdown of the outcome or outcomes in which students avoided myside bias. For students who considered both sides in only one outcome, it is notable that a large percentage of students from the conditions that participated in the wolf unit were more likely to consider both sides when writing about the wolf question. Students in the control condition were most likely to consider both sides in only one outcome measure, and they were most comfortable entertaining both sides of the pinewood derby dilemma.

Figure 4.3
*Percentage of students who articulated reasons on both sides, in one or two outcomes*

We can also calculate conditional probabilities as an indication of how likely students were to avoid myside bias if they had avoided it in another outcome. If a student
in CG considered both sides on one of the outcomes, the probability that they would consider both sides on one of the other two outcomes was 50%. If a student in DI considered both sides on one of the outcomes the probability that they would consider both sides on one of the other two outcomes was 25%. If a student in CTL considered both sides on one of the outcomes the probability that they would consider both sides on one of the other two outcomes was 11%.

4.2.2. Elaboration of reasons. Next, I investigated the degree to which students supported their position with elaborated reasons. The unit of analysis was the C-unit. A single C-unit represents a single reason, or an elaboration of a reason. Table 4.4 provides the average number of c-units coded as elaborated reasons for each outcome measure. Each student would have a discrete number of elaborated reasons per text, so we treated elaborations as count data. Investigating the different counts for each outcome, I found that in every case the distributions roughly conformed to a Poisson distribution. Some of the distributions were somewhat over-dispersed, so I fit two models, a multilevel log-linear model assuming a Poisson distribution and a multilevel log-linear model assuming a negative binomial distribution. Using model fit strategies and statistics described earlier, I found that assuming a Poisson distribution provided a better fit, however, all parameters reported as significant were significant when assuming either distribution. Table 4.5 includes the parameter estimates for the final models.

For elaborations in the wolf letter, students in the CG ($\beta=.72$ SE=.13, $p<.01$) and DI ($\beta=.71$ SE=.22, $p<.01$) outperformed students in CTL, while DI and CG were not significantly different. For the whale interview, CG provided more elaborated reasons than students in control ($\beta=.27$ SE=.08, $p<.01$) but there was no difference between DI
and CTL or DI and CG. When counting elaborations in the whale interview I did not differentiate between elaborations that were provided spontaneously or when prompted.

For the pinewood essay, there were no reliable differences between conditions in number of elaborations.

Table 4.4
*Average number of elaborated reasons in each outcome*

<table>
<thead>
<tr>
<th>Cond</th>
<th>Outcome</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG</td>
<td>Wolf letter</td>
<td>3.57</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>Whale interview</td>
<td>1.21</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>Pinewood essay</td>
<td>1.84</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>Wolf letter</td>
<td>3.18</td>
<td>4.00</td>
</tr>
<tr>
<td>DI</td>
<td>Whale interview</td>
<td>0.87</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>Pinewood essay</td>
<td>1.58</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>Wolf letter</td>
<td>1.60</td>
<td>4.00</td>
</tr>
<tr>
<td>CTL</td>
<td>Whale interview</td>
<td>0.57</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Pinewood essay</td>
<td>1.43</td>
<td>1.33</td>
</tr>
</tbody>
</table>

Table 4.5
*Model building for elaborations of reasons in each outcome*

<table>
<thead>
<tr>
<th></th>
<th>-2LL</th>
<th>Intercept</th>
<th>β₁</th>
<th>SE</th>
<th>β₂</th>
<th>SE</th>
<th>β₃</th>
<th>SE</th>
<th>β₄</th>
<th>SE</th>
<th>β₅</th>
<th>SE</th>
<th>β₆</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wolf Letter</td>
<td>1082.00</td>
<td>-1.53</td>
<td>0.02**</td>
<td>0.00</td>
<td>1.35*</td>
<td>0.45</td>
<td>0.72**</td>
<td>0.13</td>
<td>0.71**</td>
<td>0.22</td>
<td>0.03*</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whale interview</td>
<td>1389.91</td>
<td>-0.13</td>
<td>0.01**</td>
<td>0.00</td>
<td>0.39*</td>
<td>0.23</td>
<td>0.27**</td>
<td>0.08</td>
<td>0.18</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pinewood essay</td>
<td>2128.98</td>
<td>0.19</td>
<td>0.01**</td>
<td>0.00</td>
<td>0.48</td>
<td>0.30</td>
<td>0.13</td>
<td>0.17</td>
<td>0.14</td>
<td>0.02</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 ** p < .001
Chapter 5

DISCUSSION

5.1 Overview

This discussion will begin with a summary of the important findings focusing on the consideration of both sides in the different outcomes. Next, I will present findings from the sample of students who completed all three outcomes as evidence that CG students had generalized the consideration of both sides as a habit of mind. Following this, I will review the different characteristics of the discourse in the two conditions. Using those characteristics, I will discuss transfer of the consideration of both sides in the CG students’ reasoning as a function of collaborative practices in their condition. Finally, I discuss the lack of transfer of the consideration of both sides in the DI students’ reasoning as a function of students using a different habit of mind generalized from their discourse practice.

5.2 Summary of important findings

In this study, I looked for evidence that myside bias would be reduced for students who learned about the controversial topic of wolves in one or the other of two instructional frameworks. In the direct instruction condition students were led by a teacher through readings and activities that outlined the controversy about wolves with the goal of enriching the students’ knowledge about the complexities of an issue involving various stakeholders with different and often antithetical goals. In the collaborative reasoning condition, peer-led groups worked through the same readings and similar activities that highlighted various points of view with the goal of instigating
authentic argumentation about the controversy. Neither condition received any systematic instruction about the forms that arguments should take.

Both of these instructional conditions increased the likelihood that, in an essay about the wolf controversy, students would take a position on the issue and consider a reason on the opposing side, or present reasons on both sides without taking one of the positions as their own. Students in both CG and DI conditions outperformed students in control classrooms.

Comparable percentages of students in the CG condition considered reasons on both sides of the issue when asked to make a decision about the analogous whaling controversy as when asked the wolf question (see Figure 4.2). In contrast, there was a dramatic decrease in the percentage of students in the DI condition who considered both sides of the whaling controversy. Not only did students in DI not spontaneously consider reasons on both sides, but also a majority of them (see Table 4.2) needed an elaborate prompt (e.g., “What would the people that disagreed with you say?”) before they articulated a reason for the opposing position about the whale controversy. CG students seldom needed to be prompted to consider both sides. The knowledge-enrichment that occurred in the DI condition during the wolf unit was directly applicable to the whaling controversy and arguably should have facilitated a framing of the task and the transferring-in of relevant knowledge based on the surface- and deep-level similarities. But, for DI students, an in-depth understanding of one controversy did not reduce biased reasoning about a similar controversy.

The pinewood derby essay was designed as a test of far transfer considering the nearly complete lack of conceptual overlap between the topics. While fewer students
considered reasons on both sides in the pinewood essay, the percentage of CG students was still about twice the percentage of DI and CTL students who considered both sides in the pinewood derby essay.

5.3 Evidence for consideration of both sides as a habit of mind

The control group in this study provides a good baseline for how likely it is that an average student from this population would consider reasons on both sides of these three controversies. For all three controversies, very few (12-13%) of the CTL students articulated reasons on both sides of the issue. This means that collaborative group work contributed to a nearly four-fold increase in consideration of both sides in the wolf letter and whale interview, and nearly a three-fold increase in the pinewood essay.

These results provide strong and unique evidence that students who engage in dialogic argumentation about one topic are more likely to consider reasons on both sides of that topic as well as transfer that strategy to other topics. Results from the reduced sample of students who completed all three outcome measures show that if a CG student avoided myside bias on one outcome then the probability that they would avoid it on a second outcome was 50% as compared to 25% for a DI student and 11% for a control student. These conditional probabilities provide evidence that for some students the disposition to consider both sides had become a ‘habit of mind,’ a heuristic used when reasoning about complex controversies without a single or obvious answer.

5.4 Internalizing norms of discourse

How did students develop what appears to be a habit of mind? Evidence to answer this question comes largely from differences in the discourse in collaborative interaction and direct instruction. First, discussions in the DI condition generally did not encourage
the dialogic construction of knowledge through argumentation because teacher-led
discussions tended to follow a recitation style where students were expected to provide
only short bursts of requested information (Applebee, Langer, Nystrand, & Gamoran,
2003; Nystrand, 1986; Wells & Arauz, 2006). Second discussions in the CG condition
appeared to encourage dialogic argumentation because, in the absence of a teacher, the
students needed to elaborate and connect their own thinking to be understood. In
collaborative groups, students could not as readily assume common knowledge or shared
ideas as when speaking to an expert (e.g., teacher), and, because of this, students typically
needed to negotiate their ideas with the others in their group. Both of these conjectures
were supported by the discourse analysis completed in this study. Students in the CG
condition were not only engaging in more elaborated and connected talk with each other,
they were also elaborating their thinking more than the teacher and students in the DI
condition. The analysis of challenges (Table 3.3) also provided strong evidence that
students were elaborating their thinking in response to other students’ requests and
challenges.

5.5 Collaborative group work and attention to others’ viewpoints

It is my contention that over the course of the wolf unit students in the CG
condition began to understand what was expected of them in their collaborative groups
through a mixture of explicit requests from one another and tacit expectations. In
collaborative groups who repeatedly work together, certain discourse norms tend to
emerge based on negotiated roles, task demands, and various aspects of social
accountability (Webb, 1989; Li et al., 2007). The discourse analysis from this study
suggests that students in CG groups requested explanations and provided explanations for
one another. From their essays, we know that students in groups held different and quite varied opinions, and, so, we can assume that their explanations involved a variety of positions and reasons that were held and defended by fellow students. It is this active embodiment of the argument by the discussants that I hypothesize is the strongest factor in reducing myside bias.

While students in both the DI and CG classrooms were hearing and reading about the opinions of others, students in the CG classrooms were providing their own opinions as well as voicing the opinions expressed in the unit. This active arguing with others likely encouraged an awareness of each other’s commitment set (e.g. supporting reasons, concessions, agreements, see Walton & Krabbe, 1995). Along with tracking each other’s commitment sets, students would have become aware of strategies of reasoning that were more or less effective in persuading others. Other work from the collaborative reasoning project has shown that students attend to and reuse reasoning strategies voiced by others that they see as effective (Anderson et al., 1997; Lin et al., 2013; Morris et al., 2014). We know from previous work and the current study that requests for reasons and challenges are commonplace in collaborative group work. For students in CG, keeping track of other students’ arguments and regularly having one’s assumptions and reasons challenged led to an increased awareness of other positions in relation to one’s own. This is a departure from typical classroom discourse.

In a typical classroom, the teacher controls the flow of information through questioning. Students need only to supply as much information as the teacher requests, often only a word or a phrase (e.g., ‘And if wolves are killed the number of rabbits will…’). Thus, the student is responsible for only discrete bits of information (e.g.
‘increase’). In such situations, language planning and production are attenuated. The children lack agency inasmuch as they are not responsible for the whole constructive process.

Contrast this with argumentation in collaborative work where students have the opportunity to solicit reasons in order to understand or challenge another student’s thinking. To argue successfully they need to fully explain their own views without contradicting themselves. Participation in a collaborative group requires awareness of fellow discussants and their ideas. I contend that it is this repeated process of holding potential audiences in mind while planning arguments is what generalizes and is eventually internalized as awareness of audience. The weighing of both sides or dealing with counterpositions in independent reasoning stems from the imperative to attend to the views of others during discussions. In summary, students in CG would have more experiences actively engaging with discussants holding a variety of opinions, and this experience led to a more active engagement with both sides of the issue. It is still necessary to account for why such a strategy was effectively transferred to other reasoning tasks.

5.5.1 Transfer of ‘consideration of both sides’ heuristic. Transferring the strategy of considering both sides to topics other than the wolf unit would require that the students frame the other two reasoning tasks as similar to the wolf letter. Research by Engle and associates has shown that transfer and, they suggest, by necessity framing, is positively influenced by providing, “opportunities for students to actively contribute to larger conversations that extend across time, people, places and activities” (p. 605, Engle, Nguyen, & Mendelson, 2011; Engle, Lam, Meyer, & Nix, 2012). They refer to the
provision of such opportunities as *expansive framing* and they achieve this expansive
framing through a number of non-explicit contextual cues. These cues included
discussing past and future uses for the knowledge, connecting discrete facts with a larger
conceptual organizer, and ensuring that students felt that they were the author of their
own ideas, as opposed to simply a “spokesperson for the text.” Although Engle and her
colleagues were not examining collaborative groups, the conceptual parallel is obvious.
Students in the CG condition were working closely with the same classmates for six
weeks. During that time students were referring to past activities and explanations, and
keeping some track of each others’ commitment sets as they stabilized or shifted over
time. The whole of the unit was set up as a fact-finding mission to influence a future
decision on one central issue. For most students, it would have been apparent that the
ideas they were negotiating and their strategies for doing so were relevant beyond the
moment in which they were being used. In short, working in collaborative groups served
to expansively frame their learning, increasing the likelihood that they would see
connections between the wolf unit and other topics.

In conclusion, the students in CG learned the value of considering both sides
because such considerations would have been a reoccurring aspect of their daily
discussions. They were also more apt to then transfer this strategy because they perceived
their ideas and learning processes as relevant beyond their daily activities.

**5.6 Direct instruction and a different habit of mind**

The performance of the DI students in writing the wolf letter suggests that
knowledge-enrichment is an effective tool for building up a student’s understanding of
the multisided nature of one controversy. Compared to the control students, twice as
many DI students considered reasons on both sides of the wolf controversy. The much-reduced number of DI students who considered both sides in the whale interview suggests that whatever knowledge led to the reduction in myside bias about the wolf issue did not positively impact reasoning on a separate topic with both shallow and deep conceptual similarities. Similarly, there were markedly fewer students who considered both sides when writing the pinewood essay than when writing the wolf decision letter. I contend that this failure to transfer consideration of both sides to different topics was likely due to the students’ framing of all three tasks as opportunities to engage in the recitation of knowledge; the most common means of communication in the classroom. In the next section I propose why students’ framing the task as a means of demonstrating knowledge led to consideration of both sides in the wolf letter, but not in the other two reasoning tasks.

5.6.1 Reasoning about the wolf controversy. The idea that students are aware of potential audiences and target their writing to these perceived audiences is well supported (Gunel, Hand, & McDermott, 2009; Midgette, Haria, & MacArthur, 2008). It is likely that the student texts analyzed here were also targeted to an audience, and that that perceived audience was most likely their teacher as the teacher is, far and away, the primary audience for all academic language production in the classroom (Applebee et al., 2003; Cazden, 2001). The idea that considering both sides of the wolf controversy was something a teacher would appreciate most likely originated in the design and from the implementation of the wolf unit.

The readings and activities in the wolf unit outlined the many stakeholders who held different positions on the wolf problem, and so a teacher would need to delineate
and perhaps group the different positions in order to provide a framework for understanding the unit. In this way, a teacher would not be presenting a single side but modeling the consideration of both sides. Discerning students would perceive that articulating all of the reasons was valued, and proceed with such weighing of ideas in their writing. Like so many prior recitations, those students who made that inference could reproduce the facts and thinking the teacher had indexed as important.

5.6.2 Transfer of a different habit of mind. Still, this explanation does not fully account for why most DI students elaborated only one position in the whale interview. Wouldn’t students in DI want to demonstrate their ability to articulate reasons on both sides of the whaling issue, if they perceived that articulating reasons on both sides was a demonstration of proficiency to their teacher? I argue that this is primarily due to the norms of teacher-led instruction having the two fold effect of, (1) leading students to rely on an unsophisticated reasoning strategy of simply telling what they know about a subject, and (2) failing to help students identify the consideration of both sides of the controversy as a strategy to be applied to future topics.

Most discussions about the unit were tethered to the teachers in the DI classrooms. As discussed earlier, the teacher is the audience for most writing and class discussions. According to my analysis of the classroom discourse, the students in the DI condition did not practice articulating their ideas or challenging other’s thinking. There was little engagement between students with most of the discussion dictated by the teacher. I contend that repeated exposure to IRE forms of discourse led students to abstract a different sort of habit of mind that guided their text production, one that forefronted processes of demonstrating knowledge so the teacher could properly assess
your learning. When it came time to reason about the whale controversy and pinewood essay, students likely framed those tasks using the sort of ‘text production-as-assessment’ frame that they had used for the wolf letter, and, indeed most writing and speaking in their classes. In the wolf unit, this frame would have led students to include reasons on both sides as that was presented as a valuable relational concept by the teacher over the course of the unit. Lacking an in depth understanding of the material and a detailed understanding of how the teachers were thinking about the controversy, the students defaulted to showing their understanding using reasoning strategies common to this population.

So, one reason students did not consider both sides on the other tasks is because they were using a simplistic guiding heuristic of providing what they know. A second and related reason is that it is unlikely that students would have successfully generalized the consideration of both sides as a valued reasoning strategy to be used in other reasoning tasks. This failure to abstract the strategy is because, in contrast to students in CG, the DI students had little practice with the intra-personal processes of planning and encoding their reasoning to be received by multiple audiences as well as tracking the views of their peers. There were less instances of having to craft their reasoning to persuade another, because persuading the teacher is not a common practice. There were fewer instances of students encountering opposing views as embodied by their peers. It was at least partly this lack of practice in taking up those views and seeing others do the same, that contributed to students’ failing to sufficiently generalize the weighing of different sides as an effective strategy since its effectiveness is socially mediated and can only be truly demonstrated in situ.
In conclusion, the DI instructional framework was effective for helping students build up an understanding of the pros and cons of killing the wolves. One third of the DI students were able to articulate reasons on both sides of that issue but the failure to consider both sides on the two transfer tasks suggests that the reasoning was localized to their understanding of the wolf unit. In part, this failure to transfer the consideration of both sides was likely due to student reasoning primarily driven by the goals of recitation and presenting only as much as they were asked. I suggest that this a different sort of habit of mind, one that is less desirable in education but fostered nonetheless in classrooms everywhere (Nystrand, 1986).

One of the more exciting aspects of this study is that a six week intervention with only modest changes to the instructional practice, had a remarkable impact on disrupting this pervasive and generally uncreative habit of mind, and replaced it with an approach to reasoning that is crucial in decision making and thinking in all domains.

5.7 Limitations and future directions

The current study assessed reasoning on three different tasks, two of these tasks representing different “distances” of transfer. Analyzing these three outcomes allowed a triangulation of reasoning processes that afforded strong, if preliminary, evidence that some individuals can come to have fruitful ‘habits of mind’ when producing argumentative texts. Specifically, we have seen that some students transfer the strategy of considering multiple viewpoints across different topics, and this is more likely when given the opportunity to engage in authentic arguments that require considering multiple viewpoints.
One question that we can begin to answer with the current data is: do students who demonstrated a willingness to consider reasons on both sides continue to do so? The near complete lack of conceptual overlap between the wolf letter and pinewood essay suggest that the heuristic is generalized enough to be applied across considerable conceptual distance, but given that these assessments were administered only days apart, we cannot speak to its viability in the long run.

Although the discourse analysis included here is extensive, and unique in the literature, there are other analyses that might yield results that would bolster the arguments made here. At first, the selection and transcription of 4-minute excerpts was a practical decision because accurately transcribing open-ended and overlapping children’s discussion with multiple speaking floors is a labor-intensive process. We stopped transcribing after the selection of 144 excerpts presented here because the results from this set of excerpts were so robust and varied so little from class to class that there was little cause to sample more. Also, there is anecdotal evidence and field notes that bear out the story we were seeing in the data.

In the near future, I plan to identify one student in each condition who considered both sides on at least two outcomes. I will track their participation and group talk looking specifically for episodes where they integrate counterarguments in their own thinking. This case study approach is based on Nuthall’s (2007) pioneering work tracing the origin of ideas across time and in and out of both academic and informal social interactions.

There is a need to further characterize the collaborative discussions in terms of the argumentative goals that are taken up by the discussants (Walton, 1996). For example, are groups that are engaged in persuasion dialogues more likely to see a reduction in
my side bias than those engaged in consensus seeking? There are a few studies that show students who are seeking to persuade are less likely to address counter positions (Baron, 1995; Nussbaum & Kardash, 2005) and other work has shown that active dissent is better for knowledge building (Johnson & Johnson, 1985). By looking at group’s argumentative goals I can assess (1) if groups or individuals tended to adhere to one scheme (2) if certain schemes were associated with certain gains/deficits in independent reasoning processes. These are but a few of the intriguing questions that deserve further attention.

5.8 Conclusions

The current study has shown than dialogic argumentation in peer-led groups about authentic issues can greatly increase the likelihood that underserved students will consider reasons on both sides of complex socio-scientific policy issues as well as schoolyard ethical dilemmas. It has been argued here that this reduction in bias is due at least in part to students seeing the value of accounting for multiple perspectives in their social argumentation, as this was necessary to engage in the CG discussions, thus increasing the chances that they would use that strategy in other inter- and intrapersonal reasoning contexts.

From a practical standpoint, this study should encourage educators to include more genuine peer-led dialogue in their classroom in an effort to foster the development of strategies that a consensus of educators has deemed to be indicative of school and career readiness. The integration of one’s position with the position of another is touched on in two of the anchor standards in the Common Core Standards (2010, CCS.ELA-LITERACY.CCRA.SL1 & CCS.ELA-LITERACY.CCRA.SL3).
In a broader sense myside bias is a detriment to productive civil discourse. If we wish to steel future generations against the manufacture of consent resulting from the consolidation of news organizations, we must ensure that they we take seriously our charge to prepare our students to assess their knowledge and biases critically. This study has shown that a modest intervention can overcome one of the most pervasive biases in human reasoning.
WORKS CITED


