

The Effects of Collaboration on Student Writing Development

Natalie Boyd

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy
under the Executive Committee of the Graduate School of
Arts and Sciences

Columbia University

2018

© 2018
Natalie Boyd
All Rights Reserved

The Effects of Collaboration on Student Writing Development

Natalie Boyd

ABSTRACT

Dialogic argument activities have been shown to facilitate the development of argumentative writing in young adolescents. The present study investigates the extent to which collaborative writing has a further facilitative effect, serving as a bridge between the dialogic and individual writing contexts. Over the course of one school year, a total of 54 students in two low-performing 7th grade classes participated in a twice-weekly dialogic argument curriculum of known effectiveness that included various kinds of dialogic activities addressing a sequence of four topics and an individual essay as the culminating activity for each of the topics. In a quasi-experimental design, one class was randomly chosen as an experimental group and the other as a comparison group. The participation of the two classes in the curriculum was identical except that in one class students had an additional activity toward the end of each 15-session topic unit, during which they were asked to collaborate with a classmate who held the opposing view on the topic and produce a jointly written essay. The comparison group also wrote an interim essay but did so individually rather than collaboratively.

Compared to students who only wrote individually, collaborative writers performed better on their subsequent final individual essays on the topic. They anticipated the arguments of the other side better, and countered them using an integrative argumentation structure more often. Further, they repeated ideas less often and had more unique idea units in their essays.

To explore the collaborative processes possibly underlying the differences between the groups, analyses of digital voice recordings from the collaborative writing activity were examined. In addition, the transfer of ideas from the collaborative to subsequent individual essays was examined. The recordings of verbal dialogue between the pair engaged in

collaborative essay writing show an increase over the year in metacognitive dialogue pertaining to their task. Furthermore, in their subsequent individually-written essays, students utilized and built on ideas presented by their partner. Most notable was inclusion in the individual essay of arguments and evidence supporting the opposing partner's position, particularly when the ideas presented supported the opposite side of the argument. Both of these developments support the view that collaborative writing aids in the development of an argumentative mindset that transforms inter-individual dialogue into intra-individual reflection.

TABLE OF CONTENTS

LIST OF TABLES	ii
LIST OF FIGURES	iv
ACKNOWLEDGMENTS	v
Chapter 1: INTRODUCTION & LITERATURE REVIEW	1
Development of Argumentative Skills.....	1
The Argumentative Mindset & Writing.....	5
Development of Metacognitive Skills.....	7
The Current Study.....	11
CHAPTER 2: METHODS.....	13
Participants.....	13
Procedure	20
The Essays: Pre-, Interim, & Post-Topic	26
Scoring the Final Topic Essays.....	28
Scoring the Collaborative Recordings	30
CHAPTER 3: RESULTS.....	33
CHAPTER 4: DISCUSSION.....	52
Development of the Argumentative Mindset.....	52
Growth progression in essays	52
Integrative arguments in essays	53
Yours, Mine, Ours.....	55
Using ideas from collaboration.....	55
The Process of Collaboration.....	57
Discourse vs. writing	57
Metacognitive collaborations	58
Limitations & Future Directions.....	60
Conclusion	63
REFERENCES	64
LIST OF APPENDICES.....	68

LIST OF TABLES

Table 1. Number of participants for each group, by topic.	15
Table 2. Mean number of different types of idea units on pre-topic 1 essays by group.	17
Table 3. Proportions of topic 1 pre-essay statements.....	18
Table 4. Proportions of types of idea units in topic 1 pre-essay.	18
Table 5. Percentage of students who ever used each type of idea unit in topic 1 pre-essay.	19
Table 6. Description of argumentation curriculum daily activities for each topic, with essay days shaded.	21
Table 7. Topics and prompts for each unit within the current study.....	21
Table 8. Summary of the groups and how they worked for the three essay assignments.....	26
Table 9. Types of idea units used during scoring.....	29
Table 10. Types of dialogue used in collaborative writing recordings.	31
Table 11. Mean number of idea units between groups for all post-topic essays. Proportions of idea units to total statements in parentheses.....	37
Table 12. Mean number of different idea units by group for topic 1 post-essay.	38
Table 13. Mean number of different idea units by group for topic 2 post-essay.	38
Table 14. Mean number of different idea units by groups for topic 3.	39
Table 15. Mean number of different idea units by groups for topic 4.	39
Table 16. Percentage of students who ever used each type of idea unit in post-topic essays.....	40
Table 17. Proportions of types of idea units in post-topic essays by topic and group.	40
Table 18. Coefficients for binomial mixed models analysis across topic, with treatment group set as reference group.	41
Table 19. Exponentiated coefficients for binomial mixed models analysis across topic, with treatment group set as reference group.	41
Table 20. Differences between groups in topics 1 and 4 means.	43
Table 21. Differences at topic 1 pre-essay and topic 4 post-essay, by group.	44
Table 22. Coefficients for binomial mixed models analysis on pre-test and post-test, with treatment group and pre-test set as reference.	45
Table 23. Exponentiated coefficients for binomial mixed models analysis on pre-test and post-test, with treatment group and pre-test set as reference.....	45
Table 24. Percentage of groups who used each type of discussion by topic.....	48
Table 25. Mean frequency of occurrence of discussion types in topics 1 and 4.....	48
Table 26. Mean number of each type of idea unit by source and topic.	49
Table 27. Percentage of students who ever used each type of idea unit by source and topic.	50

Table 28. Discourse between dyads during topic 2 (animal testing). 59

LIST OF FIGURES

Figure 1. Stages of argumentative development in adolescents.	3
Figure 2. Number of students present for each topic.	15
Figure 3. Total number of idea units and repeated statements for each group for each of the four post-topic essays.	34
Figure 4. Number of different types of idea units for post-topic essays for treatment and comparison groups.....	35
Figure 5. Number of M+ and O- reasons used by each group in post-topic essays for each topic.	36
Figure 6. Mean number of idea units for each group across topics.	37
Figure 7. Proportions of metacognitive and non-metacognitive discourse in collaborative recordings for each topic.	47

ACKNOWLEDGMENTS

It takes a village to raise a researcher, and I am deeply indebted to the village that surrounded me during this process. This includes my dissertation committee for all the time and thought given to this document. In particular, I'd like to acknowledge Laura DeRose, for her guidance on adding more information about the process of teaching to this document; Lisa Son, for her thoughtful suggestions about metacognition; Bryan Keller, for all the hours spread over almost two years that he gave to me to go over the statistics; Dolores Perin, for chairing my dissertation committee and for her thoughtful feedback focused on making this ready for future publication; and, of course, Deanna Kuhn, for all the years she's spent moving me from an educator to a researcher. I have learned so much from you. I also owe a debt of gratitude to Flora Matos for her help collecting the data used in this study.

In addition, I'd like to thank my colleagues and friends who have seen me through this process, including Davie Floyd, Lia Papathomas, Julia Zavala, and Nicole Zillmer. Thanks also go to Isabel Gutierrez, Barbara Seater, Brandyn Heppard, and Deb Corbett for their support, encouragement, and guidance as I made it through the final days of this process. Finally, thanks to my family, especially my parents, for pushing me and supporting me and making this possible.

Chapter 1: INTRODUCTION & LITERATURE REVIEW

Argumentative writing is a key 21st century skill (Binkley et al., 2012; Dede, 2009; Kuhn, 2015). The ability to understand multiple viewpoints, evaluate options, and argue in support of the best possible option are all a part of constructing meaning from the multitude of information available today. However, argumentation is cognitively taxing, and difficult to teach (Howe, 2010; Kuhn & Crowell, 2011; Kuhn, Zillmer, Crowell, & Zavala, 2013; Newell, Beach, Smith, & Vanderheide, 2011). Some research has shown that dialogic activities can help students develop argumentative writing (Howe, 2013; Kuhn, Hemberger, & Khait, 2016; Newell et al., 2011), though less research has been done on how collaborative writing can benefit students. In this chapter, I review the literature available on the development of argumentative skills through collaboration. Following that, I highlight the rationale and focus of the current study.

Development of Argumentative Skills

Because of its complex nature, and the various skills involved in arguing well, argumentation does not necessarily develop naturally in adolescents (Felton & Kuhn, 2001; Kuhn & Crowell, 2011). There is evidence that dialogic collaboration can offer students a way to cultivate argumentative skills (Howe, 2010; Kuhn & Crowell, 2011; Kuhn et al., 2013; Newell et al., 2011).

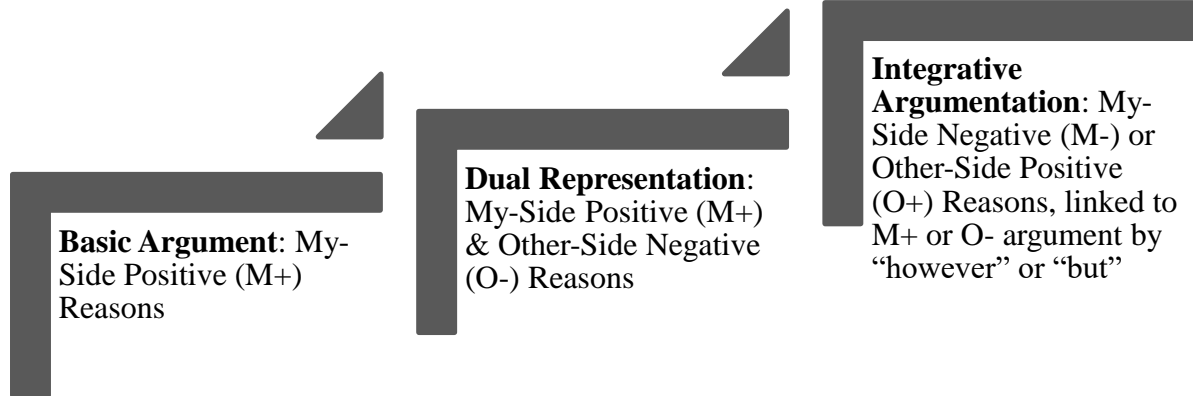
It might seem intuitive that arguing with others leads to argumentative development, but all arguments are not created equally. Even young children can state a position and declare their position superior to that of others, so why don't their arguments naturally develop as they argue with the people around them?

The answer lies in the type of structured dialogic argumentative tasks given. Through dialogue with peers, especially peers who hold opposing viewpoints, students can begin to evaluate ideas and evidence (Kuhn et al., 2013; Newell et al., 2011; Zillmer, 2016). Further, these students begin to understand how to construct arguments that address opposing viewpoints (Graff, 2008; Kuhn et al., 2013; Newell et al., 2011).

Kuhn and colleagues have identified the general way that students' argumentation skills develop (Kuhn, Hemberger, & Khait, 2014; Kuhn et al., 2016). Figure 1 shows a summary of the three levels of argumentative development. Students start with single representations of an argument, whereby they offer only reasons that support their side. Graff (2008) identifies this type of "my-side" only argumentation as characterized by a "missing interlocutor." That is, students are writing without recognizing or taking into account the opposing position. This type of argument does not require dialogic processes; in fact, this is the level of argument that most beginning students produce (Kuhn et al., 2014; Newell et al., 2011).

The next level of argumentation involves dual representations of an argument. At this point, students express using reasons (and possibly evidence) that support their own side, but now they are also able to include reasons and evidence that challenge the other side's position. These dual representations emerge over time when students participate in a dialogic argumentation curriculum (Kuhn et al., 2014, 2013).

Figure 1. Stages of argumentative development in adolescents.



The most advanced levels of argumentation involve first representing arguments inconsistent with one’s own position and, subsequently, constructing integrative arguments, in which students connect arguments against their own position or in support of the other side to arguments consistent with their own side. These integrative arguments can be called ‘however’ clauses or ‘this-but-that’ arguments due to the structure of the argument (e.g., a my-side negative reason, followed by ‘however’ or ‘but,’ most often followed by a counter in the form of a my-side positive reason) (Kuhn et al., 2014). With integrative argumentation, students have constructed a representation of Graff’s (2008) missing interlocutor, an internal voice that allows them to consider the other side and construct sophisticated responses to it.

Integrative arguments do not develop in all students, but they often do with prolonged participation in a dialogic curriculum (Crowell, 2011; Kuhn et al., 2014). The time taken is not surprising, as children find it harder to process facts against their beliefs than adults do (Kuhn & Pease, 2006). But the dialogic focus of collaboration in the curriculum supports students’ developing skill in anticipating others’ arguments and countering them, holding both their own side and the opposing side’s arguments in mind at once. Doing so is facilitated by external dialogue with a real rather than imagined interlocutor.

Not all dialogue is created equally, though. For decades, there has been a focus on collaboration in education, particularly when it comes to cooperative learning (e.g., Johnson, Johnson, & Stanne, 2000; Johnson & Johnson, 1989; Slavin, 1990). Cooperative learning can take many forms, of varying effectiveness (for a review of different types of cooperative learning, see Johnson et al., 2000). And while not all forms of cooperative learning involve avoiding group discord, many educators choose to spend time building a positive community in the classroom and avoiding disagreement.

It is likely not beneficial to avoid disagreement altogether. Social environments are often filled with conflict by nature, especially those involving children and adolescents. Yet, the very nature of disagreement can cause gains in students, both cognitively and socially (Howe, 2010; Jarvis, Newman, & Swiniarski, 2014; Kapur, 2008; Skoumios, 2009). This growth via discord is known as “sociocognitive conflict” (Skoumios, 2009). Ample research supports the idea that sociocognitive conflict can contribute to the development of both social skills (Järvelä et al., 2015; Johnson et al., 2000; Mercendetti, 2010; Rogat, Linnenbrink-Garcia, & DiDonato, 2013; Skoumios, 2009) and cognitive (particularly argumentative) skills (Howe, 2010; Kapur, 2008; Kuhn & Crowell, 2011; Mercer & Howe, 2012; Skoumios, 2009). For example, Howe (2010) found that students who discussed a subject with someone holding an opposing view demonstrated gains on post-tests several weeks later, regardless of whether the pairs came to agreement during collaboration or not. Apparently, it is the struggle (and, as Howe argues, the subsequent reflection on the argument) that benefits the students, not the resolution.

Arguing with their peers, then, allows students to develop an argumentative mindset, wherein they are capable of taking the other side into consideration when arguing. This mindset is acquired as students’ interactions and collaborations with others help create a voice in their

head that allows them to hear and counter arguments from the other side, just as during the dialogic process, they heard others argue (Felton & Kuhn, 2001; Kuhn & Crowell, 2011; Kuhn et al., 2014). This internalization of the dialogic process also helps to explain why Howe and others have found that collaborating with peers who disagree is not an easy task but yields the most benefits (Howe, 2010; Newell et al., 2011).

It is important to note that dialogic collaboration not only serves as the tool with which students develop this argumentative mindset, but also as the field within which they use it. Arguing with others allows students to become better at argumentation, but it also shapes the way they argue with others (Kuhn et al., 2014, 2013). This, in turn, leads to further practice and honing of these skills, creating an upward spiral of skill development.

The Argumentative Mindset & Writing

Thus far, we have looked at the ways in which argumentative skills develop via the creation of an argumentative mindset that fills in Graff's "missing interlocutor" (2008). But how might collaborative writing offer additional benefits, above and beyond those offered by non-writing dialogic activities? To answer that question, we must look at the ways in which argumentative writing is both a social and cognitive task.

The social aspect of argumentation involves looking at the context of collaboration, whereas the cognitive perspective looks at the specific, individual processes dedicated to argumentative writing (Newell et al., 2011). Both, of course, are relevant to the use of dialogic activities to help develop argumentative writing. For example, an accomplished argumentative essay will address the other side's perspective. This skill can be examined as a part of an individual's argumentative development (i.e., a cognitive framework) or as a part of the development of understanding others' perspectives (i.e., a social framework).

Much of the research on collaboration in the classroom has focused on the cognitive element of collaboration. Schwartz (2008), for example, writes that in collaboration, “cognition is propagated from mind to mind...in such a way that it: (a) is shared between the constituents of the group, (b) creates representations within and between [members] of the group, and (c) combines itself into a dynamic coordinated system” (pg. 390). That is, as team members work together, the sharing of ideas leads to higher cognitive levels. As you and I collaborate, we create a language and understanding that becomes the basis of our collaborative work.

It is difficult to argue that cognitive elements are not in play during collaboration. After all, the very process of discussing a task with another requires some cognitive interplay. The interplay of the social and cognitive tasks required for collaboration helps students grow as thinkers (Chinn & Clark, 2013; Vass & Littleton, 2010). Collaborative argumentation can lead to “explicit elaborative processing” (Chinn & Clark, 2013, pg. 321). That is, students working together elaborate on their ideas through explanation and argumentation. This allows them to build deeper understanding as they discuss with others.

Howe has found that when two students who hold opposing viewpoints on a topic are asked to talk and come to an agreement, their performance on post-tests measuring argumentative reasoning increases even when the partners do not actually come to an agreement (see, for example, Howe, 2010, 2013; Mercer & Howe, 2012). The argument itself is enough to create cognitive gains. Further, those gains are still present weeks later (Howe, 2010).

It is in the dialogue of the partners that the students are growing. Collaboration and dialogue are not easy tasks; students must utilize both cognitive and social skills simultaneously. This creates a unique challenge for students, but it appears that the struggle with collaborators is what leads to cognitive gains (Howe, 2010; Kapur, 2008; Kiili, Laurinen, Marttunen, & Leu,

2012). That is, dialogue (particularly argumentative dialogue) contributes to the growth of thinking.

Thus argumentative writing lies at the intersection of cognitive and social skills. This is precisely why dialogic argumentation tasks can be such powerful tools in the development of argumentative skills and in particular argumentative writing skills. Research has shown that the two develop along similar paths—and, in fact, argumentative skills are often evaluated via argumentative writing (Kuhn et al., 2016). Students start by writing about their own side and then develop the ability to address and weaken the opposite side. They further may develop the ability to find weaknesses in their position and/or strengths in the other side’s position. Finally, some students develop the ability to integrate arguments by linking two opposing statements (such as an other-side strength and other-side weakness) using words like “however” or “but” (Kuhn et al., 2016).

As argumentative skills in general and argumentative writing skills specifically develop similarly, it is not unreasonable to believe that the same mechanisms are involved in each. As noted, one theory of the development of argumentative skills along those lines is that dialogic argumentation allows students to develop the voice of the other side in their head, which eventually translates to more flexible cognition, as they are able to hold both their own views and opposing views in their heads (Graff, 2008; Kuhn et al., 2016, 2013). But what underlies this development?

Development of Metacognitive Skills

It can be argued that metacognitive skills enable the argumentative mindset, allowing students to think flexibly about arguments and anticipate others’ ideas. Metacognition involves “higher order thinking which involves active control over the cognitive processes engaged in

learning,” such as strategy and evaluation of approaches to a task (Livingston, 2003, pg. 2). Dialogic activities can boost metacognition (Näykki, Järvelä, Kirschner, & Järvenoja, 2014; Panadero & Järvelä, 2015; Zillmer, 2016). But how do they do that?

Järvelä et al. (2015) suggest three elements that should be taken into consideration when designing a learning environment to support the development of metacognition. First, the learning environment should raise awareness of one’s own and others’ learning processes. That is, the context in which students work should cause them to become aware of how they and their partners work.

The second element to be understood when designing a learning environment to boost metacognition is that it should support the discussion of learning processes. The authors call this “externalization,” whereby students discuss how they are approaching a learning task and which strategies work (and which don’t) (Järvelä et al., 2015).

The final element that educators should keep in mind when designing a learning environment is that it should teach the regulation of learning processes. That is, students should learn how to change their learning strategies as they work if they find that they are not succeeding (Järvelä et al., 2015).

Of course, these three elements build upon each other. As students become aware of their learning processes, they are likely to discuss these learning processes, which then leads to self-regulation of their learning. There is evidence that these elements do in fact build upon each other. For example, self-regulated learning develops across time as students collaborate with others (Grau & Whitebread, 2012), which suggests that meta-talk (as in Järvelä et al.’s second element) can lead to individual self-regulatory growth.

Using these elements as a framework, it becomes clear how metacognition may develop through dialogic collaboration. It creates a context in which students are not only collaborating on the content of the task but collaborating on the regulation of the learning processes involved in the task. As students regulate and discuss strategy during dialogic argumentation tasks, they are engaged in building metacognition via the three elements identified by Järvelä et al. (2015).

One aspect of metacognition that seems to work particularly well within the context of dialogic collaboration is that of evaluation. Both the evaluation of ideas in general and the evaluation of the strengths of various arguments have been shown to develop in students who engage in dialogic collaboration (Kuhn & Crowell, 2011; Newell et al., 2011). As is the case with argumentative development, this growth in evaluative skills might be due to the nature of discourse. As students talk (and argue) with others, they become more and more skilled at viewing arguments critically. Thus, the dialogic nature of the activity leads to enhanced evaluative skills. This conclusion is supported by evidence that the more time students spend within a dialogic argumentation curriculum, the better they get at metacognitive dialogue (Kuhn et al., 2013). Thus, the more they talk to others, the better they become at this type of evaluative skill.

Students' ability to evaluate the ideas of others has been developed via dialogic collaboration (Frijters, ten Dam, & Rijlaarsdam, 2008; Newell et al., 2011; Zillmer, 2016). This makes sense, as students learn to weigh options presented by others. In particular, the longer students are exposed to dialogic argumentation, the better they become at evaluating the validity and strength of arguments and evidence (Kuhn et al., 2013; Zillmer, 2016). Dialogic curricula have been shown to increase metacognitive skills (e.g., Kuhn et al., 2013; Xiao & Carrol, 2013; Zillmer, 2016). Partly, this can be attributed to the same phenomenon occurring repeatedly

during the growth of students' argumentative skills, namely, the exposure to and analysis of others' viewpoints. As students become aware that not everyone thinks (or learns) the same, they become more aware of what strategies work best.

Students also build upon one another's metacognitive understanding. For example, Zillmer (2016) found that students engaged in dialogic argumentation with a partner engaged in "metatalk," about the discussion (as opposed to topic-related discussion). Students discussed strategy, evaluated the strength of arguments (both their own and others'), and regulated their learning experience through discussion with their partners. This meta-level discourse became stronger with time and was even more pronounced in students who kept the same partner over time, versus those who changed partners. This result may reflect the way that students are able to continuously build upon one another to help scaffold their own (and their partner's) metacognitive development.

Research has also shown that the more two students work together, the more meta-talk they engage in. Zillmer (2016) compared dyads who worked together consistently to construct arguments against other side pairs to those dyads who switched partners. She found that the students who stayed with their argumentative partner spent more time engaged in metacognitive-focused discourse, and developed their metacognitive skills together as they helped each other to regulate their discussion. This could be a result of a growing level of comfort with a specific partner.

From these research findings, we can see that metacognitive skills develop in tandem with argumentative skills. Dialogic collaboration can help support metacognitive development, in large part because of the constructivist nature of the discourse. That is, students engaged in constructing strategies (especially with regards to metacognition and regulation) through

dialogue with other students begin to see how to utilize those strategies on their own. They become fluent in meta-talk, which then feeds into their thinking about others' arguments, thus making them better at argumentation.

The Current Study

Most of the research on the effects of dialogic activities on argumentative writing development has focused on non-writing dialogic activities. However, dialogic writing activities may provide further benefits to students, as students develop the argumentative mindset described above in a near-transfer situation. That is to say, if collaborative non-writing activities help students develop argumentative writing skills, how much more so might collaborative writing activities?

Early research suggests that collaborative writing might have a positive effect on the development of the “other” as students write (Gélat, 2003; Kuhn et al., 2016; Newell et al., 2011). For example, collaborating on writing activities help students become more aware of and adept at writing for an audience (Gélat, 2003).

In addition, the research discussed earlier about the benefits of sociocognitive conflict within collaboration suggests that not just collaborative writing, or even dialogic collaborative writing, but perhaps dialogic writing activities that incorporate disagreement could lead to dense collaboration, and as a result, argumentative writing gains (Howe, 2010; Mercer, 2013; Mercer & Howe, 2012; Newell et al., 2011; Skoumios, 2009).

The current study explores the benefits of disagreement-based collaborative writing. It builds upon the dialogic argumentative curriculum designed by Kuhn (Kuhn & Crowell, 2011; Kuhn et al., 2016, 2013). A summary of the curriculum can be found in the methods section. This curriculum involves multiple opportunities for dialogic engagement with others and has

been shown to improve argumentative writing as well as dialogue (Kuhn & Crowell, 2011; Kuhn et al., 2016, 2013). But none of the collaborative activities in this curriculum include writing collaborations, leading to the question of whether collaborative writing could further benefit students.

Three research questions were asked to examine both the process and outcomes of collaborative writing within an argumentative curriculum. The research questions and their associated hypotheses are:

Research Question 1: Does collaborative writing lead to better argumentative writing development than individual writing does?

Hypothesis 1a. Collaborative writing enhances individual argumentative writing.

Hypothesis 1b. Differences between collaborative and individual groups increase most in the later part of the intervention (topic 4) compared to the earlier part (topic 1).

Hypothesis 1c. Students in the collaborative writing condition will show greater gains from pretest to posttest assessments of argumentative writing than will those in the individual condition.

Research Question 2: How do collaborating pairs use metacognitive discourse in collaborative pre-writing?

Hypothesis 2. The benefit of collaboration will be reflected in increased metacognitive discussion during the collaborative writing activity.

Research Question 3: Do students use ideas acquired during collaborative writing in subsequent individual essays and, if so, how?

Hypothesis 3a. In their final, individual essays, students will use their own ideas and incorporate their partner's ideas that support their own side more than they will use partner ideas that support the partner's side.

Hypothesis 3b. Students will use more partner ideas in their independent essays in later (topic 4) earlier (topic 1) during the dialogic intervention.

CHAPTER 2: METHODS

Participants

Students from the two seventh grade classes were enrolled in the argumentation curriculum developed by Kuhn and used extensively in her research (Kuhn & Crowell, 2011; Kuhn, Zillmer, Crowell, & Zavala, 2013). The curriculum was implemented in place of students' regular, 50-minute social studies class twice a week throughout the course of the school year. The other two social studies class sessions each week were dedicated to regular American history content, separate from the argumentation curriculum. This regular social studies curriculum did not overlap with the argumentation curriculum and was taught by the regular social studies teachers for the classes. Both classes had similar activities, including some writing assignments, in their social studies classes throughout the year.

In contrast, the argumentation curriculum was taught by the author of this study and another doctoral student at Teachers College, neither of whom taught in the school regularly. The author of the current study had worked for three years prior in schools with this curriculum, whereas the second doctoral student was new. At the beginning of the year, the second teacher served as a helper and facilitator, while the author of this study took the lead teacher role. As the year progressed and the second teacher became more comfortable with the curriculum, she began to take a lead teacher role on some days. Teacher effects were controlled for by having the lead teacher for one group on each day also serve as the lead teacher for the other group.

The school students attended was a public inner city middle school in a low-income neighborhood. While the school was not able to release most demographic data on the specific students in the study, the demographics for the school as a whole are available. The school's

racial and economic make-up includes 96.3% of students who are Hispanic or black, 29.6% who are English Language Learners (ELLs), and 22.1% with some type of disability. Further, 84% of students at the school are eligible for free lunch.

Despite not having access to the demographic data of the specific students in the classes, observation of the participants suggested that the racial and economic demographics closely mirrored those of the whole school. Similarly, the number of students who received accommodations due to an IEP seemed consistent with the figure of 22.1% of students with some type of disability. Gender demographics were easier to obtain through observation, and the current study included 10 girls and 16 boys in the intervention group and 16 girls and 13 boys in the comparison group.

Data were analyzed on 26 students in the treatment (collaborative) group and 29 students in the comparison (individual) group. On the whole, students in both groups were several grades below grade level in reading and writing, a common occurrence at the school. Attendance varied throughout the year, and some students were excluded from the final data sets because they missed either an interim or final essay writing activity. In addition, two students in the comparison group were eliminated because they were not able to complete at least one writing assignment due to language barriers. Students addressed a sequence of four topics across the year. Figure 2 shows the number of students in each group that participated in each topic (the two students eliminated from the comparison group due to language barriers are not included).

For a student to be included in the final data set, they had to have been present for both the interim and final essays for at least three of the four topics. Thus, those who were present for fewer than three topics were excluded from the final data set. In addition, any student present for

an interim but not final essay in any topic was not included in that topic's data set. The final number of students in each group for each topic can be seen in table 1.

Figure 2. Number of students present for each topic.

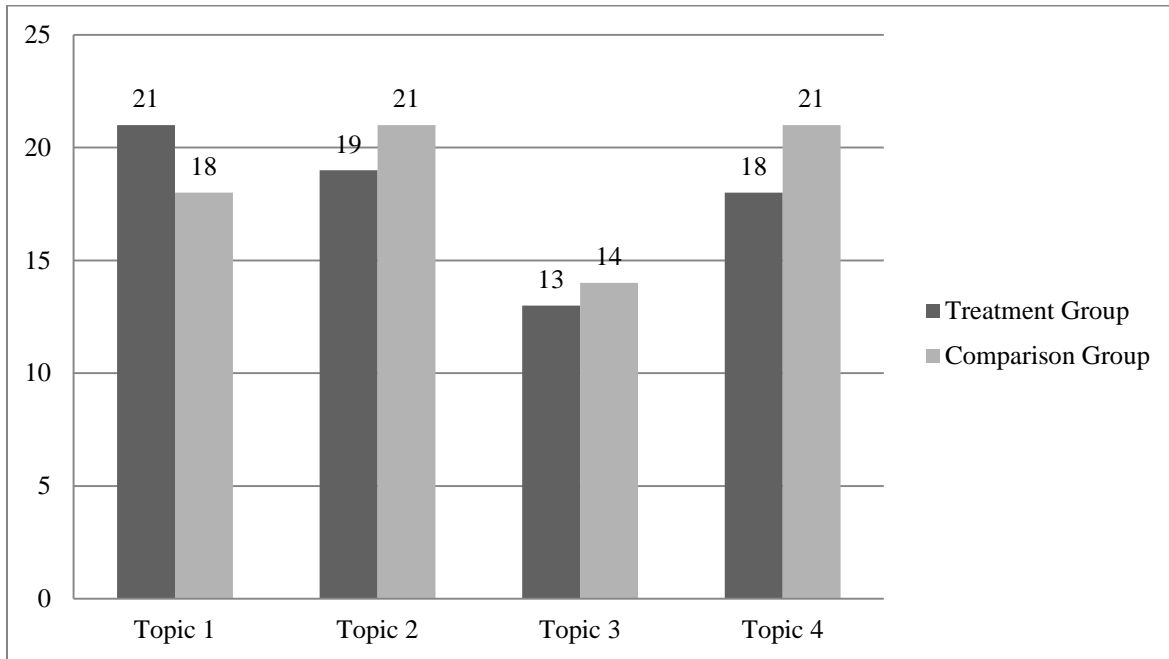


Table 1. Number of participants for each group, by topic.

Topic	Treatment Group	Comparison Group
1	n=19	n=16
2	n=19	n=21
3	n=11	n=13
4	n=17	n=18

The students in the comparison group were part of an English as a Second Language (ESL) classroom, and varied in their English proficiency, although all students were able to communicate with other classmates in their native language (Spanish or Arabic) and all were able to communicate moderately well in English. In contrast, the students in the treatment group

were in a mainstream classroom, although many of them were English Language Learners (ELLs), and most were at the same level of proficiency in English as those students in the comparison group who were not excluded due to language barriers.

To examine possible differences in the two groups due to their English proficiency, 19 treatment group students' and 16 comparison group students' pre-instructional essays, given on the first day of topic 1, were compared via a Poisson regression analysis. A robust description of the coding of the essays is presented later, but in short: statements within the essay are coded as unique idea units, repeated ideas, or no reason. Idea units are the number of unique idea units in the essay, repeated ideas are ideas previously mentioned in the essay, and no reason statements are coded as such because they are off-topic, nonsensical, or a simple statement of opinion with no reasoning. This follows the tradition of looking at the ideas behind the writing summarized in Newell et al. (2011), instead of focusing on more surface characteristics of writing, such as word count or grammar.

Idea units are further divided into subtypes. Again, these are described in more detail subsequently. The three types of idea units found in the pre-instructional essays were supportive of the writer's position (my-side positive, or M+), oppositional to the other side (other-side negative, or O-), or supportive of the opposite side of the writer's position (other-side positive, or O+). These types provide a way of examining the thinking behind the writing, and have been used extensively in conjunction with the curriculum described in this study (Kuhn et al., 2016).

Poisson regression was chosen as a statistical analysis for the pre-instruction essays because the data analyzed are count data, and not continuous data. Whereas linear regression requires dependent variables to be continuous, Poisson regression analysis allows for count data.

The analysis of the groups' essays showed similar performance across the two groups. Perhaps surprisingly, the comparison group (with English language learners) produced more idea units on average in these pre-topic essays than the treatment group (2.13 vs. 1.11, see table 2), working against the study's hypothesis. However, this difference in group means for overall idea units was not significant.

Further analysis was done on different types of idea units. Neither group wrote any negative reasons about their own side, which is not surprising at the pre-instruction phase. The means for other types of idea units appear in Table 1 (explanation of the different idea types is presented below and in table 9). The Poisson regression did not show any significant differences between those means for any of the types of idea units.

Table 2. Mean number of different types of idea units on pre-topic 1 essays by group.

Idea Type	Group	Mean
Pre-essay M+	Treatment	.85
	Comparison	1.19
Pre-essay O-	Treatment	.20
	Comparison	.75
Pre-essay O+	Treatment	.05
	Comparison	.19
Pre-essay M-	Treatment	.00
	Comparison	.00
Pre-essay Repeat	Treatment	.00
	Comparison	.19
Pre-essay However	Treatment	.00
	Comparison	.06
Pre-essay Total Idea Units	Treatment	1.11
	Comparison	2.13

The proportion of idea units, repeated ideas, and no reason statements for the topic 1 pre-essays is in table 3. These proportions seem to support the Poisson regression findings that the comparison group's essays were slightly better at the beginning of the year, with a relatively

larger proportion of unique idea units and smaller proportions of repeated and no reason statements, compared with the treatment group.

Table 3. Proportions of topic 1 pre-essay statements.

	Treatment Group	Comparison Group
Idea Units	.553	.739
Repeated Ideas	0.00	.065
No Reason Statements	.447	.261

Table 4 shows the proportions of specific types of idea units in the pre-topic 1 essays for each group. As with the Poisson, these proportions show that the treatment group had a higher proportion of my-side positive (M+) statements and a lower proportion of other-side negative (O-) and other-side positive (O+) statements. Thus, their pre-instructional essays were, on the whole, not as strong as the comparison group's.

Finally, table 5 compares the percentage of students who used the different types of idea units in the pre-topic 1 essay. Here, we see that not only were class means and proportions slightly better in the comparison group, a larger percentage of students in the comparison group were using the different types of idea units. Of particular interest is the larger percentage of students using other-side negative (O-) ideas and *however* phrases, as these are advanced practices, representing dual representation and integrative arguments respectively.

Table 4. Proportions of types of idea units in topic 1 pre-essay.

	Treatment Group	Comparison Group
M+	.762	.559
O-	.190	.353
O+	.048	.088

Table 5. Percentage of students who ever used each type of idea unit in topic 1 pre-essay.

	Treatment Group (n=19)	Comparison Group (n=16)
M+	73.7%	81.3%
O-	21.1%	75%
O+	52.6%	18.6%
Repeated Ideas	0%	18.6%
However	0%	6.25%

A larger percentage of students in the treatment group used other-side positive (O+) ideas. While the use of O+ ideas is an advanced practice, it is important to note that no students in the treatment group used a *however* phrase in their pre-instructional essay. Including O+ ideas without countering them is far less advanced than incorporating them into a *however* phrase (as seen below).

Overall, the data offer a picture of the two classes that suggests the treatment group was no more advanced than the comparison group at pretest. Observations of the two classes reinforced the conclusion that there was little difference between the two. Often, in the district where the study was done, students born in the United States are assigned to general education classes when their only language at home is not English. As a result, general education classes can look very much like ELL classes, and that was the case here. With a few exceptions (most of which were eliminated from the data set, as described above), the ELL students did not appear to be behind the students in the other class in terms of language. Both groups had a mix of students who were quite good at reading, writing, and conversing in English, and students who struggled with one or more of those skills.

One additional note about a possible difference between the groups concerns behavioral issues. While both classes had students who were well-behaved and some who struggled with behavior, the ELL class on the whole was a more attentive and well-behaved class. Behavioral

issues with the treatment group regularly cropped up, often resulting in time being spent going back over the instructions several times when students weren't listening.

Both the analysis of the pre-instructional essay and the observation of the behavioral dynamics of the classes, then, suggest that the comparison group began the year with some small advantages, despite being an ELL class. Balancing those advantages (slightly better first essay, better behavior) against the fact that it was an ELL class, and adding in the fact that many of the students in the treatment group were also ELL students, equivalency between groups at the beginning of the year was assumed.

Procedure

As mentioned above, the current study was done within the framework of a yearlong argumentation curriculum developed by Kuhn (Kuhn et al., 2014), presented twice a week during the students' regular social studies class. Here we examine the curriculum in general and the writing component (and its manipulation in the current study) specifically.

The curriculum is divided into topics, during which students choose a side and develop arguments and counterarguments about the issue, culminating in an in-class debate and a final persuasive essay. The final essay is written individually during one class period.

Table 6 provides an overview of activities by class period for each topic. It is worth examining in detail in order to understand the learning environment in which the current study was conducted. The year was divided into four 15-session units centered on a political or social topic, such as whether kidney sales should be legal or whether juveniles should be tried as adults. Table 7 includes the topics and prompts for each unit. These topics and prompts have been used in previous studies with similar populations, and have been found to be successful at engaging students and allowing for rich debate in the classroom (Kuhn et al., 2014, 2016).

Table 6. Description of argumentation curriculum daily activities for each topic, with essay days shaded.

Class Session(s)	Type of Work	Activity
1	Independent	Straw poll & pre-topic essay
2	Small group	Reasons generation
3	Small group	Reasons ranking
4-9	Same-side pair	Dialogues with other side pair (reflection sheets)
10-11	Small group	Showdown prep (summary reflection sheets)
12	Treatment: other-side pair Comparison: independent	Interim essay
13	Whole class	Debate
14	Whole class	Debate debrief
15	Independent	Post-topic essay

Table 7. Topics and prompts for each unit within the current study.

Topic Number	Topic	Prompt
1	Soda tax	Should a tax be charged on soft drinks and that money go to lowering the price of healthy foods?
2	Animal research	In medical research labs across the country animals are used to test new medications. This testing makes it possible to develop new medications that can save human lives. Should companies be allowed to conduct this research upon animals?
3	Juvenile justice	Teens who commit serious crimes may be tried and sentenced in the adult court system. Or they may be tried in a court system for juveniles. Which is better?
4	Kidney sales	Humans have two kidneys. They need at least one working kidney to live. If both their kidneys stop working, it is possible for them to get a transplanted kidney from someone who is willing to give up one of their kidneys. But new kidneys are in short supply; people needing them often have to wait years. A poor couple heard that a man will pay them \$10,000 to sell him a kidney to save the life of his 12-year-old son. The husband wants to do it because they need the money, but the wife is unsure because it would be her kidney they would sell to the man and she is afraid it could cause problems. Should people be allowed to take money for their kidneys or should this be forbidden?

As the curriculum was delivered twice a week, interspersed with school holidays and schoolwide events, the 60 sessions (15 sessions for each of four topics) were distributed across the entire school year. School holidays and events delayed some sessions, so each topic did not

take up the exact same number of weeks. However, the four topics were relatively regularly spaced across the year, with two topics finished before the winter recess, a third finished before the spring recess, and the fourth and final topic taking place between spring recess and the end of the year. In all cases, delays of sessions impacted both classes, so the two groups were always in sync across the year.

Each topic started with a straw poll and pre-topic essay, which students completed individually during one 45-minute class period on the first day of the topic. On a piece of paper, students were asked to choose a side (pro, con, or undecided), rank how certain they were of their opinion, and then to write a justification of their position. This justification was the pre-topic essay. Appendix 1 shows the straw poll and pre-essay handout students were given on the first day of the first topic.

Using the straw polls, the researcher divided each class into two groups (pro/con) of approximately the same size. To accomplish uniformity of group size, students who indicated they were undecided were distributed to the smaller group. If the groups were still uneven in size, the students in the larger group who indicated that they were the most unsure of their position were assigned to the opposing (smaller group) side.

The redistribution of students generally did not require moving many students other than those who were undecided. In the first topic, for example, the treatment group consisted of 11 students who favored pro, 9 students who favored con, and 4 students who marked undecided. In the comparison group, 11 students marked pro, 5 students marked con, and the remaining 3 students marked undecided. For both groups, those who marked undecided were distributed in such a way to make the groups as equal as possible in size, while those who marked pro and con were left on the side they chose.

During the second class session, students worked in small groups of four to six students (two pro groups and two con groups) to create reasons in support of their side. Each student was asked to write reasons for their side on index cards, and then the groups came together to share each other's reasons and group them into similar piles. Finally, the group stapled the similar reasons together, so that by the end of class, each small group had a card (or more than one card stapled together) for each reason in support of their side.

At the third session, students worked in their small groups again. This time, they looked through the reasons produced by their group the previous class and grouped them according to how strong each reason was: so-so, good, or best. By the end of the period, each group had a ranked set of reasons in support of their side, from their strongest to their weakest reasons.

Sessions four through nine were conducted as six consecutive, technology-enabled dialogues during which students worked in same-side pairs and debated using a Google document with an opposite-side pair during each class. Students worked with the same partner throughout the dialogues, while the dyad they debated rotated each day. Each chat session began with a different side (i.e., session 1 started with the pro side, session 2 with the con side, and so on). During the dialogues, same-side dyads were asked to respond to the opposing pair, working on producing counterarguments. In addition, at each session, students were asked to complete a reflection sheet on which they reflected on their own and the opponents' arguments and counterarguments, and how they could strengthen (or weaken) them. Appendix 2 shows the two reflection sheets, one focusing on own side and the other on opponent's side; these were alternated over sessions.

Following the dialogues, students returned to their original small groups for two sessions to prepare for their final, class-wide debate. The first of these debate prep sessions involved

having students think about how to best counter the other side. In this session, students were given “Other” reflection sheets that had been filled out during the earlier tech-enabled dialogue sessions (see Appendix 2). The groups were asked to sort the reflection sheets based on the argument made by the other side.

Once the reflection sheets were grouped based on the other side’s argument, students analyzed the responses to the argument listed on the reflection sheet(s), as well as any other counters they could think of. As a group, they chose the strongest counterargument(s) and completed a summary reflection sheet with the best counter(s) listed.

The second of the two Showdown prep sessions followed a similar pattern. In this case, the small groups worked with the “Own” reflection sheets filled out during the online dialogues (see Appendix 2). They sorted them based on the initial argument presented by their own side, then identified the strongest counters from the other side, and their strongest counters (i.e., rebuttals) to the other side’s counters. The students then completed a summary reflection sheet, as they had done at the previous session. Both summary reflection sheets can be seen in Appendix 3.

Following the two debate prep sessions, students spent a class period writing an interim topic essay. This was the only part of the curriculum that varied between groups, with one class (the comparison group) working independently, and the other (the treatment group) working with an opposing-side partner to construct the persuasive essay. The essay worksheet for topic 1 is in Appendix 4; it was the same for both the treatment and comparison groups. For both classes, instructions were read aloud to the class, and students were given the rest of the 50-minute period to work on the essay. The oral instructions to each class can be found in Appendix 5, and are discussed in detail below.

The next class period was the full-class debate. During the debate, one student at a time from each side came to the front of the classroom and spent two minutes debating verbally. Debaters or their small group members were able to call a 1-minute time out during which the small groups could huddle and strategize, with no limits on the number of “huddles” a group could have. The class meeting following the debate was devoted to a debrief, wherein students were able to see and discuss recorded and transcribed segments from the debate and evaluate them.

At the final session, students individually wrote a final position essay on the topic, which served as a post-topic essay. The sheet on which the students wrote their essays was the same as the one used for the interim essay (see Appendix 4). Oral instructions for this essay are included in Appendix 5 and discussed in detail below.

This sequence repeated itself four times throughout the year, resulting in four topics covered in the same way each time. There was one exception, an anomaly in the collection of data for topic 3 that impacted both groups. Due to a school-wide event, the students were not in class on the day they were scheduled to write the final essay for topic 3. Instead, students were given the essay to write as homework. As a result, a number of them did not return the final essay, reducing the n for the present study. The essays returned were from a self-selected group of students. Even so, the essays that were returned were shorter than other essays, indicating that perhaps students did not spend as much time writing the essay at home as they might have in class. As will be seen, the data for topic 3 do not fit the trends seen across the other topics. However, it is left in this study to give a full view of what was collected.

The Essays: Pre-, Interim, & Post-Topic

While it is important to understand the argumentation curriculum as the educational context in which student essays were written, for the purposes of this study, there were three class periods of major interest in each topic: day one, during which students wrote their individual pre-topic essays; day 12, during which students wrote the interim essay individually or with a partner; and day 15, the final day for the topic, when students wrote their individual post-topic essays. As the curriculum cycled through four topics, those same three writing sessions occurred four times across the span of the year, for 12 essays total. Table 8 identifies the two groups, comparison and treatment, and how they worked (independently or collaboratively) for each of the three essays. As the table shows, both groups wrote individual pre-topic and post-topic essays. The sequence was repeated for each topic, thus a total of four times.

Table 8. Summary of the research design.

Class	Pre-Topic Essay	Interim Essay	Post-Topic Essay
Treatment Group	Individual	Collaborative	Individual
Comparison Group	Individual	Individual	Individual

The interim essay was the only procedural difference between the groups: the comparison group worked independently on this essay, while the treatment group worked with a partner from the opposing side (i.e., a “pro” student was paired with a “con” student). Students in the treatment condition were assigned an opposing-side partner through a process meant to

randomize student pairings as much as possible. The researcher created a list of students on the pro side, and then blindly drew names of con-side students with which to pair them.

Four sets of oral instructions are included in Appendix 5, representing one set each for the pre-topic and post-topic essays, which were identical for both conditions, and two sets for the interim essays, one for the comparison group and one for the treatment group. Because the actual handout given to the students for their interim essays was the same for both groups, the oral instructions were the only difference (other than process) between the groups. As a result, it is important to specify the oral instructions for the interim essay in more detail.

The instructions began the same. The students were told that the purpose of the interim essay was to prepare for the debate by helping them organize their thoughts and arguments. Both groups were also told that they would have all class period to work, and they were given instructions regarding filling out the top of the paper (see Appendix 4). However, there were two additional instructions that the collaborative group was given that the individual group was not. The first involves the process of writing, and the other the technology used.

The process instructions explain the condition and expectations for collaboration. The researcher explained that the students had been paired with a student from the opposing side. The instructor also told the students that they would work together with their partner, and that both partners would have to agree before anything was written on paper. This statement was included to prevent pairs from having one student do all the work, while the other student sat back, although it had the added benefit of encouraging a more lively debate.

Besides explaining how the students would be collaborating, the instructor also included a word about the technology to be used during collaboration. Specifically, each pair was given a handheld audio recorder to record their conversations as they collaborated on the joint essay.

These recordings were later analyzed, as discussed below. The instructor explained that they would be expected to keep the recording going the entire time and not to stop it at any time.

Being seventh graders, many of the pairs disobeyed the instruction to keep the recording going. As a result, some of the conversations were missed, although thanks to the diligence of the instructors in the classroom, who continually restarted stopped recordings and reminded students to keep them going, most of the conversations were preserved on the recordings.

Scoring the Final Topic Essays

Each student was assigned a numeric ID produced by a random number generator, and the final essay written at topic conclusion associated with that student was linked with that numeric ID upon transcription, not with the student's name. (The names and IDs were listed in a separate document.) After transcribing the essays, they were put in order of the ID associated with each essay, which meant that the groups' essays were mixed. The essay documents were then saved, closed, and left alone for five weeks, at which point the researcher returned to code them. Although this does not produce a completely blind procedure (as at one point the researcher did see the names associated with the IDs and with the essays), by the time the essays were coded, the researcher retained no conscious memory of which essay belonged to which student (or group).

Coding of the essays was based on a coding scheme used by Kuhn and Crowell (2011) and others in connection with the argumentation curriculum described. The coding involves first dividing the essay into unique "idea units." Idea units consist of a phrase or sentence that expresses a single idea. This follows the tradition of Walton (1998), who views argumentation in a dialectical framework, and Newell, Beach, Smith, and Vanderheide (2011), who discuss

argumentation as a dialogue consisting of ideas built upon one another. Thus, each distinct idea unit is a statement that can carry the essay forward.

The first time a reason appears in an essay, it is coded as an idea unit; if the idea is repeated, all subsequent repetitions of that reason are coded as a repeated reason. Also coded were “no reason” units—these are off-topic, nonsensical, and/or a claim with no reason given.

Remaining idea units were coded one of four types (see Table 9). Idea units scored as M+ are arguments that are supportive of the writer’s position (i.e., my-side positive reasons). An O+ unit is a unit that supports the opposing side’s position (i.e., other-side positive reasons). M- units are ones that weaken the writer’s position (i.e., my-side negative reasons), and O- units are those that weaken the other side’s position (i.e., other-side negative reasons).

Additional coding of adjacent idea units was performed and coded as *however* clauses. They consist of an adjacent pair of idea units (usually O+ followed by O- or M+), explicitly linked by words such as “however” or “but” that indicate their connection. Because of the grammatical construction of *however* clauses, they can be called “this-but-that” clauses.

Table 9. Types of idea units used during scoring.

Type	Explanation
M+	My-side positive: reasons supportive of the writer’s position
M-	My-side negative: reasons against the writer’s position
O+	Other-side positive: reasons supportive of the position opposite of the writer’s position
O-	Other-side negative: reasons against the position opposite of the writer’s position
<i>However</i>	This-but-that: two opposing idea units linked together with “however” or “but”
Repeated	Idea already presented in the essay
No reason	Statement with no reason given (such as a statement of opinion or one that is off-topic)

Kuhn et al. (2016) found that students generally begin arguing by simply presenting a series of my-side positive (M+) reasons. As students progress, they begin using a mixture of M+

and O- (other-side negative) reasons, a dual representation of consistent statements. Later, inconsistent statements (O+ and M-) may emerge. Finally, the most advanced type of argumentation involves integrative arguments, namely the *however* clauses described above. This coding system, then, provides a way to focus on the thinking underlying the writing, as opposed to surface measures such as word count or grammatical construction.

In order to establish inter-rater reliability, 10 essays from the first topic were coded separately by the two researchers who served as co-teachers of the curriculum. Analysis of these 10 essays yielded an agreement of 94% for segmenting into idea units, and an inter-rater agreement of 91% for assignment of category. The remaining essays were scored solely by the author.

Scoring the Collaborative Recordings

In addition to the essays, the recordings of the verbal dialogue that occurred during the collaborative writing process in the treatment condition were also coded. In this case, each utterance was scored based on a coding scheme used by Zillmer (2016), which looked at various metacognitive and non-metacognitive utterings. Table 10 lists Zillmer's types of dialogue.

As with the coding of the essays, 10 transcripts used in Zillmer's study were coded, and an agreement rate calculated between the present author and Zillmer's scoring, resulting in an inter-rater agreement of 87%. The recordings from the present study were then coded by the author.

Table 10. Types of dialogue used in collaborative writing recordings.

Dialogue	Type	Definition	Example
Planning	Metacognitive	Strategizing and decision-making	<i>“Do you think that the taxes should go down?”</i>
Regulating	Metacognitive	Focusing attention on changing behavior	<i>“All right, ready? We’re getting sidetracked.”</i>
Evaluating	Metacognitive	Assessing an argument	<i>“I don’t think it will make a difference.” “It kind of does, though, make a difference. Technically.”</i>
Soliciting help	Metacognitive	Asking for partner to assist	<i>“Wait, what was the homeless person argument?”</i>
Off-topic	Non-metacognitive	Talk centered on something other than the task	<i>“You stupid, I’m not.”</i>
Content discussion	Non-metacognitive	Non-planning discussion of the topic	<i>“Sodas are less expensive, but if there’s the tax people can buy the healthy foods they want.”</i>
Technology	Non-metacognitive	Raising issues with the recorders	<i>“Wait! We wasn’t recording this whole time.”</i>
Surface level division of labor	Non-metacognitive	Deciding who is doing what, with no strategy involved	<i>“I always write. Go ahead, get your pencil.”</i>
Dictation	Non-metacognitive	Telling the other person what to write	<i>“And you could write this: ‘We also think that you should say all sodas are destroying because it’s not healthy for you.’”</i>
Talk to teacher	Non-metacognitive	Commentary directed at the teacher	<i>“Wait, can I write this, Ms. Boyd?”</i>
Reading from paper	Non-metacognitive	Reading out loud what has been written	<i>“Um...it says, ‘Dear Mayor DeBlasio, I think soft drinks should be taxed because it pays, it pays, um, our teachers and...’”</i>
Announcing status	Non-metacognitive	Explaining what is happening in the group	<i>“Ok, so he’s writing.”</i>
Writing out loud	Non-metacognitive	Saying what is being written, while it is being written (commentary by student writing)	<i>“I’m gonna start with ‘I believe that’ because it’s always about ‘I believe that...there should be a tax, a tax on soda because, because people, certain people want to go on diets, but they cannot afford the healthy foods.’ I’m gonna change that – ‘want to eat healthy but cannot afford, but cannot afford, but cannot afford the healthy foods and this tax will change that. If there is a tax, this tax will also...um...this tax will also...’”</i>

As mentioned above, during the collaborative process, each dyad was given an audio recorder with which to record their conversations. The recorders had labels to identify each one from the others, from B01 to B12. During the collaborative writing session, the ID of the recorder used by each pair was noted. Recordings were transcribed with the pair ID (B01-B12), and the generic titles “Student 1” and “Student 2” for each partner (based on which partner spoke first on the recording), listed as the only identifying information. Similar to the essays, the transcripts were then left alone for four weeks, at which point the researcher scored them with very little conscious memory of which voices were associated with which recordings.

Any collaboration, especially collaboration between 7th graders, is likely to create some friction. The recordings revealed some affective issues regarding partners getting along with each other or “liking” each other. However, the subjects were not with the same partner for all four topics, and therefore the amount of dislike experienced during the collaboration were most likely randomly distributed across the subjects and topics. That is, any student was likely to have a personal dislike (or no personal dislike) for their partner at any given topic.

CHAPTER 3: RESULTS

Hypothesis 1a. Collaborative writing enhances individual argumentative writing.

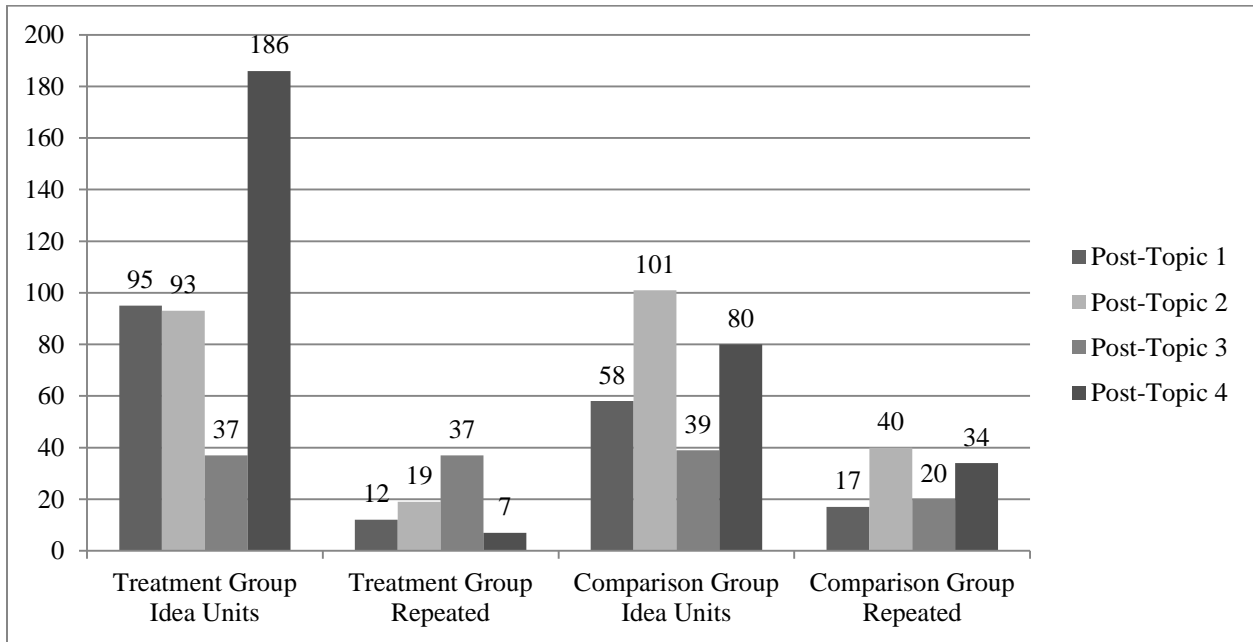
Hypothesis 1b. Differences between collaborative and individual groups increase most in the later part of the intervention (topic 4) compared to the earlier part (topic 1).

Hypothesis 1c. Students in the collaborative writing condition will show greater gains from pretest to posttest assessments of argumentative writing than will those in the individual condition.

To examine the effects of the dialogic writing activity on students' argumentative writing, post-topic essays for each of the four topics were examined. At the end of the first topic, students in the comparison group were writing individual essays with a mean of 167.79 words, while the treatment group had a mean of 163.45 words. By the end of the second topic, the comparison group's essays had a mean of 135.14 words, while the treatment group's essays had a mean of 121.79 words. The post-topic essay for topic 3 was done for homework, as discussed above, and both groups produced shorter essays. The comparison group had a mean of 102.38 words in their third topic essays, while the treatment group had a mean of only 69.27 words. For the first three topics, then, students in the comparison group were writing more words. However, the post-topic essay for the final topic saw a different trend: the comparison group had a mean of 217.78 words, while the treatment group had a mean of 277.88 words.

The number of words isn't the only measure of success, though. The focus of the current study was on the thinking underlying writing, and to get at that, it's important to examine the number and types of ideas in the essays. Figure 3 shows the total number of unique idea units and the total number of repeated statements for each condition across the four topics.

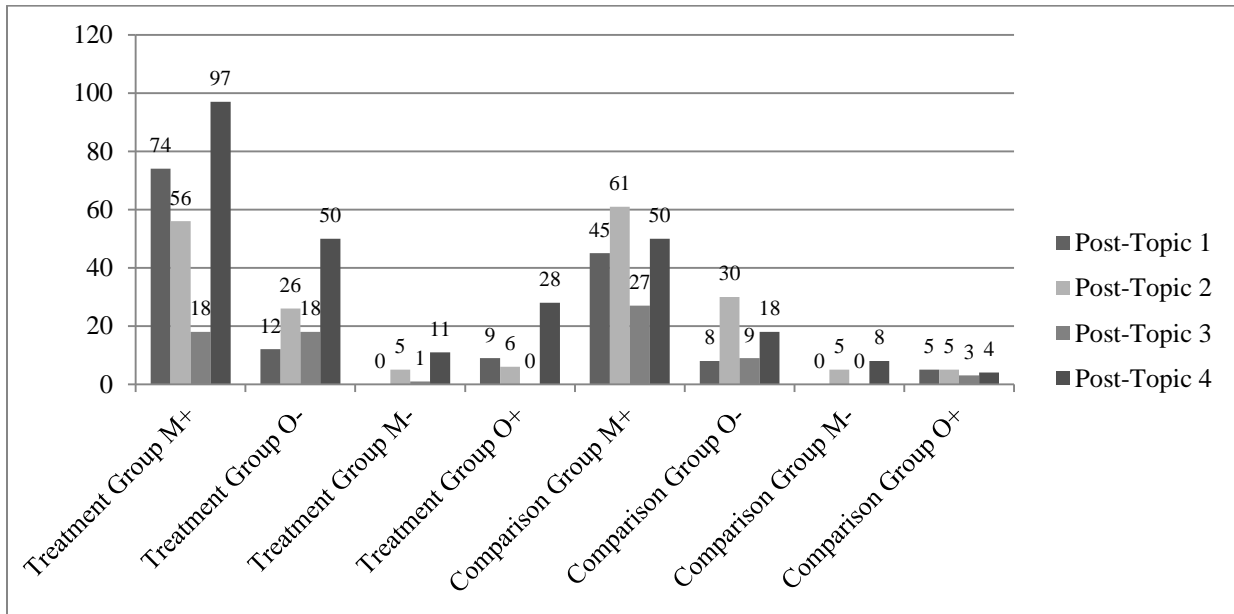
Figure 3. Total number of idea units and repeated statements for each group for each of the four post-topic essays.



While each group grew from the first essay to the last, in terms of the number of unique idea units, the growth of the treatment group was more dramatic. Further, the comparison group’s number of repeated statements increased from topic 1 to topic 4, whereas the treatment group used fewer repeated statements by the end of the year.

The raw data for the different types of idea units further shows growth patterns for the different groups (see figure 4). From topic 1 to topic 4, both groups increased the number idea units that are consistent with their position, either by supporting the writer’s side (M+) or weakening the other side (O-). However, the treatment showed more dramatic gains than the comparison group, particularly for O- reasons. Further, when examining the growth for ideas that challenge the writer’s position (M-) or strengthen the opposing position (O+), the treatment group showed more by topic 4 than topic 1. In contrast, the comparison group showed a smaller increase in M- reasons and stayed about the same in O+ reasons.

Figure 4. Number of different types of idea units for post-topic essays for treatment and comparison groups.

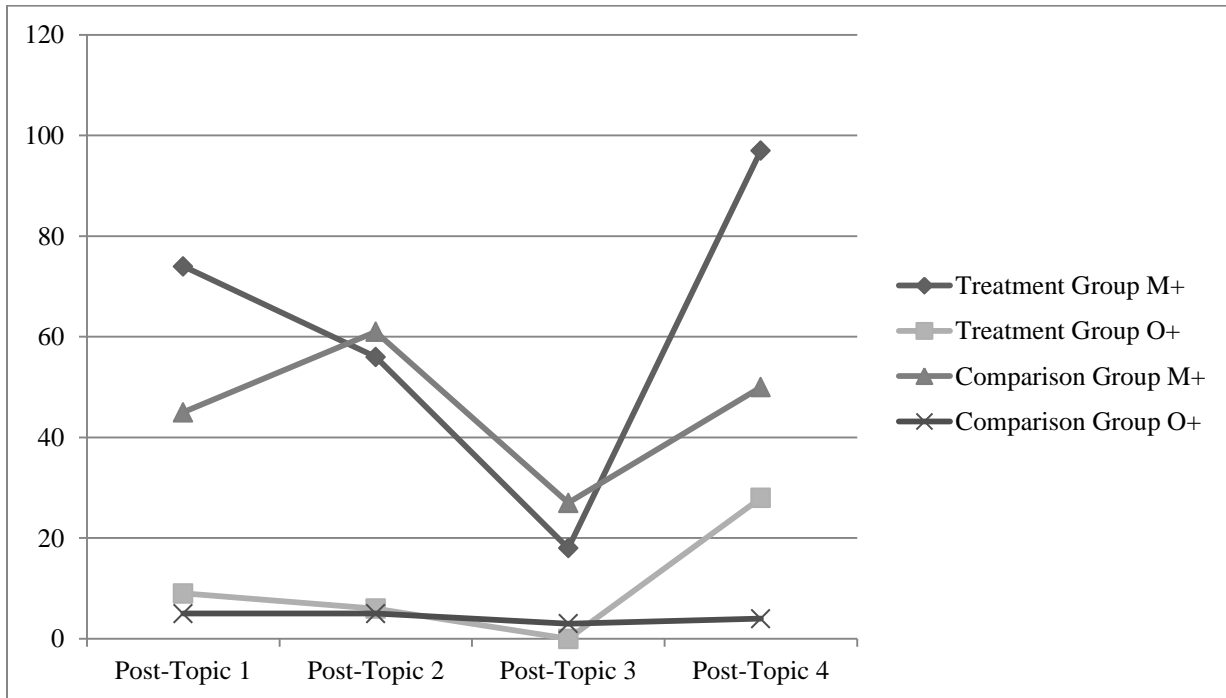


Of particular interest is the shifting way that students use M+ and O+ reasons.

Specifically, using more M+ reasons over time represents less growth than using more O+ reasons. Thus, we would expect true growth in argumentation to involve not only an increase in the number of M+ ideas but also an increase in the number of O+ ideas by the end of the year.

Figure 5 shows the number of M+ and O+ reasons used by each group in the post-topic essays for each topic. Both groups used more M+ reasons than O+ reasons at all points during the year. However, the treatment group showed an increase in both the number of M+ and the number of O+ ideas used by the end of topic 4.

Figure 5. Number of M+ and O- reasons used by each group in post-topic essays for each topic.



To examine these trends statistically, a Poisson regression was performed on the final, post-topic individual essay for each of the four topics, comparing across conditions the number of unique idea units, number of M+, O-, M-, and O+ idea units, number of repeated idea units, and the number of *however* compound units.

The treatment group showed more gains than the comparison group in many areas. However, these differences emerged as the year progressed. In the first topic, the difference in the mean number of idea units used in the essays approached significance. The treatment group was likely to use 1.37 (95% CI, .995 to 1.912, $p=.05$) more unique idea units than the comparison group (see table 11), but there were no other significant differences between groups (see table 12).

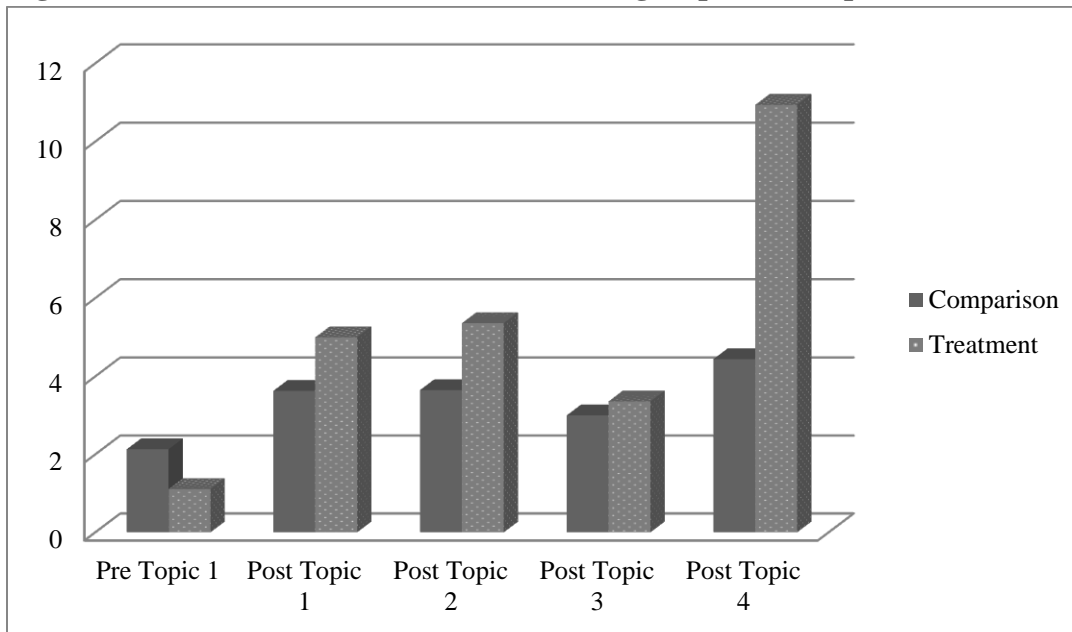
Table 11. Mean number of idea units between groups for all post-topic essays. Proportions of idea units to total statements in parentheses.

Topic	Treatment Group	Comparison Group	Difference (Treatment – Comparison)
1	5.00 (.586)	3.63 (.483)	1.37
2	5.36 (.783)	3.65 (.461)	1.71**
3	3.36 (.881)	3.00 (.574)	.36
4	10.94 (.861)	4.44 (.530)	6.50***

**p<.01

***p<.001

Figure 6. Mean number of idea units for each group across topics.



For topics two and three, there were significant differences between groups on several specific types of idea units, and a significant difference in the number of idea units overall for topic 2 (1.71 (95%CI, 1.262 to 2.304), $p < .01$), but not for topic 3 (tables 13 and 14). By the fourth and final topic, there were significant differences in most idea unit types (see table 15), as well as the number of idea units overall (6.50 (95%CI, 5.934 to 7.239), $p < .001$). (The exception was M- ideas, which remained low in both groups.) By the final topic, the difference in the mean number of idea units across groups had grown from 1.37 for the first topic to 6.50 (table 11).

Thus, by the end of the year, students in the collaborative condition were producing 6.50 more unique idea units on average than those in the individual condition.

Table 12. Mean number of different idea units by group for topic 1 post-essay.

Type	Treatment Group	Comparison Group	Difference (Treatment – Comparison)
M+	3.89	2.81	1.08
O-	.63	.50	.13
M-	0.00	0.00	0.00
O+	.47	.31	.16
Repeat	.63	1.06	-.43
However	.58	.25	.33

Table 13. Mean number of different idea units by group for topic 2 post-essay.

Type	Treatment Group	Comparison Group	Difference (Treatment – Comparison)
M+	2.89	2.71	.18
O-	2.05	.24	1.81**
M-	.32	.33	-.01
O+	.42	.05	.37*
Repeat	.37	2.76	-2.39**
However	.47	.14	.33

*p<.05

**p<.01

The individual Poisson regressions for each topic highlighted trends across topics in terms of counts. To enrich this analysis at the individual level, table 16 shows the percentage of students in each group who ever used each type of idea unit across the four topics.

Unsurprisingly, the percentage of students using at least one M+ reason in their post-topic essays was extremely high from the beginning to the end of the year. However, growth in other types of idea units can also be seen, as a higher percentage of students moved towards using O-, M-, O+, and *however* compound units. Here, too, the same trends begin to emerge: students in the

collaborative condition showed more growth in terms of the percentage of students using higher level argument forms.

For example, at the end of topic 1, the two groups were about even in terms of the percentage of students using dual representation (O-) arguments. By the end of the year, though, the treatment group had 94.1% of students using O- arguments, while the comparison group only had 50% using O- arguments. Similarly, while the end of topic 1 saw a higher percentage of students in the collaborative condition using integrative (*however*) arguments, by the end of the year, the gap between the groups was even wider. And while a larger percentage of students in the comparison group were using M- arguments by topic 4 (50% vs. 35.3%), the lower percentage of students using *however* units (33.3% vs. 76.5%) suggests that they were not using M- arguments in an integrative way.

Table 14. Mean number of different idea units by groups for topic 3.

Type	Treatment Group	Comparison Group	Difference (Treatment – Comparison)
M+	1.64	2.08	-.44
O-	1.64	.69	.95*
M-	.09	0.00	.09
O+	0.00	.23	-.23
Repeat	.18	1.54	-1.36**
However	0.00	.08	-.08

*p<.05

**p<.01

Table 15. Mean number of different idea units by groups for topic 4.

Type	Treatment Group	Comparison Group	Difference (Treatment – Comparison)
M+	5.71	2.61	3.10**
O-	2.94	.78	2.16**
M-	.65	.89	-.24
O+	1.65	.17	1.48**
Repeat	.41	1.89	-1.48**
However	1.65	.44	1.21**

**p<.01

Table 16. Percentage of students who ever used each type of idea unit in post-topic essays.

	Topic 1		Topic 2		Topic 3		Topic 4	
	Treatment Group (n=19)	Comparison Group (n=16)	Treatment Group (n=19)	Comparison Group (n=21)	Treatment Group (n=11)	Comparison Group (n=13)	Treatment Group (n=17)	Comparison Group (n=18)
M+	100%	93.8%	100%	100%	100%	92.3%	100%	88.9%
O-	36.8%	37.5%	84.2%	19.0%	72.7%	46.2%	94.1%	50.0%
M-	0%	0%	26.3%	33.3%	9.1%	0%	35.3%	50.0%
O+	36.8%	25.0%	21.1%	4.8%	0%	23.1%	82.4%	11.1%
Repeated	52.6%	100%	26.3%	81.0%	18.2%	53.8%	35.3%	61.1%
However	36.8%	18.8%	31.6%	14.3%	0%	7.7%	76.5%	33.3%

However, these regressions do not control for the length of essay. To get a fuller picture of how the groups compared as the year progressed, an examination of the proportions of each type of idea units to the total number of idea units was necessary. Table 17 shows the proportions of each type of idea unit by topic and group. A follow-up binomial regression model examined the proportions of M+, O-, M-, and O+ idea units for each topic. The coefficients for all comparisons across all topics are summarized in Table 18, and the exponentiated coefficients for all comparisons across all topics are summarized in Table 19.

Table 17. Proportions of types of idea units in post-topic essays by topic and group.

	Topic 1		Topic 2		Topic 3		Topic 4	
	Treatment Group (n=19)	Comparison Group (n=16)	Treatment Group (n=19)	Comparison Group (n=21)	Treatment Group (n=11)	Comparison Group (n=13)	Treatment Group (n=17)	Comparison Group (n=18)
M+	.779	.776	.509	.814	.486	.692	.522	.625
O-	.127	.138	.361	.071	.486	.231	.269	.225
M-	.000	.000	.056	.100	.027	.000	.059	.100
O+	.095	.086	.074	.014	.000	.077	.151	.050

The trends seen in the proportional difference between the groups over time offered a slightly different view of the trends than that seen in the Poisson analysis. As table 18 shows, significant differences between the groups emerge over the course of the year of the year. Of most interest is topic 4, where the proportion of O+ to M+ in the collaborative group was

significantly higher than that proportion in the individual group by the end of the year ($F(1, 171)=5.138, p<.05$, at topic 4). However, the proportion of M- to any other type of argument was lower in the collaborative group at topic 4 than in the comparison group, a result that must be interpreted cautiously due to the relatively low use of M- by either group in any topic.

Table 18. Coefficients for binomial mixed models analysis across topic, with treatment group set as reference group.

	Topic 1			Topic 2			Topic 3			Topic 4		
	O-	O+	M-	O-	O+	M-	O-	O+	M-	O-	O+	M-
Reference Category: M+	.081	-.082	-	2.035***	-1.861	-.108	-1.051	17.375	-.805	.492	1.453*	-1.202*
Reference Category: O-	-	-.182	-	-	.261	-2.275**	-	2.425	-.805	-	.962	-1.912*
Reference Category: O+	-	-	-	-	-	-2.208	-	-	-45.132**	-	-	2.971**

* $p<.05$

** $p<.01$

*** $p<.001$

Table 19. Exponentiated coefficients for binomial mixed models analysis across topic, with treatment group set as reference group.

	Topic 1			Topic 2			Topic 3			Topic 4		
	O-	O+	M-	O-	O+	M-	O-	O+	M-	O-	O+	M-
Reference Category: M+	1.085	0.921	-	7.65	0.155	0.898	0.350	∞^{\dagger}	0.447	1.636	4.275	0.301
Reference Category: O-	-	0.834	-	-	1.298	0.103	-	11.301	0.447	-	2.617	0.148
Reference Category: O+	-	-	-	-	-	0.110	-	-	$-\infty^{\dagger}$	-	-	0.051

[†]Exponentiated coefficients are not estimable by maximum likelihood due to quasi-complete separation.

Table 19 shows the exponentiated coefficients for the binomial regression. These coefficients represent the odds of the treatment group using a particular type of idea unit proportionally more than the reference category. For example, in topic 4, students in the

collaborative condition were 4.275 times more likely to use a larger proportion of O+ to M+ ideas in their essays than students in the individual condition.

A note about extreme values in the tables above is warranted, especially those in topic 3. Two proportions in topic 3 have both large coefficients and very large or very small exponentiated coefficients (O+ to M+ and M- to O+). Both of these are a result of a statistical issue arising from the methodological issue with topic 3. Specifically, because the post-topic 3 essay was done as homework, the number of essays returned and the number of idea units in both groups was very small. As a result, the raw proportions of ideas included 0.00, which resulted in quasi-complete separation so the coefficients were not estimable to maximum likelihood.

Overall, the binomial regression gives a similar picture of the development of argumentative writing that the Poisson models do, though perhaps a less dramatic picture. Specifically, the binomial regression shows that the treatment group moved more towards using other-side positive reasons in their essays as the year progressed.

To compare across time, a mixed models Poisson regression was performed on the final, post-topic individual essay for topics 1 and 4. The within group factor was topic, and between group factors were number of unique idea units, and number of M+, O-, M-, and O+ idea units, repeated idea units, and *however* clauses. For this analysis, students had to have been present for the mid-topic essay date and the final topic essay date for both topics 1 and 4. The n for the treatment group was 14, and the n for the comparison group was 13. Table 20 summarizes the results of the analysis.

In mean number of idea units, and in all types of idea units except for M-, there were significant differences in growth from topic 1 to topic 4. In other words, as students were in the curriculum longer, the differences between the groups increased. The difference between groups

was greater for number of unique idea units at topic 4 (6.50) than at topic 1 (1.38), $F(3, 56)=3.97, p<.05$. In addition, the difference between groups was greater for number of M+ idea units at topic 4 (3.10) than topic 1 (1.08) ($F(3, 56)=3.77, p<.05$); greater for number of O- idea units at topic 4 (2.16) than topic 1 (.13) ($F(3, 56)=5.09, p<.01$); and greater for number of O+ idea units at topic 4 (1.48) than topic 1 (.16) ($F(3, 56)=3.70, p<.05$). Finally, the difference between groups approached significance for *however* clauses at topic 4 (1.21) vs. topic 1 (.33) ($F(3, 56)=2.36, p<.10$).

Table 20. Differences in mean idea units between groups in topics 1 and 4.

Type	Topic 1	Topic 4
Total Idea Units	1.38	6.50*
M+	1.08	3.10*
O-	.13	2.16**
M-	-.24	.65
O+	.16	1.48*
Repeat	-.43	-1.48
However	.33	1.21

* $p<.05$

** $p<.01$

To examine pre-test and post-test, comparisons of essays were undertaken of pre-topic essay for topic 1 and the post-topic essay for topic 4. A mixed effects Poisson regression with the between-groups factor of type of idea unit and the within-groups factor of time (pre-test, post-test) was performed. The n for the treatment group was 27, and the n for the comparison group was 28.

The mean change of each group from pre-test to post-test is in Table 21. Unsurprisingly, based on what we've already seen of the growth trajectories of the two groups, the treatment group showed greater growth in several key areas. The mean number of idea units grew in the treatment group by 9.86, while only growing by 2.69 in the comparison group. The Poisson model found a highly significant interaction term of topic and group ($F(1, 50)=21.25, p<.001$), indicating that the difference between the groups when it comes to growth from pre-test to post-test was highly significant. Considering individual types of idea units, growth for M+ ideas was greater for the treatment group ($F(1, 50)=12.66, p<.001$). In addition, for O- clauses, the treatment group grew significantly from pre- to post-tests ($F(1, 50)=10.72, p<.01$). The interaction of time and class was likewise highly significant for O+ ideas ($F(1, 50)=13.96, p<.001$). In addition, the difference in growth between groups when it came to *however* clauses was significant ($F(1, 50)=4.43, p<.05$), indicating that the fact that the collaborative group increased usage from pre-test to post-test by 1.64 was significantly higher than the .54 increase in *however* clauses for the individual group.

Meanwhile, the comparison group increased the mean number of repeated idea units in their essays from 2.00, while the treatment group only increased repeated idea units by .43. The Poisson model indicates that the interaction between class and time in this respect was also significant ($F(1, 50)=8.89, p<.01$). This means that the comparison group was more likely to simply repeat ideas in their essays, whereas the treatment group was less likely to do so by the end of the year.

Table 21. Differences at topic 1 pre-essay and topic 4 post-essay, by group.

Type of Idea Unit	Treatment Group (n=27)	Comparison Group (n=28)	Difference in Means (Treatment – Comparison)
-------------------	------------------------	-------------------------	--

# of Idea Units	9.86	2.69	7.17***
M+	4.71	1.54	3.17***
O-	2.79	.23	2.56**
M-	.64	.92	-0.28
O+	1.71	0.00	1.71***
Repeat	.43	2.00	-1.57**
However	1.64	.54	1.1*

*p<.05

**p<.01

***p<.001

Table 22. Coefficients for binomial mixed models analysis on pre-test and post-test, with treatment group and pre-test set as reference.

	O-	O+	M-
Reference Category: M+	1.42	2.14	-1.12
Reference Category: O-	-	.96	-1.92
Reference Category: O+	-	-	-3.16

Table 23. Exponentiated coefficients for binomial mixed models analysis on pre-test and post-test, with treatment group and pre-test set as reference.

	O-	O+	M-
Reference Category: M+	4.12	8.52	.327
Reference Category: O-	-	2.62	.146
Reference Category: O+	-	-	.043

A follow-up mixed models binomial regression was performed on the pre-test and post-test, with the collaborative group pre-test set as a reference. The coefficients are in table 22 and the exponentiated coefficients are in table 23. None of the coefficients were significant, although

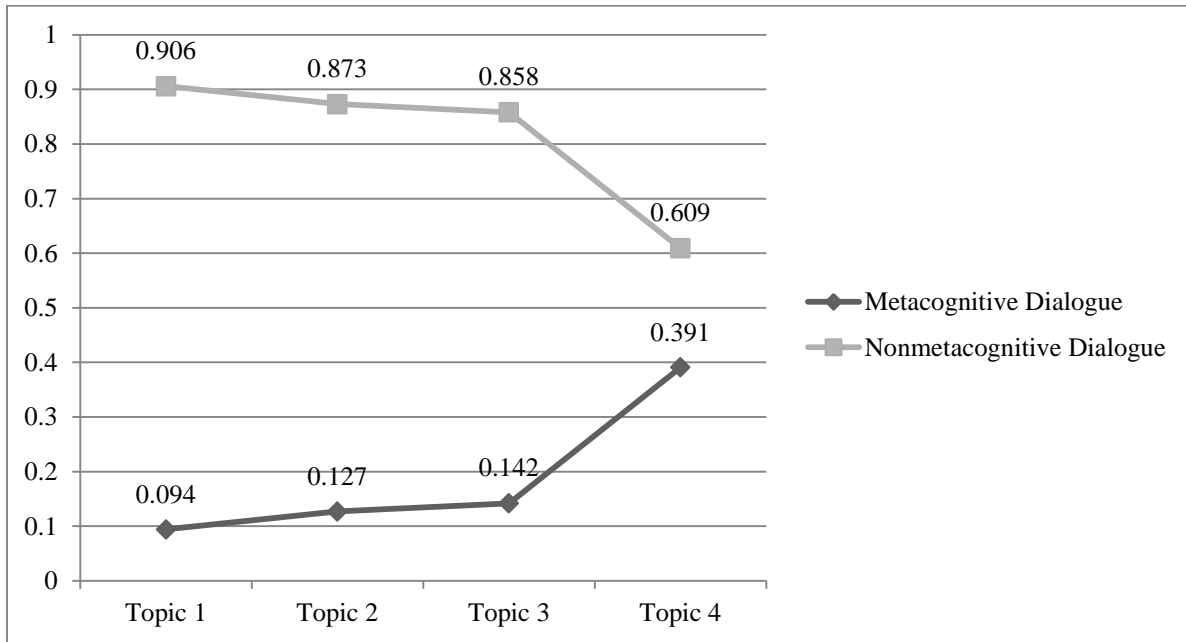
the binary of M+ and O- approached significance ($F(1, 253)=3.39, p<.10$). Likewise, the comparison of M+ and O+ across time and group approached significance ($F(1, 208)=2.71, p=.10$).

Hypothesis 2. The benefit of collaboration will be reflected in increased metacognitive discussion during the collaborative writing activity.

A comparison of different types of collaborative discourse was made, based on the recordings from topic 1 and topic 4. In topic 1, there were 11 dyads, whereas in topic 4, there were only 6 due to malfunction in some of the audio recorders. The comparison shows that in both the first and last topics, both metacognitive and non-metacognitive discourse was used. Figure 7 shows the proportion of discourse for each topic that was made up of metacognitive and non-metacognitive talk. As the year progressed, the collaborations included more metacognitive and less non-metacognitive discussion, with the largest change occurring between topics 3 and 4.

The percentage of pairs who used each type of discourse across topics appears in table 24. From the beginning of the year, all dyads used both metacognitive and non-metacognitive dialogue. From topic to topic there were variations in the types of metacognitive discourse used by pairs. For example, 9 of 11 pairs (82%) used the metacognitive strategy of planning in topic 1, while only 4 of 6 (67%) used it in topic 4. However, all groups used some types of metacognitive dialogue (planning, evaluating, or regulating) in topics 1 and 4. Further examination of the use of metacognitive dialogue over time indicates a general upward trend. Table 25 shows the mean frequency of occurrence of the different types of discourse in topics 1 and 4.

Figure 7. Proportions of metacognitive and non-metacognitive discourse in collaborative recordings for each topic.



Use of metatalk increased from a mean of 6.27 metacognitive utterances per dialogue in topic 1 to a mean of 14.33 in topic 4, while the total number of utterances stayed approximately the same (36.08 in topic 1 and 36.66 in topic 4). In contrast, non-metacognitive utterances per dialogue decreased from topic 1 (29.81) to topic 4 (22.33). It is also important to note that each individual type of metatalk saw increases in their means across topic. Planning discussion increased from 3.18 in topic 1 to 7.33 in topic 4, evaluating discussion increased from 1.09 in topic 1 to 4.33 in topic 4, and regulating increased from 2.00 in topic 1 to 2.67 in topic 4.

To test whether trends in the use of different categories of metacognitive and non-metacognitive talk were significant, chi-square tests were performed on each sub-category (e.g., planning, evaluating). These tests showed significance in the non-metacognitive category of surface level division of labor ($\chi^2 (27, 32)=42.402, p<.05$), indicating that the proportion of utterances that involved surface level division of labor (who would write, for example) decreased

significantly over the course of the four topics. However, no other individual sub-category of discourse showed significance, indicating that no single type of discussion changed significantly over the course of the year.

Table 24. Percentage of groups who used each type of discussion by topic.

Type of Discussion	Description of Discussion	Topic 1 (n=11)*	Topic 4 (n=6)*
Metacognitive	Planning	82%	67%
	Evaluating	55%	50%
	Regulating	82%	83%
	<i>Any Metacognitive</i>	<i>100%</i>	<i>100%</i>
Non-Metacognitive	<i>Any Non-Metacognitive</i>	<i>100%</i>	<i>100%</i>

*n represents the number of dyads

Table 25. Mean frequency of occurrence of discussion types in topics 1 and 4.

Type of Discussion	Description of Discussion	Topic 1 (n=11)*	Topic 4 (n=6)*
Metacognitive	Planning	3.18	7.33
	Evaluating	1.09	4.33
	Regulating	2.00	2.67
	<i>Total Metacognitive</i>	<i>6.27</i>	<i>14.33</i>
Non-Metacognitive	<i>Total Non-Metacognitive</i>	<i>29.81</i>	<i>22.33</i>
Total Utterances		36.08	36.66

*n represents the number of dyads

A follow-up one-way ANOVA was performed to examine the effect of time on metatalk vs non-metatalk. The result showed some interesting trends. Comparing topic 1 to topic 4, the mean increase in metacognitive dialogue instances of 8.06 approached significance ($F(3, 28)=2.22, p=.10$), as did the mean decrease in non-metacognitive dialogue of 7.49 ($F(3, 28)=2.55, p<.10$).

Hypothesis 3a. In their final, individual essays, students will use their own ideas and incorporate their partner’s ideas that support their own side more than they will use partner ideas that support the partner’s side.

Hypothesis 3b. Students will use more partner ideas in their independent essays later (in topic 4) than earlier (topic 1).

Tracing idea units is difficult to do, as it is impossible to know exactly where the ideas students use come from. However, comparing student recordings with student final essays provides a starting point. Final topic student essays were coded based on whether each idea unit was brought up by the student (“my idea”), the student’s partner (“partner’s idea”), or not brought up at all (“new idea”) during the interim essay collaboration. Thus, a “My M+” idea is one that supports the writer’s side and was brought up by the writer during the interim collaboration. Similarly, a “Partner O-” opposes the other side and appeared in the student’s final individual essay, but was brought up by the student’s partner during interim collaboration.

Table 26. Mean number of each type of idea unit by source and topic.

Type	Source	Topic 1	Topic 4
M+	Student	1.11	2.70
	Partner	.21	.50
	New	2.68	1.50
O-	Student	.16	1.10
	Partner	.05	.30
	New	1.74	1.70
M-	Student	.32	.30
	Partner	.05	.10
	New	.42	.50
O+	Student	.05	.10
	Partner	.05	.80
	New	.11	.70

Table 27. Percentage of students who ever used each type of idea unit by source and topic.

Type	Source	Topic 1 (n=19)	Topic 4 (n=10)
M+	Student	53%	90%
	Partner	21%	40%
	New	89%	80%
O-	Student	11%	60%
	Partner	5%	20%
	New	68%	70%
M-	Student	26%	30%
	Partner	5%	10%
	New	32%	20%
O+	Student	5%	10%
	Partner	5%	60%
	New	5%	40%

Qualitative comparisons of the different sources of four major types of idea units (M+, O-, M-, and O+) show that the percentage of students likely to use partner ideas is higher if the ideas support the other side (O+) than if they support the student's side (M+) (see table 27). To be included in this comparison, students had to be present for both the interim and final, individual essays. The n for topic 1 was 19, while the n for topic 4 was 10.

Unsurprisingly, students were much more likely to use ideas that supported their own side (M+) or criticized the other side (O-) than ideas that criticized their own side (M-) or supported the other side (O+). But the source of those ideas seems to matter. For example, in topic 4, 9 of 10 students used M+ ideas that they had explicitly offered up during collaboration, while only 4 of the 10 used M+ ideas of their collaborative partner. In contrast, only 1 of 10 students used an idea that supported the other side (O+) that they came up with during collaboration, while 6 of the 10 incorporated an O+ idea offered up by their partner during collaboration.

The reason for this may be simple numbers: an opposite-side partner during collaboration is less likely to come up with an M+ reason (which, for them, is an O+ reason). As a result, there were likely fewer partner M+ statements (i.e., an idea that supports my side that is from my opposite-side partner) during collaboration than partner O+ statements (i.e., an idea that supports the opposite side that is from my opposite-side partner). Thus, students are exposed during collaboration with an opposite-side partner to more O+ and M- ideas than they would be if they were working on their own (or with a same-side partner).

CHAPTER 4: DISCUSSION

The present study asked whether collaborative writing offers a further opportunity for growth of argumentative writing above and beyond that provided by practice in independent writing. Results show that dialogic writing activities enhanced students' performance in argumentative writing. In particular, students who engaged in collaborative writing were more likely to consider alternative claims and arguments. The process of collaboration aided in this development, as across the span of the school year, students became more likely to use metacognitive dialogue in their collaborations and to use their partners' ideas in their subsequent individual essays. The following section further addresses the findings and their implications, followed by areas for future research.

Development of the Argumentative Mindset

Growth progression in essays

Working with others is a complex process, and the impact it has on writing is likewise multifaceted. In the present study, students in the collaborative condition showed more improvement than those in the individual condition, differences that emerged throughout the year. By the end of the year, students in the collaborative group produced more unique idea units and used more O-, O+, and *however* clauses, a finding supported by both count and proportional analyses. This indicates that the students who collaborated with others understood and anticipated the arguments of the opposite side better than those who worked on their own.

All this suggests the development of an argumentative mindset. Through collaboration, students are learning how to think strategically as they write. Of particular interest is the growing divide between groups for the use of O+ and *however* clauses. Since these indicate that students are aware of and thinking about how the other side might argue, this suggests that working with a

student from the opposite side allowed students in the treatment group to develop an argumentative voice inside their head. In Vygotskian terms, the opposite-side collaboration is allowing the interpsychological discourse to become intrapsychological discourse (Vygotsky, 1978). This in turn allows them to address the other side's arguments, as opposed to simply presenting their own.

Of course, students in both conditions were exposed to arguments from the other side in other parts of the curriculum. The further growth in argumentation in the treatment group seen in the current study may be attributable to the exposure to the other side's positions during a writing activity. That is, the near transfer situation of writing collaboratively heightened the inner interlocutor in the treatment group.

Integrative arguments in essays

Qualitative examinations of student essays and collaborative recordings bolster the idea of the development of an argumentative mindset. Often, students address specific issues in their individual essays that were brought up during the collaborative writing session. One example of this emerged in topic four, where several students addressed the argument that making kidney sales legal would result in people being forced to sell their kidneys. Some students addressed the issue with little context (e.g., simply writing a sentence like, "No one should be forced to sell a kidney"), and some put it into a however clause (e.g., "Some people say that they'll be forced to sell a kidney, but that's not the case"). Either way, students seemed to directly address an issue that came up during collaboration, indicating a move from interpsychological to intrapsychological.

It is perhaps not surprising that students collaborating with others from the opposite side are able to think more about what the other side might say. They are actively engaging with opposite-side arguments while in the process of writing, which allows for near transfer when

moving to individual writing. In fact, during the collaborative writing process, many dyads became so wrapped up in debate that they were able to get little down on paper, a point discussed in more detail below. But the fact of the matter is, the process of working with another student from the opposite side allows students to engage in a way that encourages them to think flexibly about arguments.

An example of the development of this argumentative mindset can be found in the use of integrative arguments in one student's final essay of the year. Student 24 was a student in the collaborative condition present for all sessions across the year. His first essay of the year was relatively unimpressive, consisting of two idea units, an M+ and O-. This indicated some dual representation, but no integrative argumentation. By the final essay of topic 4, however, student 24 had six *however* phrases in his essay, an indication that he had developed an inner interlocutor that allowed him to utilize integrative argumentation.

Beyond the sheer number of integrative arguments in student 24's final essay, the type of integrative arguments used by him is of interest. Two of the six *however* phrases were reused ideas from the collaborative writing session. They each consisted of a partner O+ linked to a "my" O- idea, indicating that he was simply remembering and recording the back-and-forth between his partner and himself during collaboration. But one of his *however* phrases consisted of a partner O+ countered with an O- that was not used during collaboration. Further, half of his *however* phrases consisted of O- and O+ phrases not used during the collaboration. The use of new ideas not seen during the collaborative writing suggests that more than a simple remembering of the collaborative debate was going on; he seems to have developed the ability to anticipate arguments from the other side and counter them, more evidence for the development

of an argumentative mindset consisting of an inner interlocutor that allows for integrative argumentation.

Yours, Mine, Ours

Using ideas from collaboration

As mentioned above, as students in the collaborative condition progressed through the year, they began to address in their individual essays issues that arose during their collaborative writing session. It was as though they took the collaborative conversations into their subsequent independent writing session. The ways in which they utilized these dialogues to help them compose individually became more complex as they moved from topic to topic.

It is virtually impossible to perfectly trace the origin of ideas seen in a student's essay, especially in a dialogue-rich environment such as the one produced by the argumentation curriculum in which the present study was done. Students may have heard arguments from many different sources, including in their same-side small groups, in the dyadic electronic dialogues, and of course in the interim collaborative writing session. Thus, there is no way to know for sure where a student first heard an idea or what triggered them to use it in their final, individual essay.

However, keeping in mind the differences in the final essays between the conditions in the current study and comparing the final independent essays to the audio recordings of the collaboration offers a starting point for understanding the way that student collaborators might be influenced by their dialogic writing session. As discussed above, students who engaged in dialogic writing mid-topic increased the number of O+ ideas as the year progressed. Examining the source of this increase in O+ ideas offers a glimpse into what might have been going on. In topic 1, only 5% of students in the treatment group carried over O+ ideas introduced by their interlocutor from the interim collaborative essay to the final individual essay. By topic 4, 60% of students were using their partner's O+ ideas in their subsequent individual essay. This suggests

that, as time went on, they were becoming better at incorporating the opposite-side position of their collaborators into their individual efforts, a trend that supports the idea that an argumentative mindset was being built through the collaboration.

Likewise, the percentage of students who used their own M+ ideas from collaboration increased dramatically, from 53% in topic 1 to 90% in topic 4. This could indicate that students were internalizing all parts of the collaborative writing session, not simply what they heard from their partners. Interestingly, though, the percentage of students who used M+ ideas presented by a partner almost doubled, from 21% in topic 1 to 40% in topic 4. While this is still a relatively small number, it must be noted that a partner M+ is the same as an O+ to the partner who introduced it during collaboration. These types of utterances were much less likely to occur than partner O+ utterances (which would be an M+ to the partner who presented it), and therefore it is telling that more and more, students were using them. This perhaps hints that they were being used more in collaboration as the year progressed. That is, students may have been more likely to present an idea that supported the opposite side during the dialogic process as time went on, leading to their partner subsequently using that idea in a later independent essay.

All of this suggests that it is the difficult but impactful work of being paired with an opposite side partner that could result in the most argumentative gains for students. As they worked through the messy process of trying to write an essay with someone with an opposing stance, they began to develop an understanding of the complexity of an issue, particularly as it relates to differing sides of a debate. Thus, they become more likely to produce opposite-side supportive statements during collaboration and are more likely to use opposite-side supportive statements in their later essays.

The Process of Collaboration

By looking at both the development of writing through the course of the year and the ways in which students use their opposite-side partners' ideas in their individual essays, we can see that the collaboration is having an impact on their individual writing. But what is going on during the collaborative writing process that is causing these changes? To answer that question, we must look at the paired writing practice, where an examination of the audio recordings gives a hint as to how collaboration might impact argumentative writing development.

Discourse vs. writing

As mentioned above, students often became so involved in the verbal debate with their interlocutors that the collaborations yielded little in the way of actual writing. On recordings, teachers circulating can often be heard encouraging dyads to write. The assumption of many teachers is that nothing on paper is indicative of non-work. In fact, immediately following the first interim essay of the year, a quick perusal of the essays led to the hasty conclusion that the comparison group did more work because their essays were longer and more robust than the collaborative essays.

However, the actual recordings tell a different story. Often, dyads were actively involved in debating and internalizing the issues surrounding the topic. In fact, much of the paired writing time was spent deciding which side (pro or con) to write from. Students debated each other passionately not only on the merits of the arguments for each side, but on which side would be the best strategic choice for writing. Time spent debating and working through the process of writing, though it might not show on paper, is not time wasted. On the contrary; it is this grappling of arguments during the process of planning an essay that may lead to students internalizing the other side's arguments later, during their individual essay writing.

Teacher direction to write things down on the recordings often interrupted this intra-collaborative debate, putting an end (temporary or permanent) to discussion in favor of moving towards writing things down. Recordings indicate that students themselves understood the importance of these arguments. In the second topic, for example, one student became frustrated with her partner for not debating with her, pushing her to challenge ideas and present her own so that they could get the best ideas down on paper. “You’re supposed to argue with me,” she says at one point, and when her partner asks why, she replies, “because that’s how we figure out what to write.” Even that early in the year, students seemed to understand the importance of the process of collaborative writing.

The teachers (both the author of this study and another researcher familiar with the curriculum) differed from the students in their focus, though, which is not surprising. Our focus, especially early in the year, was clearly on product, not process. Our interruptions of the dyads’ verbal sparring often derailed the debate between the students. In contrast, a process-centered approach to this type of exercise, including allowing students to spend as long as they need debating instead of writing, might have strengthened the development of the argumentative mindset in students.

Metacognitive collaborations

When left to debate, the type of dialogue that dyads engaged in during collaborative writing evolved as the year progressed. Near the beginning of the school year, student pairs were focused more on non-metatalk than near the end of the year, when more groups were engaging in metacognitive dialogue. This suggests that student pairs were beginning to emerge as metacognitive collaborators. This can be seen in the recordings, which show more complex collaborative dialogue as time went on. For example, as the year progressed, dyads began to engage with other dyads during collaboration, leading to cross-collaborative discourse.

An early example of this cross-collaborative discussion occurred during topic 2, which focused on the ethics of animal testing. As two students debated the difference in experimenting on animals versus humans, a nearby student chimed in, bolstering the argument of one of the original interlocutors. A portion of this dialogue can be seen in Table 28.

Though the actual argument in this example (that animals are human) is not empirically correct, the fact of the cross-collaboration demonstrates the ways in which students engaged with each other and the material to enrich their collaborative environment. They are actively shaping their learning environment to meet their needs, and the result is a more complex collaborative arrangement. This, in turn, is likely to make them more sophisticated collaborators.

Table 28. Discourse between dyads during topic 2 (animal testing).

Pair 1, Student 1	Pair 1, Student 2	Pair 2, Student 3
How is a animal - [gets up from table] [returns] You should- How is a animal a human?!		
	I said a human!	
		A animal is a human.
How?		
		Apes. Apes. Evolution.
A animal walks on four legs not two.		
	But guess what? Guess what? A animal, guess what – is living right? So -	
But still its not a human!		

There is some evidence that the collaborations themselves became more sophisticated as the year went on, as evidenced by an increase in the metacognitive dialogue used during collaboration. In the beginning of the year, dyads were almost five times more likely to utter non-metacognitive dialogue during collaboration than metacognitive dialogue. By the final topic of the year, that gap had narrowed, and almost 40% of utterances during collaborative were

metacognitive in nature. And while individual categories of metacognitive and non-metacognitive dialogue were mostly non-significant, movement towards the general category of metacognitive collaboration from topic one to topic four indicates that students were deepening their conversations.

Thus, when examining the underlying reasons for the argumentative writing gains seen in the treatment group, we can see the ways in which what happened during collaboration impacted individual writing later. Prolonged exposure to an opposite-side partner in a near transfer situation (i.e., debate and discussion during the process of writing, as opposed to debate and discussion during other parts of the curriculum) clearly had an impact. Part of this exposure included metacognitive discourse, which is clearly an indication of deepening thinking about argumentation. However, this metatalk also likely fed the deepening understanding of argumentation, both indicating and contributing to the development of an argumentative mindset in the interlocutors as the year developed.

Limitations & Future Directions

This study demonstrated the impact collaborative writing activities can have on the development of argumentative writing skills, but it does have some limitations. As discussed earlier, the two groups were composed of different types of students. While the treatment group was a general education classroom, the comparison group was a class of English language learners (ELLs). Analysis of the pre-instructional essay for both groups found that they were statistically equivalent, a conclusion supported by observations, including the fact that many students in the treatment group were ELLs themselves, and many students in the comparison group had high English proficiency. As a result, equivalence between the groups was assumed.

Still, the possibility that the comparison group's emerging English could have impacted their development in argumentative writing cannot be ruled out.

It is also important to remember that the interim and final topic essays were done in class (with the exception of topic 3, discussed below). Language barriers could have made the time limitations on writing more difficult for the comparison group. It is possible, then, that the differences between groups would have narrowed had they been given unlimited time to work on their essays.

Another limitation with the study is the number of classes used. As a quasi-experiment, each class is a single unit, and any issue that impacted one class but not the other could confound the results. While there were no observed differences between the classes in terms of teaching styles, events that impacted one group but not the other, or other major issue, there is no way to completely rule out the fact that something could have influenced the growth of one class and not the other.

A further limitation of this study involved the students used in the analysis. This study reflects the real-world issue of absenteeism. Students came and went from the classroom, transferring in and out as the year progressed and missing days due to absence or special events that pulled one or two students from the class. As a result, the sample varied across topics. The final analysis included only students that were present for the essay days for at least three of the four topics, the same students were not present at every measured point. While this is a serious limitation on the conclusions that can be drawn based on this data, it is also a problem that many teachers face daily.

One possible limitation might also lie in cultural assumptions around argumentation. In many cultures, arguing (even in the context of a class debate) is seen as aggressive and negative.

Further, even in cultures that value argumentation, such as that of the mainstream United States, often don't value argumentation equally for both genders. Thus, some students might have felt internal cultural and/or gendered pressure not to argue as passionately or to be as vocal.

Observationally, the opposite seemed to hold true for many students. As the students in this study were predominately students of color, in a poor neighborhood, and attending a school with low test scores, much of their schooling is in a highly regimented environment. Speaking their minds is often treated as insubordination. Anecdotally, many of the students that struggled with behavior issues in other classes turned into the best-behaved students in the program, as they found a space (often the only space) to express their views and have a voice. Even students from cultures where they might be expected to avoid speaking out often became the more vocal and enthusiastic debaters in class. Despite these observations, though, there still might have been students who chose to remain quiet due to cultural and gendered expectations.

One final limitation of the current study lies in the lack of a transfer task. That is, students were not given an essay or other argumentation task outside of the topics covered in the curriculum. This means that there isn't a way to know for sure whether the benefits seen in the study transfer outside of the specific topics the students engaged with deeply.

This study represents a starting point for thinking about the near transfer situation of working with an interlocutor during argumentative writing before subsequently writing independently. The results suggest that collaborative writing can aid in argumentative development, but future research may be able to better clarify under what conditions dialogic writing enables growth. One future step would be to examine the impact of working with a partner from the same side in comparison to working with a partner from the opposite side, as the treatment group in the present study did.

Another topic for exploration involves ways to mitigate the teachers' natural instinct to focus on the product of collaboration, while still keeping the near transfer situation. The recordings indicate that having the space to debate during the process of collaboration, without pressure to get words down on paper, might contribute to argumentative writing development. However, taking away the writing part of the collaboration could diminish the benefits of near transfer seen in moving from writing with an interlocutor to writing alone. Future research, then, could look at the close balance for teachers keeping the writing assignment, while stressing the process of collaboration instead of the finished writing piece.

Conclusion

Collaboration is both a 21st century skill to be learned and a tool through which other skills, such as argumentative writing and metacognition, can be developed. Our current educational standards place a high value on both collaborative and argumentative skills; however, opportunities for students to grow these skills with others are arguably few. Further, the educational climate is such that teachers often feel pressured to think about product instead of process. Allowing students the space to engage deeply with each other can help them build social, cognitive, and metacognitive skills that can serve them well in many areas. This study contributes to a growing body of literature that suggests collaborative writing may be that space.

REFERENCES

- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First Century Skills. In *Assessment and Teaching of 21st Century Skill* (Vol. 1, pp. 17–66). <https://doi.org/10.1007/978-94-007-2324-5>
- Chinn, C. A., & Clark, D. B. (2013). Learning through collaborative argumentation. *International Handbook of Collaborative Learning*, (February), 314–332. <https://doi.org/10.4324/9780203837290.ch18>
- Crowell, A. (2011). *Assessment of a three-year argument skill development curriculum*. Retrieved from <http://academiccommons.columbia.edu/catalog/ac:132281>
- Dede, C. (2009). Comparing Frameworks for “21st Century Skills.” In *21st Century Skills: Rethinking How Students Learn* (pp. 51–76).
- Felton, M., & Kuhn, D. (2001). The Development of Argumentive Discourse Skill. *Discourse Processes*, 32(2), 135–153. https://doi.org/10.1207/S15326950DP3202&3_03
- Frijters, S., ten Dam, G., & Rijlaarsdam, G. (2008). Effects of dialogic learning on value-loaded critical thinking. *Learning and Instruction*, 18(1), 66–82. <https://doi.org/10.1016/j.learninstruc.2006.11.001>
- Gélat, M. (2003). Taking Others’ Perspectives in a Peer Interactional Setting while Preparing for a Written Argument. *Language and Education*. <https://doi.org/10.1080/09500780308666855>
- Graff, G. (2008). *Clueless in academe: How schooling obscures the life of the mind*. Yale University Press.
- Grau, V., & Whitebread, D. (2012). Self and social regulation of learning during collaborative activities in the classroom: The interplay of individual and group cognition. *Learning and Instruction*, 22(6), 401–412. <https://doi.org/10.1016/j.learninstruc.2012.03.003>
- Howe, C. (2010). Peer Dialogue and Cognitive Development: A two-way relationship? In K. Littleton & C. Howe (Eds.), *Educational Dialogues: Understanding and Promoting Productive Interaction* (pp. 32–47). New York: Routledge. Retrieved from https://books.google.com/books?hl=en&lr=&id=_buLAgAAQBAJ&pgis=1
- Howe, C. (2013). Scaffolding in context: Peer interaction and abstract learning. *Learning, Culture and Social Interaction*, 2(1), 3–10. <https://doi.org/10.1016/j.lcsi.2012.12.005>
- Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., ... Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Educational Technology Research and Development*, 63, 125–142. <https://doi.org/10.1007/s11423-014-9358-1>
- Jarvis, P., Newman, S., & Swiniarski, L. (2014). On “becoming social”: the importance of collaborative free play in childhood. *International Journal of Play*, 3(1), 1–16.

<https://doi.org/10.1080/21594937.2013.863440>

- Johnson, D. W., & Johnson, R. T. (1989). Social Skills for Successful Group Work. *Educational Leadership*, (December 1989/January 1990), 29–33. Retrieved from http://www.researchgate.net/profile/David_Johnson50/publication/234590538_Social_Skills_for_Successful_Group_Work/links/00b7d5261b54ebe322000000.pdf
- Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). Cooperative Learning Methods : A Meta-Analysis. *Methods*, 1(JANUARY 2000), 1–33. Retrieved from <http://www.tablelearning.com/uploads/File/EXHIBIT-B.pdf>
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. <https://doi.org/10.1080/07370000802212669>
- Kiili, C., Laurinen, L., Marttunen, M., & Leu, D. J. (2012). Working on Understanding During Collaborative Online Reading. *Journal of Literacy Research*, 44(4), 448–483. <https://doi.org/10.1177/1086296X12457166>
- Kuhn, D. (2015). Thinking Together and Alone. *Educational Researcher*, 44(1), 46–53. <https://doi.org/10.3102/0013189X15569530>
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science : A Journal of the American Psychological Society / APS*, 22(4), 545–552. <https://doi.org/10.1177/0956797611402512>
- Kuhn, D., Hemberger, L., & Khait, V. (2014). Argue with me: Argument as a path to developing students' thinking and writing. Retrieved from https://www.researchgate.net/publication/259176367_Argue_with_me_Argument_as_a_path_to_developing_students_thinking_and_writing
- Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the Development of Argumentive Writing in a Discourse-Rich Context. *Written Communication*, 33(1), 92–121. <https://doi.org/10.1177/0741088315617157>
- Kuhn, D., & Pease, M. (2006). Do Children and Adults Learn Differently? *Journal of Cognition and Development*, 7(3), 279–293. https://doi.org/10.1207/s15327647jcd0703_1
- Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing Norms of Argumentation: Metacognitive, Epistemological, and Social Dimensions of Developing Argumentive Competence. *Cognition and Instruction*, 31(4), 456–496. <https://doi.org/10.1080/07370008.2013.830618>
- Livingston, J. A. (2003). *Metacognition: An Overview*. Retrieved from <http://eric.ed.gov/?id=ED474273>
- Mercendetti, D. (2010). *Connecting Social Skills and Cooperative Learning*. State University of New York, College at Brockport. Retrieved from http://digitalcommons.brockport.edu/cgi/viewcontent.cgi?article=1073&context=eht_theses
- Mercer, N. (2013). The Social Brain, Language, and Goal-Directed Collective Thinking: A

- Social Conception of Cognition and Its Implications for Understanding How We Think, Teach, and Learn. *Educational Psychologist*, 48(3), 148–168.
<https://doi.org/10.1080/00461520.2013.804394>
- Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction*, 1(1), 12–21. <https://doi.org/10.1016/j.lcsi.2012.03.001>
- Näykki, P., Järvelä, S., Kirschner, P. A., & Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning-A process-oriented case study in a higher education context. *International Journal of Educational Research*, 68(February), 1–14.
<https://doi.org/10.1016/j.ijer.2014.07.001>
- Newell, G. E., Beach, R., Smith, J., & Vanderheide, J. (2011). Teaching and Learning Argumentative Reading and Writing : A Review of Research. *Reading Research Quarterly*, 46(3), 273–304.
- Panadero, E., & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist*, 20(3), 190–203. <https://doi.org/10.1027/1016-9040/a000226>
- Rogat, T. K., Linnenbrink-Garcia, L., & DiDonato, N. (2013). Motivation in Collaborative Groups. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *The International Handbook of Collaborative Learning* (1st ed., pp. 250–267). New York: Routledge.
- Schwartz, N. H. (2008). Exploiting the Use of Technology to Teach: The Value of Distributed Cognition. *Journal of Research on Technology in Education*, 40(3), 389–404. Retrieved from <http://search.proquest.com/docview/274711611?accountid=41453>
- Skoumios, M. (2009). The effect of sociocognitive conflict on students' dialogic argumentation about floating and sinking. *International Journal of Environmental and Science Education*, 4(4), 381–399.
- Slavin, R. E. (1990). Research on cooperative learning: Consensus and controversy. *Educational Leadership*, 47(4), 52–54. Retrieved from http://www.understandingbydesign.net/ASCD/pdf/journals/ed_lead/el_198912_slavin3.pdf
- Vass, E., & Littleton, K. (2010). Peer Collaboration and Learning in the Classroom. In K. Littleton, C. P. Wood, & J. K. Staarman (Eds.), *International Handbook of Psychology in Education* (p. 808). Emerald Group Publishing. Retrieved from <https://books.google.com/books?hl=en&lr=&id=tNepO17yQO4C&pgis=1>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes. Full-Text. (N.D.)*. <https://doi.org/10.1007/978-3-540-92784-6>
- Walton, D. (1998). *The new dialectic: Conversational contexts of argument*. Toronto: University of Toronto Press.
- Xiao, L., & Carrol, J. M. (2013). The Effects of Rationale Awareness on Individual Reflection Processes in Virtual Group Activities. *International Journal of E-Collaboration*, 9(2), 78–

95. <https://doi.org/10.4018/jec.2013040104>

Zillmer, N. (2016). *Metacognitive Dimensions of Adolescents' Intellectual Collaboration. Dissertation.* Columbia University. <https://doi.org/10.1017/CBO9781107415324.004>

LIST OF APPENDICES

Appendix A. Straw poll & pre-essay handout given to students on the first day of the first topic.....	69
Appendix B. Reflection sheets used during sessions 4-9 (electronic dialogues).....	70
Appendix C. Summary reflection sheets used in sessions 10-11 (small group Showdown prep).	72
Appendix D. Topic 1 essay worksheet.....	74
Appendix E. Oral Instructions for essays.....	75

Appendix A. Straw poll & pre-essay handout given to students on the first day of the first topic.

PRE-TOPIC OPINION POLL

Name: _____ **Date:** _____

Question: Should a tax be charged on soft drinks and that money go to lowering prices of healthy foods?

Yes No Undecided

How sure are you of your opinion? (Circle one)

Certain Very Sure Sure So-so Not very sure Not sure at all

Why do you believe the way that you do?

Appendix B. Reflection sheets used during sessions 4-9 (electronic dialogues).

"OTHER" REFLECTION SHEET

Team members _____

Date _____

Let's think...Starting with the other side's argument

One of the other side's
MAIN ARGUMENTS was:

Our **COUNTERARGUMENT**
against their argument was:

Give a specific example of an
improved, more effective
COUNTERARGUMENT.



"OWN" REFLECTION SHEET

Team members _____

Date _____

Let's think...Starting with our argument

One of our **MAIN ARGUMENTS** was:

Their **COUNTERARGUMENT** against our argument was:

Our **COMEBACK** was:

How can this **COMEBACK** be improved?
Is there a more effective comeback?



Appendix C. Summary reflection sheets used in sessions 10-11 (small group Showdown prep).

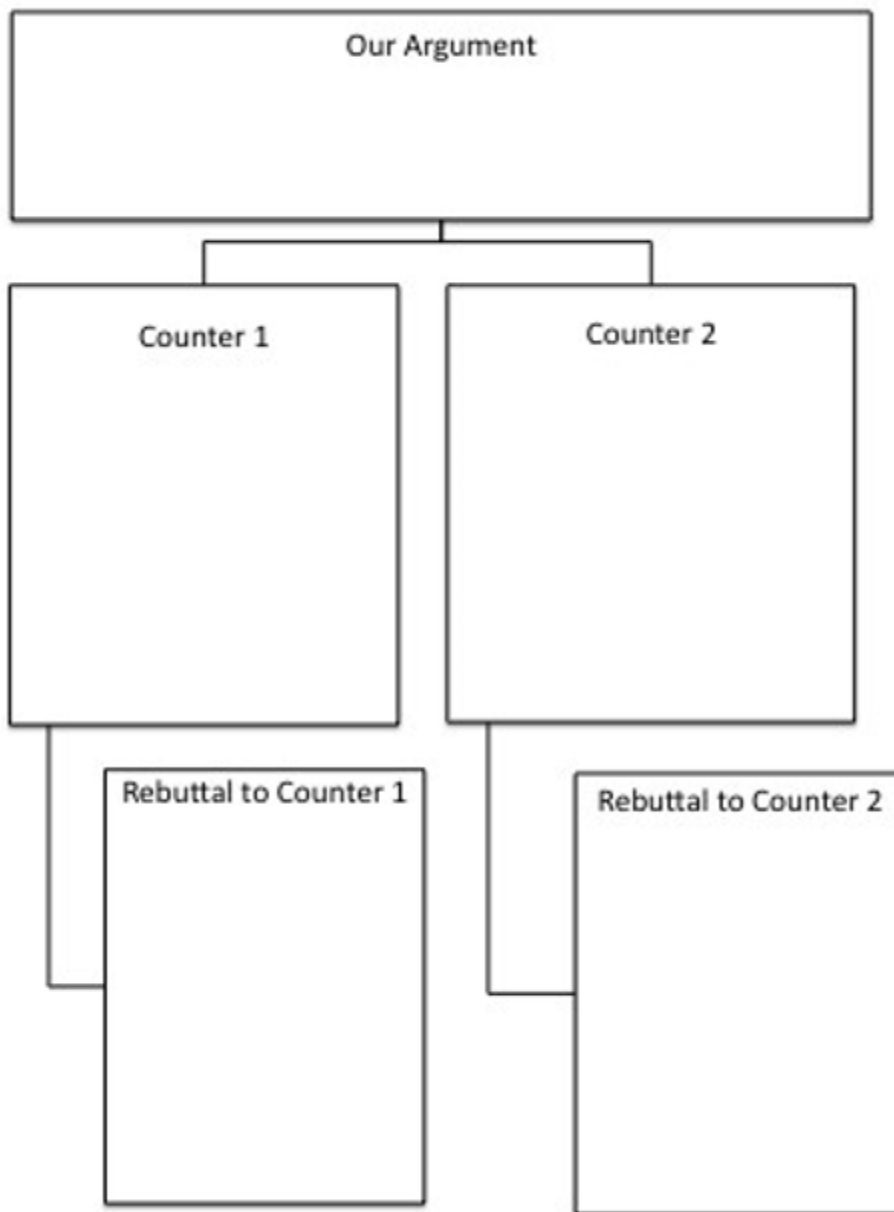
“OTHER” SUMMARY REFLECTION SHEET

THEIR ARGUMENT:

OUR BEST COUNTERARGUMENT:

ANOTHER COUNTERARGUMENT:

"OWN" SUMMARY REFLECTION SHEET



Appendix D. Topic 1 essay worksheet.

PERSUASIVE LETTER

Name and Class: _____ Date: _____

Question: **Should a tax be charged on soft drinks and that money go to lowering prices of healthy foods?**

Yes No

How sure are you of your opinion? (Circle one)

Certain Very Sure Sure So-so Not very sure Not sure at all

Write a letter to Mayor De Blasio to convince him that he should or should not charge a tax on soft drinks and put that money toward lowering the prices of healthy foods.

Appendix E. Oral Instructions for essays.

Pre-topic essay instructions for control and experimental groups:

“Today, we are going to write down some ideas that we have on the topic of whether or not taxes should be charged on sodas. I want you to follow along as I read the question at the top of your paper. ‘Should a tax be charged on soft drinks and that money go to lowering prices of healthy foods?’ Now, I want you to take a moment to circle what you think, yes, no, or undecided, which means that you’re not sure. [pause to let students write] Now, below that, you’ll see another question: ‘How sure are you of your opinion?’ I want you to circle the one that best represents how sure you are of your opinion: certain, very sure, so-so, not very sure, or not sure at all. [pause to let students write] Now, I’m going to give you the rest of the period to tell me why you think the way that you do. You may use the bottom of this sheet, and the back, and if you still need more room, raise your hand and I’ll bring you another sheet of paper.”

Interim topic essay instructions for control group:

“Today is our last day to get prepared for our Showdown. So what you’re going to do is to work on your own today to get your mind around some of the ideas you’ve been talking about in your groups and in your pairs. We’ll be writing our pre-showdown letters today. At the top of your paper, you’ll see a question that you should be familiar with by now: ‘Should a tax be charged on soft drinks and that money go to lowering prices of healthy foods?’ Please circle what you think, yes, no, or undecided. [pause to let students write] Below that please answer the question, ‘How sure are you of your opinion?’ Circle the one that best represents how sure you are of your opinion: certain, very sure, so-so, not very sure, or not sure at all. [pause to let students write] You’re going to get the rest of the period today to work on writing a letter to Mayor De Blasio to convince him that your position is the better one. You may use the front and back of this paper, and if you need additional space, raise your hand and we will bring you another sheet of paper.”

Interim topic essay instructions for experimental group:

“Today is our last day to get prepared for our Showdown. So what you’re going to do is to work with a partner from the other side today to get your mind around some of the ideas you’ve been talking about in your groups and in your pairs. We’ll be writing our pre-showdown letters today. At the top of your paper, you’ll see a question that you should be familiar with by now: ‘Should a tax be charged on soft drinks and that money go to lowering prices of healthy foods?’ And below that, you’ll see the question, ‘How sure are you of your opinion?’ You and your partner are going to discuss what side you want to write your essay from, and circle the side and how sure you are of your opinion. Then, you and your partner are going to write a letter to Mayor De Blasio to convince him that the position you choose is the better one. You may use the front and back of this paper, and if you need additional space, raise your hand and we will bring you another sheet of paper. You’re going to collaborate, which means that you have to agree on

everything that goes down on the paper. Every single thing, you should discuss and decide whether to put it down or not.

In addition to writing, while you collaborate, you will be using voice recorders as you talk. Once we come around and start your recordings, you should not turn the recordings off for any reason. Do not stop the recordings for the entire period. And, once again, remember that you and your partner are working together, and so when I listen to these later, I expect to hear both people talking and working.”

Post-topic essay instructions for control and experimental groups:

“Today is our last day to get think about soda taxes. You’re going to think about all the arguments you’ve heard over the past few weeks, and write a persuasive essay about soda taxes. At the top of your paper, you should see the question, ‘Should a tax be charged on soft drinks and that money go to lowering prices of healthy foods?’ Please circle what you think, yes, no, or undecided. [pause to let students write] Below that please circle how sure you are of your opinion: certain, very sure, so-so, not very sure, or not sure at all. [pause to let students write] You’re going to get the rest of the period today to work on writing a letter to Mayor De Blasio to convince him that your position is the better one. You may use the front and back of this paper, and if you need additional space, raise your hand and we will bring you another sheet of paper.”