How Construction of a Dialog Influences Argumentive Writing and Epistemological Understanding

Julia Zavala

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

COLUMBIA UNIVERSITY

2016

ABSTRACT

How Construction of a Dialog Influences Argumentive Writing

and Epistemological Understanding

Julia Zavala

Argumentive writing is not an easy skill to master. Students from middle school through college demonstrate weaknesses. In particular they fail to take a dialogic perspective, emphasizing their own position without considering addressing alternatives. Research has shown that engaging in dialog with peers is effective in enhancing students' argumentive thinking and writing. The present study examines whether college students (n=30) show similar benefits when asked to engage individually in a dialogic argumentive writing task. They were asked to construct a dialog between two people holding opposing positions on an issue. Students in a comparison group (n=30) were asked to write an essay on the same issue. Subsequently students in both groups were asked to write a brief TV script conveying their view.

Differences in students' argumentive skills produced in the dialogs and essays were examined. Results showed that the dialog group more frequently included opponent-directed statements (sum of Critical single evaluation, Compare, Integrate other, and Integrate own/other) and integrative statements (sum of Integrate own, Integrate other, and Integrate own/other) in their writing, compared to the essay group. Differences in students' writing of their TV scripts were also examined. On this assessment, the effect of the dialog largely disappeared, with students in both the essay and dialog groups focusing largely on their own position.

Students' level of epistemological understanding was also examined – that is, whether they regarded knowledge claims as largely facts (absolutist level), opinions (multiplist level), or

judgments subject to scrutiny in a framework of alternatives and evidence (evaluativist level). Level of epistemological understanding was assessed immediately after the writing task to determine if constructing a dialog influenced students to take on a more evaluativist perspective in which the need for comparison of multiple perspectives is recognized. Students who had constructed a dialog were more often assessed to be at the multiplist or evaluativist levels of epistemological understanding (and never at the absolutist level), compared to students who had written an essay rather than constructed a dialog.

Although the benefit of the dialogic writing task largely did not generalize to the more self-focused TV script writing task, these findings indicate that promoting a dialogic perspective, even without engaging in dialog with an actual person, can be beneficial in supporting argumentive thinking and writing and mature epistemological understanding.

TABLE OF CONTENTS

LIST OF TABLES iv

LIST OF FIGURES v

ACKNOWLEDGMENTS vi

CHAPTER I: Introduction 1

CHAPTER II: Literature Review 6

The Importance of Argumentation 6

Two Types of Argument 9

Rhetorical argument 9

Dialogic argument 10

What Constitutes Skilled Argument 12

Acknowledgment and weakening of opposing views 12

Use of evidence 13

Integration of information to back up an argument 16

The Development of Argument Skills 17

Research on Argumentive Writing 18

Goal instruction prior to argumentive writing 19

Discourse prior to argumentive writing 20

Epistemological Understanding 25

Relation between epistemological understanding and argument skills 26

Epistemological sensitization 28

Conclusion 30

CHAPTER III: Method 31
Participants 31
Design & Procedure 31
Assessments 32
Writing tasks 32
Initial writing task 32
Outcome writing task 34
Epistemological instrument 34
Demographic questionnaire 35
CHAPTER IV: Results 37
Coding 37
Coding scheme for writing tasks 37
Coding scheme for epistemological understanding instrument 39
Treatment Fidelity 42
Effect of Prior Experience 43
Argumentive Writing Outcomes Across Conditions 43
Differences in evaluations produced within the dialog versus the essay 43
Differences in evaluations produced in the TV Script 47
Individual differences in the use of each evaluation type in the dialog versus the essay 51
Individual differences in the use of each evaluation type in the TV script 53
Epistemological Understanding and Argumentive Writing 55
Effect of constructing a dialog on epistemological understanding 55

Relationship between GPA and epistemological understanding 56

Relationship between gender and epistemological understanding 57

CHAPTER V: Discussion 58

Purpose of Research 58

Summary of Findings 58

Limitations and Future Directions 59

Transfer of argumentive skill 59

Change in epistemological understanding 61

Effect of epistemological understanding on argumentive writing 64

Implications and Conclusions 64

References 67

Appendix A: Recruitment Flier 74

Appendix B: In-person Recruitment Script 75

Appendix C: Consent Form 76

Appendix D: Demographic Questionnaire 78

Appendix E: Writing Task Coding Process 79

Appendix F: Livia Problem Coding Scheme 82

Appendix G: Livia Problem Coding Process 86

LIST OF TABLES

Table 1 - Coding Scheme for Writing Tasks ……38

Table 2 - Coding Scheme for Level of Epistemological Understanding……...41

Table 3 - Demographic Variables…….……42

Table 4 - Fixed Coefficients and Random Effects for Each Evaluation Type Produced in the Dialog and Essay…….……...46

Table 5 - Fixed Coefficients and Random Effects for Each Evaluation Type Produced in the TV Script……..49

Table 6 - Percentage of Participants Using Each Evaluation Type in the Dialog and Essay……..…52

Table 7 - Percentage of Participants Using Each Evaluation Type in the TV Script……54

LIST OF FIGURES

Figure 1 - Toulmin's structure for analyzing an argument……..……...10

Figure 2 - Percentage of participants in each epistemological understanding level by group…...56

Figure 3 - Differences in GPA across level of epistemological understanding…….57

ACKNOWLEDGMENTS

So many people have helped me along this great journey and I am very thankful for each and every one of them.

My deepest appreciation to Dr. Deanna Kuhn, for guiding me throughout this process and for all of her support and kindness—I couldn't ask for a better adviser. Thank you to my committee members, Dr. Stephen Peverly, for challenging me to think more critically about my research, Dr. Bryan Keller, for helping me to figure out the best way to analyze my data, Dr. Bill Fifer, for making me think about my dissertation topic beyond its domain and for being a mentor who truly inspired my love for research over the past 8 years, and to Dr. Sarit Barzilai, for all of her thoughtful input. Also, a special thank you to Dr. Lisa Ecklund-Flores, for her continuous mentorship and encouragement throughout my years at Mercy College and Teachers College.

Thank you to my amazing lab mates who supported me along the way. You made it that much more fun and exciting on those tough days.

Thank you to my parents, Joe and Charity Zavala who always put education first. Your constant encouragement and support kept me focused and motivated me to never give up. Thank you to my brother Joseph and sister Jennifer for believing in me.

To my husband, Kami Hadavi, I cannot thank you enough for being by my side. You have gone through all the ups and downs with me and I am so thankful for having you there at each and every step of the way. Thank you for being my audience every time I needed to practice and for being my peer reviewer every time I needed my dissertation proofread. I love you so much and am so grateful to have had you with me on this journey. Thank you to our son, Darian, who will be arriving soon—you inspired me to work harder, and with more passion each and everyday.

DEDICATION

Dedicated to my son, Darian Hadavi. Always think well.

CHAPTER I:

Introduction

To know how to argue effectively is a critically important skill to possess. People need to know how to argue in order to understand and anticipate others' views (Kuhn & Crowell, 2011; Reznitskaya et al., 2001), to think critically (Benesch, 1999; Walton, 1989), to communicate effectively (Mercier & Sperber, 2011), and to persuade (O'Keefe, 1999; Wolfe, Britt, & Butler, 2009). It is important to ensure that young people master argumentation so that they are prepared for college, future careers (Common Core State Standards, 2010) and their roles as citizens in society (Kuhn, 2005).

According to Graff (2003), undergraduate students find it hard to make sense of their professors' intellectual discourse. They are expected to join in discussion, but they do not have the intellectual tools to "play the game". They learn to accumulate facts prior to entering college, but not to question or oppose the information they acquire. To be successful in school, they must "learn to join the argument culture" (p. 95) and to demonstrate they are capable of argumentive thinking before entering college. Several colleges and universities require students to take the SAT, which measures literacy and writing skills that are considered essential for academic success in college. Skill in argumentation is necessary for the writing section of the SAT. It asks high school students to write an essay in which they are to take a position and provide reasons and evidence to support their claims (CollegeBoard, 2015). Many graduate programs require students to take the Graduate Record Exam (GRE) that similarly includes a section that requires students to write an argumentive essay in which they are expected to be able to propose alternative explanations, as well as support their own claims and critique opposing

claims (Educational Testing Service, 2015). In regard to careers, increasingly professional positions require skill in argumentive writing. Regardless of vocation, in order for citizens to participate in a democracy, it is their responsibility to contribute to reasoned discourse regarding the society in which they live (Hileman, 2006). The decisions they make not only affect themselves in the present, but also those of future generations.

The skills required for effective argument are not innate; they develop (Kuhn, Zillmer, Crowell, & Zavala, 2013). Research has shown that most students are not effective expository writers. The National Report Card (2011) found that 74% of eighth graders and 73% of twelfth graders scored below proficient. Cooper et al. (1984) found that just 16% of college freshmen were able to produce counterarguments and rebuttals in their writing. In order to be successful argumentive writers, students must be skilled in the thinking that underlies argument. Larson, Britt, and Kurby (2009) found that undergraduate students struggled to identify unsupported, unwarranted, and acceptable arguments, performing more poorly than a comparison group of graduate students. Given these data, it is clear that students' argument skills need to be a focus of educators' attention at all education levels (Common Core State Standards, 2010; Kuhn, 2005). The Common Core (2010) standards now stress specific goals in regard to argumentive writing. For instance, by twelfth grade students should be able to generate claims and counterclaims and provide evidence to support each. However, the Common Core standards do not advise educators how to meet these standards. Endorsing the standards alone is not sufficient to enhance students' argumentive skills. Effective methods for doing so are necessary.

Recent research shows that argument skills acquired through peer-to-peer discourse over spans of five weeks (Reznitskaya et al., 2001) and three years (Kuhn & Crowell, 2011) show positive transfer to argumentive writing. Building on this dialogic approach, in the present

research I examine the effectiveness of asking students to construct their own dialogs between two opposing positions, essentially generating arguments for both sides. In an unpublished pilot study, I examined such a method with middle school students. The experimental group (n=62) individually constructed a dialog and the comparison group (n=27) simply wrote an essay, prior to a further writing task (outcome measure). Those who constructed a dialog made more comparative evaluations of the two opposing sides of an issue and made fewer unsupported claims in the outcome writing task. Furthermore, 63% of those who constructed a dialog, vs. 48% of those who did not, included statements aimed to weaken the opponent at least once. Both groups were experienced arguers as they had engaged in the multiple-year argument curriculum developed by Kuhn and colleagues (Kuhn & Crowell, 2011; Kuhn, Hemberger, & Khait, 2016). They therefore already had developed rudimentary argument skills, which the construction activity served to scaffold and reinforce. In the present work, I ask whether the method would be successful as well among students who had not had the benefit of exposure to an argument curriculum and, also, who were adults rather than adolescents.

Another factor to be examined in the present work is the role of epistemological understanding of argument. Argument skills must be valued if they are going to be applied (Kuhn, Hemberger, & Khait, 2014). The foundation for such valuing is epistemological understanding (Kuhn et al., 2014; Kuhn, 2005). Perry (1970) identified three broad levels of epistemological understanding (as cited in Kuhn, 1992, p. 167). The first level is least mature of the three and is labeled the absolutist level. At this level, judgment and interpretation are lacking; knowing is completely objective – a matter of ascertaining empirical facts. The second is the multiplist level, at which knowing becomes completely subjective. One recognizes that people may have different opinions, but they are all regarded as equally valid. No one opinion is

more right than another. The most mature epistemological level is the evaluativist. At this level, knowing achieves an integration of the objective and subjective. Although people may have different opinions, some opinions can be more valid than others, as a function of how well they are supported by evidence and arguments. Since argument skills must be valued in order for them to be applied, epistemological understanding has been regarded as a good predictor of argument skills (Kuhn, 1991; Kuhn & Park, 2005; Kuhn, Wang, & Li., 2011; Mason & Boscolo, 2004; Mason & Scirica, 2006; Weinstock, 2010; Weinstock and Cronin, 2003).

Specifically, in the present study we examine the effect of a dialogic writing task on epistemological understanding. Does the dialog construction activity support a more mature epistemological understanding? Research has shown that epistemological understanding is sensitive to influence by the nature of the task participants engage in (Kienhues, Bromme, & Stahl, 2008; Kienhues, Stadtler, & Bromme, 2011; Porsch & Bromme, 2011). In order to better examine the connection of the dialogic writing task and epistemological understanding, college students were recruited for the present study as they display more variation across the three levels of epistemological understanding than do middle school students, who are usually at the lower levels (Kuhn et al., 2011).

There are thus two main goals motivating the present study. One is to determine if a dialogic frame enhances argumentive writing. It is hypothesized that constructing a dialog will promote superior argument skills in an argumentive writing task, because the dialogic nature of argument has been highlighted by the construction activity. The second goal is to determine if the dialogic frame will affect epistemological understanding, and, if so, whether such effects are specific to epistemological level. It is hypothesized that the students who construct a dialog will

demonstrate more mature epistemological understanding as the dialog promotes active thinking and comparison of two opposing views.

CHAPTER II:

Literature Review

The Importance of Argumentation

Every decision, judgment, inference, and opinion is made with at least an implicit degree of argumentive thinking. Kuhn (1992) suggests that one's understanding of how one knows what they know falls on a spectrum where at the more rudimentary side there is little to no degree of argumentive thinking and one views knowledge as certain based on direct observation that cannot be refuted. On the more developed side there is a great degree of argumentive thinking; one views knowing as a process in which claims are judged based on evidence and arguments that are evaluated and reflected upon. Supporting and opposing evidence is weighed in relation to different claims. Understanding thinking as argument supports good thinking (Mercier & Sperber, 2011) and being a good thinker is essential to many aspects of life. The philosopher Walton (2013) called argumentation "an important skill to learn for everyday life, law, science, politics and business" (p. 1).

In education, the Common Core State Standards (2010) have stipulated that by 12th grade students must be able to acknowledge strengths and weaknesses of both own-side and opposing-side claims as well as be able to use evidence to support and weaken claims. Doing so is essential to an understanding of concepts (Osborne, 2009, December-2010, January). Scheurman (2006) stated, "[p]owerful learning occurs when students feel the significance of an explored idea" (p. 30). For instance, students that discuss and debate historical issues understand the content more deeply (Scheurman, 2006). Through argumentation, students learn to consider new or alternative views (Macagno & Konstantini, 2013), which also deepens understanding (Alverman, Qian, & Hynd 1995). Considering alternative views is essential to argument, but

people frequently ignore alternative hypotheses and evidence while allowing their own beliefs to dominate (Byrnes & Dunbar, 2014). Argumentation in pre-college and college education is important not only for deepening understanding of content, but also to help prepare students for the workplace, where many forms of argument are essential.

In the workplace, the skills required today involve capabilities such as "adaptive problem solving, critical thinking, complex decision making, ethical reasoning, and innovation" (Assessing 21st Century Skills: Summary of a Workshop, 2011, p. vii). This workshop examined data from 1969-1998 and found that *routine manual skills* like assembly line work rose from 1969 to 1980, but began to steadily decline after that. *Routine cognitive skills* (such as bookkeeping and filing) did not change from 1969 to 1980, but declined from 1980 though 1998. However, *complex communication* which involves "eliciting critical information and conveying a convincing interpretation of it" (p. 8) and *expert thinking* which involves "identifying and solving new problems" (p. 8) rose from 1969 through 1998. Deborah Boisvert, a researcher for the Boston Advanced Technological Education Connection (BATEC), presented at the workshop and further confirmed that research continues to show that critical thinking and problem solving skills are in high demand by employers. Argumentation is a skill that provides the foundation for critical thinking and promotes problem solving. It enables communication (Mercier & Sperber, 2011) that can resolve issues through persuasive dialog (Walton, 2013). For instance, lawyers, scientists, historians, businessmen/women, and medical doctors make arguments to persuade as well as draw conclusions from evidence in the work they perform. Teachers need to be able to develop students' argumentation in the classroom, but they are usually not trained in how to engage in argumentation with students (Osborne, 2009, December-2010, January).

Hence the importance of focusing on argumentation in education is critical, so people can enter the workforce with these skills in place.

Citizenship is necessary for maintaining a democratic society. Citizens have a duty to be involved in the society in which they live. President Abraham Lincoln (1863) stated that democracy is "government of the people, by the people, and for the people" (as cited in Hileman, 2006, p. 2). Government *of the people* indicates that people have the right and power to govern themselves, *by the people* indicates that people are willing to participate in the government they establish, and *for the people* indicates that the representatives elected by the people must govern in the people's best interest (Hileman, 2006).

Educating citizens to be effective arguers is thus essential for maintaining a democracy. A democratic society advocates for the citizens, but in order to do so the society must engage in critical-analytic thinking (about claims and evidence) in making collective decisions (Brookfield, 2012). Knight (1959) claimed that in order to contribute to being a responsible citizen one must possess the ability to engage in intellectual inquiry and discussion. Kuhn (2005) explains more specifically that inquiry and argument are the necessary skills in developing the mind of the individual so they can effectively not only face their own personal challenges, but also contribute to society.

In sum, argumentation is important in people's lives and it is an important skill to be mastered so that people may better understand, communicate and think in the various dimensions of life. Learning how to effectively argue better prepares individuals for college, careers, and citizenship, in essence for life itself.

Two types of Argument

There are two distinct types of argument. One is rhetorical argument and the other is dialogic argument. Rhetorical and dialogic arguments both include a social aspect; however, an individual constructs a rhetorical argument in order to persuade an audience, whereas a dialogic argument entails discourse between two (or more) people.

Rhetorical argument. Using the definition of argument from the *American Heritage Dictionary* (1981), Kuhn (1992) defines rhetorical argument as "a course of reasoning aimed at demonstrating the truth or falsehood of something" (p. 157). The argument is a *product* generated by an individual. In the revised edition of *The Uses of Argument*, Toulmin (2003) described in detail the six components that go into a rhetorical argument. He first explained that *conclusions* (or claims, C) need to be backed up with *data* (or facts, D). This answers the question "What do you got to go on?" (Toulmin, 2003, p. 90). In order to answer "How do you get [to that conclusion]?" (Toulmin, 2003, p. 90) a *warrant* (W) must be made. This helps explain the legitimacy between the data and the conclusion. However, in some cases a warrant is not enough so one will have to *qualify* (Q) the conclusion. To avoid someone not accepting the warrant made, one must *back* (B) it up with other forms of data. In addition to these components, one has to consider what other people may say so they must identify *rebuttals* (R). Figure 1 depicts Toulmin's structure for analyzing an argument.

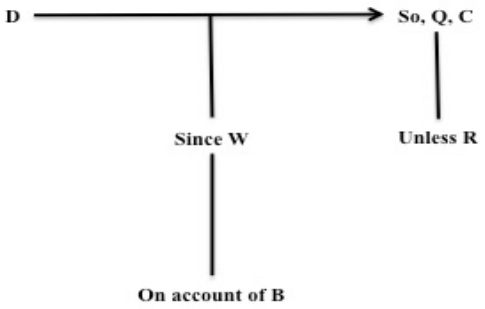


Figure 1: Toulmin's structure for analyzing an argument

Toulmin (2003) used the following example to elegantly link all six components: "Harry was born in Bermuda (D) so, presumably (Q), Harry is a British subject (C). Since a man born in Bermuda will generally be a British subject (W) on account of the following statutes and other legal provisions [the arguer would state them here] (B). So presumably, Harry is a British subject unless both his parents were aliens [or] he has become a naturalized American (R)" (Toulmin, 2003, p. 97). These elements constitute the basic structure of a rhetorical argument.

Dialogic argument. Dialogic argument can be defined as a dialogue between two people who hold opposing views (Kuhn, 1992). In other words the argument is a *process* involving two (or sometimes more) individuals. Walton (1996, 2013) called this the dialectical structure. He explained that it is "a series of back-and-forth, connected sequence of speech acts formed into an ordered profile that points in a certain direction" (Walton, 1996, p. 181). The sequence that takes place can be illustrated in the following example described by Walton (1996): Person A sends out a message to person B and person B sends out a message back to person A, but that message from person B is influenced by person A's message. Therefore it is dependent on the other person's message. For instance, a person may "corner" an opponent and ask them to clarify what they had just said (commit) so they can counter them more effectively (weaken) (Felton & Kuhn, 2001). This sequence satisfies the two goals of persuasive argumentation which are: 1) to get the

opponent to commit to premises that support the arguer's own position; 2) to weaken the opponent's claims by challenging them (Walton, 1989; Walton 1996). These goals involve fulfilling the burden of proof (or using evidence to support or weaken claims). Walton (1996) claimed that, "dialogue can provide the testing ground for the contending theses that represent both sides of the issue. This testing to determine the stronger and weaker arguments in persuasion dialogue is analogous to the testing of a scientific hypothesis by subjecting it to the experimental evidence produced in an inquiry" (p. 179). Persuasive argumentation can be adversarial; however, the arguers must be open to defeat when they come to realize that the opponent has stronger arguments that fulfill the burden of proof (Walton, 1996). Remaining closed-minded to the opponent's arguments will not accomplish anything; there is no point in arguing if the opponent's view is not considered.

There is a close link between rhetorical and dialogic argumentation. Both require the justification of a claim, acknowledgment and weakening of opposing views, and the weighing of evidence on both sides. Rhetorical argument, however, is at risk of favoring one side over others when weighing evidence for opposing views. As Kuhn (1992) put it, "the weighing of positive and negative evidence leads to favoring one assertion over its alternatives" (p. 157). Dialogic argumentation, however, better supports an inclination to consider the opposing side, since the arguers are constantly making exchanges and considering evidence to support and weaken claims.

The present study prompts students to incorporate both rhetorical and dialogic argumentation in the experimental group and only rhetorical argument in the comparison group. Those in the experimental group are asked to construct a dialog between two characters who argue for two opposing positions (dialogic argumentation). Those in the comparison group

represent only their own position. The dialogic structure we ask participants to construct is constructed by a single individual, rather than consisting of a dialog between two people, but we hypothesize it may be just as effective. As Walton (2013) explained, "whether one agent is involved or a group of agents is engaged in deciding what to do or to accept based on the evidence, rational thinking is best seen as a dialogue process in which arguments are put forward by one side and critically questioned by the other side" (p. 4).

What Constitutes Skilled Argument?

The key objectives of skilled argument are to weaken an opponent's position and strengthen one's own (Walton, 1989). However, the opponent's view must be acknowledged so that the opposing claim can be weakened. To ensure that the weakening argument is strengthened, evidence must be used to back it up. Moreover, integrating several pieces of evidence or information to back up an argument is more cognitively demanding yet effective than simply linking one argument to one piece of information.

Acknowledgment and weakening of opposing views. Kuhn (1992) stated "[a]n argument supporting an assertion is empty, indeed superfluous, unless there is an alternative—an opposing assertion" (p. 157). Arguments that only address one's own view are considered to possess *myside bias* (Wolfe, 2012). The absence of the opposing perspective makes the argument weak (Wolfe et al., 2009). The opposing position needs to be acknowledged. Otherwise, it is left unexamined and hence never compared to one's own position.

Weakening a claim is a critical aspect of skilled argumentation. In order to do so, one must acknowledge the opposing position and seek to weaken it. One can weaken one's own claim as well as weaken the opponent's claim. Outlining weaknesses in one's own claim makes for a more persuasive argument, as it lays the way for counterarguments (Wolfe & Britt, 2005 as

cited in Wolfe et al., 2009). However, weakening the opponent's claim is the major use of counterargument (Crowell & Kuhn, 2014; Felton, 2004; Felton & Kuhn, 2001; Kuhn, Goh, Iordanou, & Shaenfield, 2008; Kuhn & Udell, 2003). A counterargument is a direct response to the opponent's claim that expresses disagreement and is aimed at undermining it.

However, in the present study, one aspect of skilled argumentation is to address the opponent and critique their position. Kuhn et al. (2013) conducted a study using the same dialogic writing task as the present study (where students construct a dialog between two expert arguers on who should be mayor of a troubled city. All students were informed of the problems in the city and the solutions each candidate proposed. The solutions were considered the evidence.) They compared students' argument skill in the diaogic writing task and found that students who participated in a 3-year argument curriculum more frequently addressed and weakened the opponent than those in the control group. Moreover, in the pilot study that I conducted (mentioned in the introduction), students' dialogs and TV scripts consisted of more statements that acknowledged the opponent and aimed to weaken their position compared to students who wrote an essay. In addition to addressing and weakening the opponent, using evidence to weaken a claim strengthens an argument. As Walton (1996) put it, the use of evidence fulfills the burden of proof.

Use of evidence. Including evidence in an argument strengthens it and makes it more persuasive (Walton, 1988). It is typically used to support one's own claim but is also used to weaken the opponent's claim (Kuhn et al., 2013; Moore, 2013), although this use is less common and requires more skill. Middle school students who engaged in an argument curriculum used evidence that supported their own claims significantly more than they used evidence to weaken their opponent's claims (Kuhn et al., 2013; Moore, 2013). This is typical, as adolescents tend to

be more inclined toward myside bias, which requires less cognitive effort (Klaczynski & Gordon, 1996). However, in skilled argument, evidence is used to weaken the opponent's claim (Kuhn et al., 2013; Moore, 2013). Although a harder skill to achieve, it can be done with practice. Kuhn et al. (2013) found that students who engaged in an argument curriculum for three years produced arguments in their dialogs that had significantly more evidence that weakened their opponent's claims, compared to a non-participating control group.

To use evidence skillfully requires the ability to coordinate claims and evidence. To successfully coordinate claims and evidence one must recognize them as distinct entities (Chinn & Brewer, 1996; Kuhn, 2010). Coordinating the two involves the ability to evaluate new evidence that contradicts a preexisting theory and to accept the new evidence and revise that theory (Chinn & Brewer, 1993). Preexisting beliefs about scientific topics largely conflict with the science students learn in school (Chinn & Brewer, 1993). One who is not a skilled arguer will more readily accept evidence that confirms a prior theory and more harshly scrutinize evidence that disconfirms it (Lord, Ross, & Lepper, 1979). However, when there is a contradiction between old and new information, knowledge can be acquired and theories can be changed, but it is a challenging task, especially because people are more inclined to assimilate new evidence that confirms a theory rather than accommodate and revise old theories. There are several ways in which people respond to discrepant evidence (Chinn & Brewer, 1998). They may either ignore it if it doesn't fit their own theory, refute it for the same reason, consider that it might apply to another theory, decide that it may be something to work with later on, interpret it differently from which it was intended, reinterpret it and decide to make some changes to one's own theory, question the validity of the data, or lastly accept the discrepant evidence and revise one's own theory.

Revising a theory is challenging, and there are several factors involved (Chinn & Brewer, 1993). Firmly established, prior investment in the theory impedes change. However, if a new, plausible theory is available, revision is more likely to occur. It can also occur if the anomalous data are credible, unambiguous, and supported with multiple experiments. Lastly, deep processing, by means of discussing everyday experiences about the theory and requiring students to be accountable for justifying one's own theory can also influence theory revision. Simply accepting facts at face value, or assuming everyone has a right to their own opinion is unlikely to change a theory. One must be able to think deeply about the new evidence and how it relates to the theory.

Chinn and Brewer (1996) conducted a study that examined how undergraduate students evaluated data. This study sheds light on the coordination of theory and evidence in action. Each student was presented with a theory and evidence to support it and subsequently presented with a competing theory also containing supportive evidence. For example, some students read that dinosaurs were warm-blooded (theory) and were presented with the evidence to support that theory. They then read some competing evidence that supported the theory that dinosaurs were cold-blooded. The authors found that when students evaluated data they tended to discount evidence by identifying alternative causes. For example, the evidence that supported that dinosaurs are cold-blooded was that the skulls have small indentations where the pineal system would have been and because it was small compared to warm-blooded animals. One student discounted this evidence by explaining that those indentations could have been caused by weathering over time. By discounting, by means of invoking an alternative cause, the student maintained the theory that dinosaurs were warm-blooded. However, participants who evaluated and accepted the competing data were not able to identify "a reasonable alternative cause"

(Chinn & Brewer, 1996, p. S217). These participants weighed the likelihood of alternative causes, which did not stand up to the new competing evidence. Accepting new competing evidence may allow for accommodation of the current theory, while acknowledging that there is a discrepancy between the new evidence and the currently held theory. It takes metacognitive awareness and control to revise a theory in relation to new and old evidence. However, if the new competing evidence is simply assimilated and no theory revision takes place (as in the case of the above student who discounted the new evidence), new understanding does not emerge (Kuhn, 2010). In sum, the ability to use evidence strategically is very important, as is the ability to effectively coordinate theory and evidence.

Integration of information to back up an argument. Another aspect of skilled argumentation is to integrate several pieces of evidence or information to support or weaken an argument. Information used in an integrative rather than piecemeal way is more cognitively demanding, yet more effective as it strengthens an argument. Students who are presented with multiple sources and pay attention to the links between sources and claims are more inclined to generate arguments that acknowledge multiple viewpoints (Barzilai, Tzadok, & Eshet-Alkalai, 2015) and the ability to describe and evaluate viewpoints promotes integration of information (Barzilai & Eshet-Alkalai, 2015). Kuhn et al. (2013) found that students who were part of an argument intervention integrated more information to support or weaken a candidates' position, while students' dialogs from the control group consisted of more one-sided, single pieces of information. In the pilot study that I conducted, more individuals integrated multiple pieces of information to critique the opponent in their dialogs compared to students' essays.

The Development of Argument Skills

Argument skills are not innate; they are acquired (Kuhn et al., 2013) through education and practice. Felton and Kuhn (2001) found that community college students are better arguers than middle school students, and Kuhn (1992) found that people of various ages who have gone to college are better arguers than those who have not gone to college. Also, students from advantaged schools do better than those from disadvantaged schools, although only through junior high school age. However, low achieving students can develop argument skills through practice (Crowell & Kuhn, 2014). Crowell and Kuhn (2014) assessed students' argument capabilities prior to an argument intervention in which students engaged in a written dialog with another student on the topic of capital punishment. Each student took turns making their arguments in response to one another. The intervention spanned over three years and four topics each year. Each topic of discussion began with students working in groups to generate arguments for and against their positions. This was followed by electronic dialogs in which pairs of students on the same side of the issue argued against other pairs on the opposing side. The final event was a whole-class "Showdown" in which students debated one another. At the end of the 3-year intervention students' argument capabilities were assessed using the capital punishment topic, which had not been part of the intervention. The results showed that students who initially performed least well in argumentive discourse skill prior to engagement in the intervention showed the greatest improvements in their skill, and were performing at a level equivalent to other students after the intervention. Several researchers have found that after participating in similar argument interventions ranging from 8 weeks to 3 years, adolescents showed significant increases in use of counterarguments, compared to control groups (Crowell & Kuhn, 2014; Kuhn et al., 2008; Kuhn & Udell, 2003; Shaenfield, 2009). The intervention has

also been successful in increasing adolescents' use of evidence (Kuhn et al., 2013; Moore, 2013), including use of evidence to weaken the opponent's claim and to integrate several pieces of evidence to address a claim (Kuhn et al., 2013).

Research on Argumentive Writing

Writing serves as a way to communicate and express ideas. Argumentive writing, specifically, is effective in doing this. Skilled argumentive writing can be very impactful. Writing that includes counterarguments and rebuttals is more persuasive than writing that does not (O'Keefe, 1999; Wolfe et al., 2009). It is rated higher in quality, and the impression of the author is more positive to the reader (Wolfe et al., 2009). Including the opponent's viewpoints in argumentive writing reflects good thinking. Nussbaum and Kardash (2005) cite several studies showing that "(n)ormative models of good thinking include the evaluation of opposing viewpoints" (p. 157; also see Baron, 1988; Ennis, 1987, 1995; Scriven, 1976). However, people tend to possess myside bias in argumentive writing, not acknowledging the other side of an issue (Perkins, Farady, & Bushey, 1991). College students who consider both sides of an issue and believe that facts need to be evaluated include less myside bias in their argumentive essays (Wolfe, 2012; Wolfe & Britt, 2008). Moreover, one must understand the opposing position so it can be acknowledged and weakened with counterarguments (Hays & Brandt, 1992). This is a difficult task as a certain level of metacognitive thinking and understanding about the goals of argumentation is required (Kuhn et al., 2013). In *The National Report Card: Reading 2011,* 74% of 8th graders and 73% of 12th graders scored below "proficient" in writing skill. They had trouble persuading, explaining, and conveying experience. It is clear that argumentive writing needs to be fostered and supported. The place in which this can occur is in an educational setting.

The Common Core State Standards (2010) now require students to write argumentive essays. From kindergarden through grade 3, students should be able to develop opinions and provide reasons for those opinions. In grades 4 and 5 students should be able to support those reasons with facts. In grade 6 students should be able to "write arguments to support claims with clear reasons and relevent evidence" (CCSS.ELA-LITERACY.W.6.1). In grade 7, students should be able to "acknolwdge alternate or opposing claims, and organize the reasons and evidence logically" (CCSS.ELA-LITERACY.W.7.1.A). In grade 8, own and alternate claims should be clearly distinguished. From grades 9 through 10 students should be able to "develop claim(s) and counterclaims fairly, supplying evidence for each while pointing out the strengths and limitations of both" (CCSS.ELA-LITERACY.W.9-10.1.B). From grades 11 through 12, claims and counterclaims should be more "thoroughly" developed and the "most relevent evidence" used. Many researchers have studied how best to enhance argument skills in students' writing. Some focus on the importance of goal instruction (Ferretti, MacArthur, & Dowdy, 2000; Nussbaum & Kardash, 2005; Nussbaum & Schraw, 2007) while others focus on discourse (Kuhn & Crowell, 2011; Reznitskaya et al., 2001).

Goal instruction prior to argumentive writing. Goal instruction has shown to be effective in facilitating skill in argumentive writing. Specifically, the focus has been on generating own-side claims, counterarguments, rebuttals and increasing persuasiveness of the essay. Nussbaum and Schraw (2007) found that undergraduate students who were told what makes a good argument (i.e. position, supporting reasons, counterarguments, acknowledging arguments and counterarguments in the final conclusion) wrote essays that included a stronger integration of arguments and counterarguments, compared to a control group. Ferretti et al. (2000) found that specific goals enhanced students' writing over a general goal to persuade.

Fourth and 6th graders with and without learning disabilities were either asked to write a persuasive essay or to write an essay with specific goals in mind (generate reasons, counterarguments and rebuttals). There were no significant differences across conditions in the 4th grade students, but there were significant differences in the 6th grade students. Those who received the specific goals wrote essays that were more persuasive than those who did not. They also included significantly more alternative positions, alternative reasons, and rebuttals. However, the overall persuasiveness for their essays was only an average of 2.5 of 7, indicating they still need a lot of support. Also, only 54% of 6th graders in the elaborated goal condition included alternative propositions and rebuttals in their essays. Nussbaum and Kardash (2005) conducted a study in which they administered goal instructions to undergraduate students. They found that students who were instructed to write counterarguments and rebuttals in their essays did in fact generate significantly more of these in their writing compared to a control group and a group who were asked to simply produce reasons for their claims. Also, those who were given a text that outlined the pros and cons of both sides of an issue and were not asked to persuade an audience of their position produced more counterarguments than those in three other conditions—1) those who did not receive a text and not asked to persuade an audience of their position, 2) those who were asked to persuade and given a text, or 3) those who were asked to persuade and not given a text. The reason may be that "my-side bias is rooted in students' belief that raising counterarguments reduces overall persuasiveness" (Nussbaum & Kardash, 2005, p. 164; also see Santos & Santos, 1999). Thus asking one to write a persuasive essay would inhibit one's generation of counterarguments.

Discourse prior to argumentive writing. Researchers who have engaged students in discourse prior to persuasive argumentive writing showed that it does enhance students' ability

to produce counterarguments, rebuttals, use of evidence (Reznitskaya et al. 2001), and use of evidence for the purpose of weakening the opponent's claim (Khait, 2014). It also enhances students' ability to acknowledge positive attributes of the opponent's position and negative attributes of one's own position (Kuhn & Crowell, 2011) which lays the way for counterarguments and rebuttals (Wolfe & Britt, 2005 as cited in Wolfe et al., 2009). In sum, research shows that engaging in discourse (via face-to-face and/or electronic dialogs) enhances students' argumentive writing (Kuhn & Crowell, 2011; Khait, 2014; Reznitskaya et al., 2001).

Kuhn and Crowell (2011) found transfer of argument skills from discourse between peers to essays of individuals. After three years of an argument intervention, students showed significant gains in dual-perspective thinking in writing essays, compared to a control group. They were able not only to address positive attributes of their own position, but also to address negative attributes of the opposing side. An integrative perspective, which incorporates negative components of the own-side position and/or positive components of the other side's position, did not appear until the end of year three, and only in the experimental group was it evident. Kuhn and Crowell (2011) replicated the study using only year-one and year-two students; the results were similar for dual-perspective arguments in that they were used significantly more by the experimental group than the comparison group; however, some students showed integrative perspectives at the end of year two, which was not evident in the first cohort.

Reznitskaya et al. (2001) conducted a study in which 4th and 5th graders engaged in a 5-week Collaborative Reasoning class that focused on discourse amongst the students while the teachers coached rather than led. Upon completion of the course, each student wrote an individual persuasive essay about a text. The results showed that those who participated in the intervention wrote superior essays than the control group. Specifically, they generated

significantly more own-side arguments, counterarguments, rebuttals, and included more information from the text as evidence. Khait (2014) conducted a study in which she examined students' use of evidence in their essays. A writing task was administered to middle school students who participated in either a 1- or 2-year argument curriculum rich in fostering face-to-face and electronic dialogs between opposing pairs. Both groups wrote persuasive and discrediting essays. Results showed that those in the 2-year curriculum used evidence to weaken (the opponent's position) more than those in the 1-year curriculum. However, participants in the 2-year curriculum did not show any difference across the two essay types in evidence used to weaken.

The success of this transfer from dialog to writing may lie in the fact that discourse provides a "naysayer" (Graff, 2003). With a naysayer, an argument schema can be developed as own-side arguments are generated, counterarguments are anticipated and subsequently rebutted, and evidence is considered (Reznitskaya et al., 2001). Written records (Kuhn & Udell, 2003) or electronic transcripts (Crowell & Kuhn, 2014; Iordanou, 2012; Kuhn & Crowell, 2011; Kuhn et al., 2008; Kuhn et al., 2013; Shaenfield, 2009) used to conduct discourse facilitate metacognitive processing by allowing reflection on what was said, enhancing the ability to produce counterarguments and rebuttals, and use evidence.

Research by Reznitskaya et al. (2012) did not show transfer of argument skill from peer-to-peer discourse to individual argumentive writing. The intervention used was different than that used previously by Reznitskaya et al. (2001). The intervention conducted by Reznitskaya et al. (2012), based on the Philosophy for Children program developed by Matthew Lipman (1969), did not teach based on an argument schema as the Collaborative Reasoning intervention did. Reznitskaya et al. (2012) explain that perhaps more direct instruction may be needed for students

to develop argument schema. However, Reznitskaya et al. (2001) did not use direct instruction; rather the schema was scaffolded by the type of intervention that was implemented. The discourse itself scaffolded a schema for argument that transferred to students' writing.

Engaging in discourse during class hours may be too time consuming for some teachers. As mentioned earlier, unpublished research by the author used a method in which students individually constructed a dialog. It is similar to a "Pre-write" that has been used in the dialogic argument intervention (Crowell & Kuhn, 2014; Kuhn & Crowell, 2011; Kuhn et al., 2014; Kuhn et al., 2013). The purpose of the "Pre-write" is to facilitate transfer from individual dialogic frame to argumentive writing by encouraging students' thinking about both sides of the issue and help to organize the production of arguments, counterarguments, and rebuttals. Students were asked to fold a piece of paper in half; one side was designated for own-side arguments and the other side for opposing arguments. They started out by writing an argument for their own side, then they wrote what the other side might say in response to that argument. They would then generate a response and continue thusly, going back-and-forth between the two positions until they were satisfied with the arguments they had in place for each side. The action of arguing in this back-and-forth manner aimed to faciliate counterarguments and rebuttals, thus producing more skilled argumentive essays that incorporated both sides of the issue. In the unpublished research by the present author, students in the experimental group (n=62) engaged in a task similar to the "Pre-write". They constructed a dialog between two characters arguing for different positions on a topic while those in the comparison group (n=27) only wrote an argumentive essay on the issue. On an outcome measure, participants wrote only from the perspective of their own position. Both groups participated in an argument intervention, but some (n=29) were in their first year and others (n=60) in their second. Those who constructed a

dialog used more comparative statements in the outcome measure (a TV script) than those who wrote an essay. Moreover, constructed dialogs themselves consisted of marginally significantly more opponent directed statements and fewer one-side statements than the essays themselves. The present study replicates this work among a group of older students who had not engaged in any argument intervention, in order to determine the effect of this treatment in an untutored, but older group (Kuhn et al., 2011; Weinstock, Neuman, & Glassner, 2006).

The extent to which one values argumentation is a further factor that supports skilled argumentive writing. Khait (2014) conducted a second study in which a group of students wrote both persuasive and discrediting essays. The participants were a naïve group of middle school students who had not participated in an intervention. They engaged in six face-to-face dialogs with peers over the span of two days. The group was divided into those who argued with an agreeing partner and those who argued with a disagreeing partner. There was no difference between groups in the use of weakening evidence across the two essay types. A possible reason could lie in students' epistemological understanding. Argument must be valued in order for argument skills to emerge. This is achieved only at the evaluativist level, as elaborated below. Research shows that middle school students are unlikely to have attained the evaluativist level compared to high school students (Kuhn et al., 2011; Weinstock et al., 2006) and college level students; thus variance across the three levels (absolutist, multiplist, and evaluativist) is limited in younger groups (Kuhn et al., 2011). However, the development of the three levels throughout adolescence and adulthood continues to exist across age groups (Moshman, 2014). Moreover, educational level predicts epistemological understanding (Kuhn, 1992; Weinstock et al., 2006). Therefore, college students were recruited for the present study, as they should provide more variation across the three levels of epistemological understanding.

Epistemological Understanding

Developing argument skills on their own will not guarantee that they will be applied in other situations beyond the one in which they were developed (Kuhn et al., 2014). They must be valued if they are to be used regularly across a wide range of contexts. The foundation for valuing these skills is epistemological understanding (Kuhn et al., 2014; Kuhn, 2005). Epistemological understanding entails beliefs about knowing and the nature of knowledge (Hofer & Pintrich, 1997; Kuhn, Cheney, & Weinstock, 2000). Researchers have extensively studied the three main levels of epistemological understanding—absolutist, multiplist, and evaluativist (Barzilai & Eshet-Alkalai, 2015; Barzilai, Tzadok & Eshet-Alkalai, 2015; Kuhn, 1991; Kuhn, 1992; Kuhn et al., 2000; Kuhn & Park, 2005; Kuhn et al., 2011; Mason & Boscolo, 2004; Mason & Scirica, 2006; Weinstock, 2010; Weinstock & Cronin, 2003; Weinstock et al., 2006). As described earlier, the least mature is the absolutist level, second is the multiplist level and the most mature is the evaluativist level. At the absolutist level, judgment and interpretation are unnecessary; knowing is direct and completely objective. One believes that there is only one definitive answer to a question, based on empirical facts (Moshman, 2014). At the multiplist level knowing becomes completely subjective. All opinions are viewed as freely chosen by the knower and hence equal to one another and not to be questioned. At the evaluativist level knowing achieves an integration of the objective and subjective. One believes that people may have different opinions; however some opinions can be more correct than others. What makes opinions more valid is how well they are supported by evidence and arguments. Determining this involves an inquiry process (Moshman, 2014) and the coordination of theory and evidence (Kuhn, 2010). One at the evaluativist level understands that an existing theory may be revised given new evidence. It is understood that knowledge is constructed based on evaluation and

judgments involving argumentation and evidence. Therefore, intellectual discourse becomes valued as a means to enhance understanding (Kuhn et al., 2011).

It is clear that one at the evaluativist level can appreciate and value argumentation. In essence, "it is argument that lies at the heart of the evaluative process" (Kuhn, 1991, p. 191).

Relation between epistemological understanding and argument skills. Epistemological understanding is associated with several aspects of argument. Epistemological level (absolutists, multiplist, or evaluativist) predicts to some degree the extent to which an individual values discussion. People at the evaluativist level value discussion (Kuhn & Park, 2005; Kuhn et al., 2011), few at the multiplist level do, and, Kuhn and Park (2005) found in their study that no one at the absolutist level valued discussion. Mason and Boscolo (2004) found that when presented with a text containing two opposing views on a topic, 10th and 11th graders who had achieved either an evaluativist or multiplist level of epistemological understanding were able to write concluding paragraphs for the text in which they developed an argument in favor of one side by acknowledging the other side and generating arguments against it. The participants were also able to recognize that there was need for more research on the topic. However, the researchers combined the evaluativist and multiplist levels rather than examining them separately, which would have been more informative as to which level best predicted the outcomes. Mason and Scirica (2006) conducted a study in which the two epistemological levels were examined separately. Eighth-grade students were given a text containing two opposing views. The students were instructed to generate arguments, counterarguments, and rebuttals in response to the text. They found that students at the evaluativist level generated better arguments, counterarguments, and rebuttals than those at the multiplist level.

Prior to this research, Kuhn (1991) found that those at the evaluativist level produced significantly more counterarguments than those at the other two levels. Other research has shown that epistemological understanding is associated with students' ability to identify the *ad ignorantiam* fallacy of logic (Weinstock et al., 2006). More students at the evaluativist level than at the multiplist or absolutist levels correctly identified the fallacy in arguments that claimed to be true because of the lack of evidence against them.

The use of evidence has been associated with epistemological level. Kuhn (1991) found that those at the evaluativist level generated more evidence for arguments than those at the absolutist and multiplist levels. The difference was not significant, but noteworthy as the direction implied that those at the evaluativist level were using more evidence. Since then, researchers have found a significant relationship between epistemological level and use of evidence; epistemological level of jurors' was associated with the use of evidence to support or weaken claims (Weinstock, 2010; Weinstock & Cronin, 2003).

Barzilai and Eshet-Alkalai (2015) did not find a direct relationship between epistemological understanding and integration of multiple sources in an argumentive writing task. They assesed epistemological level for 170 college students. Students were then randomly assigned to either read conflciting or converging blog-posts on seawater desalination, next they were assessed on their viewpoint comprehension, lastly they engaged in argumentive writing on the topic. The conflicting blog-posts consisted of opposing viewpoints on the pros and cons of seawater desalination and the converging blog-posts consisted of just the pros of seawater desalination. Viewpoint comprehension measured students' ability to correctly match excerpts from the blogs' content to the author, describe authors' viewpoints, and rate how reliable each author was and justify why. The argumentive writing task prompted students to write an essay

on whether Israel should encourage seawater desalination and to justify their postion. Integration of multiple sources measured argument structure (whether students justified claims, provided a one-sided argument, two-sided argument with weak justification or a two-sided argument with full justifcation of claims) and the number of sources that supported the argument.

The results did not find epistemological understanding to be a significant predictor of integration of multiple sources in the argumentive writing task, but it did reveal that these two variables were mediated by viewpoint comprehension. Thus, students' "understanding of multiple viewpoints" (p. 100) provides the basis for the positive association between epistemological understanding and the integration of multiple sources in argumentive writing.

It is thus clear that epistemological understanding plays an important role in argument. When one has achieved an understanding of knowledge at the evaluativist level, argument skills are more likely to develop (Kuhn, 1991), as well as the valuing of argument (Kuhn et al., 2000; Kuhn et al., 2011). Although multiplists and some absolutists value and demonstrate argument skill, they do not do so as well as an evaluativist. Therefore, it is important to foster a more mature level of epistemological understanding if skilled argumentive writing is to be achieved.

Epistemological sensitization. Research shows that participants who are exposed to opposing and challenging views may show change toward more advanced epistemological beliefs (Kienhues, Bromme, & Stahl, 2008; Kienhues, Stadtler, & Bromme, 2011; Porsch & Bromme, 2011). This change can be termed epistemological sensitization.

Kienhues et al. (2008) conducted a study in which they examined whether text containing two opposing views on the safety of DNA fingerprinting would change participants' naïve epistemological beliefs (comparable to the absolutist level) to more sophisticated epistemological beliefs (comparable to the evaluativist level). They hypothesized that this would occur as

participants' initial beliefs were challenged by the new content. They also hypothesized that participants' sophisticated epistemological beliefs would not change after receiving the refutational text as their beliefs should be similar to the content. Moreover, they hypothesized that participants' naïve or sophisticated epistemological beliefs would not change after receiving a one-sided text on DNA fingerprinting. However, if change did occur it would shift to a more naïve belief as the information is one-sided and presented as certain fact. So if the sophisticated epistemological beliefs group had any doubts about their own knowledge they may take the instruction as certain fact; hence their epistemological beliefs would shift to a more naïve level.

Students were surveryed online and epistemological beliefs were measured prior to the intervention. Two instruments were used to measure epistemological beliefs, one was the Domain-specific Epistemological Belief Questionnaire (DEBQ; Hofer, 2000) and the other was the Connotative Aspects of Epsitemologcial Beliefs (CAEB; Stahl & Bromme, 2007). The DEBQ measured stability, and certainty of beliefs and the CAEB measured texture (the perception of perceived knowledge) and variability (the variability of knowledge over time). Half of 58 participants held naïve epistemological beliefs and the other half had sophisitcated epistemological beliefs. Both groups were randomly assigned to either receive the refuational text or the one-sided text on the safety of DNA fingerprinting. Post-intervention, participants' epistemological beliefs were measured.

The results were mixed. According to the DEBQ, the sophisticated group that received the refutational text significantly shifted to a more naïve epistemological belief from pre- to post-intervention and the naïve group did not change. According to the CAEB, the naïve group that received the refutational text significantly shifted to a more sophisticated epistemological belief from pre- to post-intervention and the sophisticated group did not change. Moreover, the

sophisiticated epistemological belief group that received the one-sided text significantly shifted to a more naïve epistemological belief and the naïve group did not change. The results from the CAEB confirmed their hypothesis.

The construction of the dialog in the present study required participants to actively construct a dialog containing two opposing views on an issue. It thus may have supported growth in epistemological understanding. In the present study we could not directly assess whether epistemological level changed after the construction of the dialog (since no pretest was administered). Rather, we ask if the construction of the dialog influenced participants' epistemological undertsanding as assessed by post-dialog group differences, such that more individuals who constructed the dialog would be at the evaluativist level compared to those who did not construct a dialog.

Conclusion

Argumentation is an important life skill. It is necessary for students to know how to argue effectively so that they can be prepared for college, the workplace and citizenship. Certain argumentive writing standards have been established regarding what students are supposed to know by the time they graduate from high school, but there is no clear method for how to teach to those standards. However, engaging in discourse helps students develop the skills necessary for producing sound arguments in their thinking and writing. Also, the more one values argumentation, the more skilled one's arguments will be, making it an important goal to help students achieve epistemological understanding at the evaluativist level. The present study examines whether a dialogic frame supports better argumentive writing and mature epistemological understanding.

CHAPTER III:

Method

Participants

Sixty undergraduate students from a four-year college in a suburb of New York City were recruited for this study. Age ranged from 17-55 years (M = 23.4, SD = ±6.3). The minimum GPA was 2.08 and the maximum was 4.00 (M = 3.16, SD = ±0.46). All were matriculated students except for one. Fifteen percent were freshman, 12% sophomores, 30% juniors, and 42% seniors. Twelve percent of students had a double major. More than half were psychology majors (57%). The remaining 21 different majors had less than 7% of participants in each. Eighty-three percent were females and 17% were males. This reflects the gender demographic of the entire population of the college, which consists of 11,650 students (including part-time and graduate students) among whom 73% are females and 27% males. Fifty percent of the sample identified as Black or African American, 38% Hispanic or Latino, 5% White, 2% Other, and 3% did not report race/ethnicity. Students were asked if they had ever been part of a debate team or argument program and 82% reported they had not.

Participants were randomly assigned to one of two conditions using a random number generator, 0 being the comparison condition and 1 the experimental condition. Thirty participants were assigned to the experimental condition and 30 to the control condition.

Design & Procedure

Participants were recruited via fliers (see Appendix A) posted on campus, or in person (see Appendix B) from individual classrooms where the instructor gave permission. Once the student contacted the author, an appointment was made for the activity.

The activity took place in a quiet room and lasted about 45-60 minutes. Students were allotted extra time if needed. Upon arrival, the student was greeted and given a consent form (see Appendix C) to read and sign. A copy of the consent form was provided for his/her records. The student then proceeded to complete the assessments. Once he/she was finished, the student was thanked and given a $10 Dunkin Donut gift card.

Assessments

Writing tasks. The experimental group was asked to construct a dialog between two expert arguers (on who would be better as mayor in a troubled city). Those in the comparison group were asked to write an argumentive essay on the same topic. Information regarding the problems of the city as well as the actions each candidate proposed to take was presented in writing. Participants in both groups then proceeded to a subsequent task in which they were asked to write a TV script advocating their position in favor of one of the two opposing positions (candidate A vs. candidate B as the stronger candidate for mayor). The TV script was the outcome measure of argument skill.

Initial writing task. Participants in the dialog condition received a sheet containing this information and instructions:

> Ana Cruz and Maria Diaz are running for mayor of their troubled large city. Among the city's problems are high housing costs, teen crime, traffic, school dropout, and unemployment.
>
> Chuck and Doug are TV commentators arguing about who is the better candidate. **Write a script** of what they might say. They are both experts on the city; they are both expert arguers and evenly matched. So your script should present the most well argued debate you can construct.
>
> I. Begin your script like this:
>
> **CHUCK:** Cruz should be elected mayor because she'll do better than Diaz.

DOUG: I disagree, because xxxxxx

Then continue their argument, filling in what each one might say:
CHUCK:xxxxxx
DOUG:xxxxxxx
CHUCK:xxxxxx etc.

Here is some information about Cruz' positions. She promises to:
-create job training programs
-expand city parks
-raise teachers' pay
-open walk-in health clinics
-reduce rents
-impose a teen curfew
-employ senior citizens in city schools

Here is some information about Diaz' positions. She promises to:
-improve public transportation
-open more centers for senior citizens
-revise the high school curriculum
-build a new athletic stadium
-improve health care
-build more housing

(You are not required to include all the above topics in your script.)

Participants in the essay condition received identical information about the candidates but a different instruction:

Ana Cruz and Maria Diaz are running for mayor of their troubled large city. Among the city's problems are high housing costs, teen crime, traffic, school dropout, and unemployment.

Here is some information about Cruz' positions. She promises to:
-create job training programs
-expand city parks
-raise teachers' pay
-open walk-in health clinics
-reduce rents
-impose a teen curfew
-employ senior citizens in city schools

Here is some information about Diaz' positions. She promises to:
-improve public transportation
-open more centers for senior citizens

-revise the high school curriculum
-build a new athletic stadium
-improve health care
-build more housing

Write an argumentive essay in which you consider the merits of each of these candidates for mayor. (You are not required to include all the above topics in your essay.)

Outcome writing task. Next, participants from both groups were asked to imagine that they were appearing on a TV show and that they had to write a script of what to say:

You've been asked to appear on the TV show and make a case for either Cruz or Diaz as the best candidate. Who would you choose and what would you say in your 2-minute talk? Write a short script for yourself.

Epistemological instrument. Subsequent to the writing tasks, the participant was asked to follow along as the author read aloud a short excerpt—*The Livia Problem,* which described two accounts of the fictitious Fifth Livian War. One account was from a historian from North Livia and the other account was from a historian from South Livia. Participants were then asked to answer questions. The questions were asked aloud by the author and their answers were audio recorded, to save time. The Livia problem, described below, was developed by Kuhn, Pennington, and Leadbeater (1983) and subsequently used by Kuhn, Iordanou, Pease, and Wirkala (2008), Kuhn and Weinstock (2002), Leadbeater and Kuhn (1989), Weinstock, (1999), Weinstock (2006), and Weinstock and Cronin (2003) to measure epistemological understanding. It is ficticious to prevent the participants from favoring one account over the other, thus reducing bias, and to equate participants' with respect to prior knowledge (Kuhn et al., 1983). The participant was given a sheet containing this text and followed along as the interviewer read:

North and South Livia are two small countries that existed in the 1800s in Asia. There occurred a series of conflicts between the two countries, termed the Livian Wars. Here are two brief accounts of the Fifth Livian War that took place in 1878.

A brief account of the Fifth Livian War by J. Abman, National Historian of North Livia

On 19th July 1878, during a national ceremony in North Livia to honor one of their national leaders, ceremonies were interrupted by a sneak attack from South Livia. Thus began the Fifth Livian war. Because the North Livians were caught by surprise, they were unprepared at first and the South Livians won a few early battles. Then North Livia began to win. But before the North Livians could reach a final victory, a neighboring large country intervened to stop further bloodshed. Despite their early setbacks, the later sweeping victories of the North Livians showed they would have won had the fighting continued. As a result of this war, South Livians finally recognized anything they gained from North Livians would have to be worked out through peaceful negotiations. So ended the Livian Wars.

A brief account of the Fifth Livian War by N. Ivan, National Historian of South Livia

In the last war, North Livia had beaten South Livia, taken some of its land, and refused to leave. South Livia could no longer accept this situation and spent much money to strengthen its military. On 20th July 1878, the Fifth Livian war began. The war took place with rapid, dramatic victories for South Livia, resulting in great celebration. After these dramatic victories, the South Livians suffered some minor losses. But then a neighboring large country intervened to stop further bloodshed. Despite their later setbacks, South Livia's victory seemed assured because of its position of strength. As a result of this war, South Livians felt a new self-respect. They had felt embarrassed by their previous defeats, but now they had proven they were the equals of the North Livians. Because South Livians had achieved military respect, they were willing to work out differences through peaceful negotiations, thus ending the Livian Wars.

The interviewer then presented the following questions verbally and the respondent answered orally:

1. Can you summarize what the Fifth Livian War was about and what happened?
2. Are the two historians' accounts of the war different in any important ways? In what ways are they different?
3. Could both of the historians' accounts of the Fifth Livian War be right? If no—why not? If yes—How can that be? Is one of the historians' accounts of the Fifth Livian War more true than the other?
4. Could anyone be certain of what happened in the Fifth Livian War? If yes—how? If no—Why not?

Demographic questionnaire. Upon completion of the verbal interview, the participant was asked to complete a demographic questionnaire (see Appendix D), which asked them to report their age, year of study, major, GPA, gender, race/ethnicity, and whether they were ever part of a debate team or argument program.

CHAPTER IV:

Results

Coding

Coding scheme for writing tasks. In the initial writing task, each of the solutions proposed by the candidates, Cruz and Diaz, served as evidence to either support or weaken a claim. Also examined, was whether they integrated or used the evidence in a piecemeal way to support or weaken a claim. First, participants' dialogs, essays, and TV scripts were segmented into idea units and subsequently coded using the coding scheme developed by Kuhn et al. (2013). Seven distinct evaluations have been identified presented in Table 1 below. Expanding on the terminology of the original coding scheme, the codes are termed evaluations as students evaluate the evidence to link it to a problem and by doing so they make supportive or weakening claims and either integrate evidence or use single pieces of evidence to make those supportive or weakening claims.

To establish inter-rate reliability, idea units were first identified. Idea units were defined as statements that expressed a single idea or assertion. Two researchers segmented and coded 15% of data from a previous cohort to establish reliability. The percentage agreement for identifying segments was 100%. Each idea unit was subsequently coded using the coding scheme shown in Table 1. Percentage agreement between the same two independent coders was 91%. All disagreements were resolved through discussion. Appendix E provides a detailed description of how the coding scheme was applied.

Table 1

Coding Scheme for Writing Tasks

Code	Description	Example
Non-evaluative	Claim that does not link evidence to a stated problem.	Diaz plans to directly tackle some of our city's biggest problems, such as high rent and traffic.
Positive single evaluation	Evaluation of evidence in support of favored candidate that links a single piece of evidence to a single stated problem.	… she promised to impose a teen curfew so they couldn't commit as many crimes.
Critical single evaluation	Negative evaluation of evidence of non-favored candidate that links a single piece of evidence to a single stated problem.	She'll increase crime by making criminals strong from her athletic stadium.
Integrate own	Evaluation of evidence in support of favored candidate that links multiple pieces of evidence to problems.	Diaz promises to revise high school curriculum. This lowers the rates of dropouts because students become more interested and want to stay in school. Also, since it won't cause dropouts, teens will be in school so there won't be teen crimes.
Integrate other	Negative evaluation of evidence of non-favored candidate that links multiple pieces of evidence to problems.	Those job training programs don't guarantee that you get a job. They don't guarantee that you will get enough money to pay for rent.
Compare	An evaluation of evidence comparing one piece of evidence from each of the two candidates in regard to one problem.	Diaz is improving public transportation. That resolves the traffic; Cruz says nothing about transportation in her plan.

Integrate own/other	Evaluation of evidence comparing multiple pieces of evidence from each of the two candidates in regard to problems.	…Diaz is making the high school curriculum better. This means that there isn't going to be teen crimes. Cruz says that she is going to have a teen curfew but that wouldn't work. They can still do crimes in the day and wouldn't listen to the curfew. Diaz is also going to build an athletic center so the teens can be busy.

Analyses were based on the 7 distinct evaluations. In additional analyses, evaluation types addressing the non-favored candidate were combined and considered *Opponent Directed* evaluations: Critical single evaluation + Integrate other + Compare + Integrate own/other. Evaluations that integrated multiple pieces of information to make a claim were combined and considered *Integrative Evaluations*: Integrate own + Integrate other + Integrate own/other.

Coding scheme for epistemological understanding instrument. All participants' answers were coded using a coding scheme developed by Weinstock (1999) and also used by Weinstock & Cronin (2003) in order to assign a level of epistemological understanding. It includes 22 dimensions. See Appendix F for a full description of the coding scheme, adapted from Weinstock (1999), p. 178-181 (also available in Kuhn & Weinstock, 2002). Researchers using this scheme have identified the six major levels presented in Table 2 below (Leadbeater & Kuhn, 1989; Weinstock, 1999; Weinstock & Cronin, 2003). Levels 0-2 are regarded as an absolutist level, level 3 as multiplist, and levels 4 and 5 as evaluativist (Kuhn, 1991; Weinstock, 1999). (Weinstock (1999) identified two additional levels transitional between the absolutist and multiplist levels and between the multiplist and evaluativist levels.) In the present study, a participant was assigned a level from 0 to 5 on each of the 22 dimensions, and the dominant level was regarded as the participant's overall level. These levels are characterized in Table 2 in terms

of three major dimensions, the nature of the accounts, why they differ, and how claims are justified.

Concurrent construct validity of the coding scheme has been established by Kuhn et al. (2000), in a study in which participants assigned to one of the three main epistemological levels using a different instrument were independently assigned to the same epistemological level using the Livia Problem instrument. To establish inter-rater reliability of the Livia Problem coding scheme in the present study, two researchers independently coded 20% of the data. The percentage agreement across the 22 individual dimensions was 83%, and 96% for the overall level assigned. All disagreements were resolved through discussion. Appendix G provides a full description of how the coding scheme was applied and the inter-rater reliability process.

Table 2

Coding Scheme for Level of Epistemological Understanding

Epistemological Level	Nature of accounts	Why accounts differ	How claims are justified
0 Realist (Absolutist)	Both mirror reality objectively	Difference not recognized	None needed; accounts accurately report what happened
1 Predualist (Absolutist)	Each is objective, but limited by each historian's lack of access to all the facts	Differences attributed to incomplete, but not contradictory	Combination of accounts to produce whole story
2 Dualist (Absolutist)	One account objectively correct, the other is distorted due to error or bias	Accounts contradict each other; one right, the other wrong	Judgment of which account is correct
3 Multiplist	Subjective and idiosyncratic	Overriding subjectivity makes accounts wholly incommensurate	Assertion of opinion
4 Objective relativist (Evaluativist)	Constructions from evidence with advancing objectivity as the aim	Discrepancies attributed to different emphases on events and evidence	Evaluation of claim in relation to expert knowledge and evidence
5 Conceptual relativist (Evaluativist)	Constructions from evidence based in subjective context	Subjective context and perspective determine the interpretation of evidence and events	Evaluation of claim in relation to contexts of evidence, events, and knowledge claimant

Treatment Fidelity

To further ensure that there were no differences between the two groups on each of the variables from the demographic questionnaire, statistical analyses were conducted. Table 3 displays the variables and significance for each, which shows that there were no differences between the groups on any of the variables.

Table 3

Demographic Variables

Variable		Essay Group	Dialog Group	Test statistic	p-value
Age		M=23.3, SD=7.0	M=23.5, SD=5.7	t(58)= -.16	0.872
GPA		M=3.05, SD=.56	M=3.26, SD=.35	t(34)= -1.52	0.139
Psychology Majors		60%(18)	53%(16)	χ^2 (1, 60)= .27	0.602
Gender	Male	17%(5)	17%(5)		
	Female	83%(25)	83%(25)	χ^2 (1, 60)= .00	1.000
Debate/Argument Team	Yes	10%(3)	27%(8)		
	No	90%(27)	73%(22)	χ^2 (1, 60)= 2.78	0.095
Year of Study	Freshman	20%(6)	10%(3)		
	Sophomore	10%(3)	13%(4)		
	Junior	37%(11)	23%(7)		
	Senior	30%(9)	53%(16)		
	Non-matriculated	3%(1)	0%(0)	Fisher's	0.290
Race/Ethnicity	Black or African American	55%(16)	48%(14)		
	Hispanic	38%(11)	41%(12)		
	White	7%(2)	3%(1)		
	Other	0%(0)	7%(2)	Fisher's	0.675

Note. M = mean, SD = standard deviation, and numbers in parentheses represent number of subjects.

Effect of Prior Experience

Prior exposure to engaging in an argument curriculum of any kind was considered a control variable. In this sample, only 11 students reported to have engaged in some sort of educational activity involving argument or a debate team and 49 had not. To determine if there were differences in performance between the two groups, Mann-Whitney U Tests were conducted, as the sample size was small for those who had engaged in any form of an argument curriculum. Moreover, this test was conducted because the distribution across the mean proportion of argument skills for each evaluation was not normally distributed. A 2 x 3 Fisher's Exact Test was conducted to determine if the percentage of students in level of epistemological understanding differed by group. The results showed that the distributions on the writing task and percentages of epistemological levels did not differ across the two groups (all *p's* > .05). Therefore, this variable was not considered further.

Argumentive Writing Outcomes Across Conditions

Differences in evaluations produced within the dialog versus the essay. The first hypothesis was that argumentive skill would be superior for those who constructed a dialog compared to those who wrote an essay. It was first asked whether the two groups differed with respect to length of the two tasks. The Shapiro-Wilk's Test showed that normality was violated for number of idea units. Square root transformation was applied which solved the normality problem. Thus, an independent samples t-test was conducted on the transformed data to compare the mean number of idea units in the dialogs vs. essays. Participants did significantly differ in total idea units produced in the dialog and essay, $t(58) = -2.69$, $p = .009$, $d = -.71$, where the

dialog group produced more total idea units (M = 9.33, SD = 4.49) than the essay group (M = 6.67, SD = 2.96).

To answer the first hypothesis six, Generalized Linear Mixed Models (GLMM) using the binomial probability distribution were conducted. The evaluation types (Positive single evaluation, Critical single evaluation, Integrate own, Integrate other, Compare, and Integrate own/other) were the dependent variables and group (dialog vs. essay) was the independent variable. Two additional GLMM using the binomial probability distribution were conducted to determine if there was a significant relationship between group and Opponent Directed evaluations (sum of Critical single evaluation, Integrate other, Compare, and Integrate own/other) and Integrative Evaluations (sum of Integrate own, Integrate other, and Integrate own/other). The dependent variables were not normally distributed and correlated within each individual thus making the GLMM an appropriate statistical analysis. The binomial probability distribution analyzes the use of a given evaluation type to a reference category to determine if the odds of making a specific evaluation in comparison to a reference category differs by group. Non-evaluative statements were set as the reference category as these statements do not entail a specific argument skill, i.e. they are essentially non-substantive claims without linking an action to a problem.

Table 4 includes the fixed coefficients and random effects for all evaluation types. The significance value for the fixed coefficients indicates whether or not group effects the odds of a student using a given evaluation type relative to Non-evaluative in the dialogs and essays themselves. The purpose of including the random effects is to account for the correlation between evaluation types for each student. Thus, preventing the significance values of the fixed coefficients from being overestimated, which may risk Type 1 error rate inflation. When the

random effects themselves are significant, the variability between subjects is non-zero. The results revealed that there is a significant relationship between group and three of the evaluation types (Compare, Integrate own/other, and Opponent Directed). Specifically, relative to Non-evaluative statements, the odds of making a Compare are 5.80 times higher, Integrate own/other are 3.57 times higher, and Opponent Directed are 3.10 times higher for the dialog group compared to the essay group. Thus, students' dialogs compared to students' essays consist of more comparative statements aimed at addressing the opponent and also statements that integrate information.

Table 4

Fixed Coefficients and Random Effects for Each Evaluation Type Produced in the Dialog and Essay

							95% Confidence Interval	
Evaluation	Effect	Coefficient	Std. Error	t	Sig.	Odds ratio	Lower	Upper
	Intercept	0.566	0.292	1.934	.058		-0.021	1.152
	Dialog	-0.346	0.377	-0.919	0.364	0.71	-1.108	0.416
Positive single evaluation	Random effect variance	0.859	0.396		0.030		0.348	2.118
	Intercept	-1.336	0.423	-3.156	0.003		-2.188	-0.483
	Dialog	0.387	0.538	0.719	0.477	1.47	-0.703	1.477
Critical single evaluation	Random effect variance	1.164	0.641		0.069		0.395	3.423
	Intercept	-0.507	0.355	-1.430	0.160		-1.223	0.208
	Dialog	-0.082	0.485	-0.170	0.866	0.92	-1.062	0.897
Integrate own	Random effect variance	1.406	0.645		0.029		0.572	3.453
	Intercept	-2.237	0.540	-4.145	0.000		-3.316	-1.159
	Dialog	0.494	0.686	0.721	0.474	1.64	-0.884	1.872
Integrate other	Random effect variance	1.298	0.796		0.103		0.390	4.320

	Intercept	-2.191	0.586	-3.737	0.000		-3.355	-1.028
	Dialog	1.758	0.680	2.584	0.012	5.80	0.399	3.117
Compare	Random effect variance	1.857	0.863		0.031		0.747	4.618
	Intercept	-1.351	0.379	-3.563	0.001		-2.103	-0.599
	Dialog	1.273	0.447	2.849	0.006	3.57	0.380	2.166
Integrate own/other	Random effect variance	0.504	0.372		0.175		0.119	2.141
	Intercept	-0.276	0.334	-0.828	0.410		-0.942	0.389
	Dialog	1.130	0.410	2.755	0.008	3.10	0.306	1.953
Opponent Directed	Random effect variance	0.959	0.408		0.019		0.417	2.206
	Intercept	0.053	0.289	0.183	0.855		-0.525	0.630
	Dialog	0.490	0.382	1.282	0.206	1.63	-0.279	1.258
Integrative Evaluations	Random effect variance	0.897	0.396		0.024		0.378	2.133

Differences in evaluations produced in the TV Script. Next, analyses were conducted to determine whether the two groups differed with respect to length of the TV scripts. The Shapiro-Wilk's Test showed that normality was violated for number of idea units. Square root transformation was applied and did not solve the normality problem. Thus, a Mann-Whitney U test was conducted to compare the median number of idea units in the TV scripts of the dialog and essay groups. Although the dialog group produced more idea units in the TV script (*Mdn* =

3.00) than the essay group (*Mdn* = 2.50) the difference was not significant, U = 421, z = -.45, p = .657.

Generalized Linear Mixed Models using the binomial probability distribution were conducted for the same reasons explained above. The analyses examined if there was a significant relationship between group and the evaluation types produced in the TV scripts. Two additional GLMM using the binomial probability distribution were conducted to determine if there was a significant relationship between group and Opponent Directed evaluations and Integrative Evaluations. Non-evaluative statements were set as the reference category. There was only one instance of Integrate other, and it was combined with Integrate own/other.

Table 5 includes the fixed coefficients and random effects for all evaluation types. The significance value for the fixed coefficients indicates whether or not group effects the odds of a student using a given evaluation type relative to Non-evaluative in the TV scripts. Same as above, the purpose of including the random effects is to account for the correlation between evaluation types for each student. Thus, preventing the significance values of the fixed coefficients from being overestimated, which may risk Type 1 error rate inflation. When the random effects themselves are significant, the variability between subjects is non-zero. The results revealed a marginally significant relationship between group and Positive single evaluations. Specifically, relative to Non-evaluative statements, the odds of making a Positive single evaluation are 2.90 times higher for the dialog group compared to the essay group. All others were not significant. In sum, although only marginally significant, students in the dialog group make more own-sided statements in their TV scripts. The use of other evaluation types did not transfer to their TV scripts.

Table 5

Fixed Coefficients and Random Effects for Each Evaluation Type Produced in the TV Script

							95% Confidence Interval	
Evaluation	Effect	Coefficient	Std. Error	t	Sig.	Odds ratio	Lower	Upper
	Intercept	-0.883	0.465	-1.897	0.063		-1.814	0.048
	Dialog	1.065	0.565	1.885	0.069	2.90	-0.008	2.218
Positive single evaluation	Random effect variance	1.197	0.774		0.122		0.337	4.249
	Intercept	-2.643	0.808	-3.270	0.002		-4.262	-1.023
	Dialog	0.768	1.055	0.728	0.471	2.16	-1.371	2.907
Critical single evaluation	Random effect variance	2.689	1.610		0.095		0.832	8.695
	Intercept	-0.552	0.413	-1.338	0.187		-1.382	0.278
	Dialog	0.524	0.550	0.953	0.348	1.69	-0.596	1.644
Integrate own	Random effect variance	1.305	0.780		0.094		0.404	4.213
	Intercept	-2.570	0.601	-4.280	0.000		-3.806	-1.334
	Dialog	1.412	0.841	1.679	0.104	4.11	-0.310	3.134
Compare	Random effect variance	1.724	1.322		0.192		0.383	7.750

	Intercept	-1.086	0.472	-2.300	0.027		-2.041	-0.131
	Dialog	0.631	0.685	0.922	0.362	1.88	-0.753	2.016
Integrate own/other	Random effect variance	2.114	1.150		0.066		0.728	6.142
	Intercept	-0.768	0.473	-1.625	0.111		-1.719	0.183
	Dialog	1.015	0.643	1.579	0.122	2.76	-0.282	2.312
Opponent Directed	Random effect variance	2.282	1.037		0.028		0.936	5.562
	Intercept	-0.110	0.414	-0.265	0.792		-0.939	0.719
	Dialog	0.651	0.557	1.169	0.249	1.92	-0.472	1.774
Integrative Evaluations	Random effect variance	1.920	0.841		0.022		0.813	4.531

Individual differences in the use of each evaluation type in the dialog versus the essay. Individual differences were examined to determine if the percentage of participants using each evaluation type differed by group. Students were categorized into one of two categories; they either used the evaluation type or not in the dialog or essay. Results found that percentages did significantly differ for some evaluations produced in the constructed dialogs and argumentive essays. Table 6 contains the 2 x 2 Chi-square Tests of Independence results for each evaluation, which show that the dialog group more commonly used all evaluation types except for Positive single evaluations, where percentages were equal for both groups. The percentages of students that used Compare, Integrate own/other, Opponent Directed and Integrative Evaluations were significantly greater among the dialog group.

Table 6

Percentage of Participants Using Each Evaluation Type in the Dialog and Essay

Group	Evaluation Used	Evaluation Not Used	Test statistic	*p*	Effect Size
	Non-evaluative				
Essay	67%(20)	33%(10)			
Dialog	73%(22)	27%(8)	χ^2 (1, 60) = .32	.573	.07
	Positive single evaluation				
Essay	73%(22)	27%(8)			
Dialog	73%(22)	27%(8)	χ^2 (1, 60) = .00	1.000	.00
	Critical single evaluation				
Essay	23%(7)	77%(23)			
Dialog	40%(12)	60%(18)	χ^2 (1, 60) = 1.93	.165	.18
	Integrate own				
Essay	50%(15)	50%(15)			
Dialog	57%(17)	43%(13)	χ^2 (1, 60) = .27	.605	.07
	Integrate other				
Essay	13%(4)	87%(26)			
Dialog	23%(7)	77%(23)	χ^2 (1, 60) = 1.00	.317	.13
	Compare				
Essay	13%(4)	87%(26)			
Dialog	57%(17)	43%(13)	χ^2 (1, 60) = 12.38	.000	.45
	Integrate own/other				
Essay	30%(9)	70%(21)			
Dialog	83%(25)	17%(5)	χ^2 (1, 60) = 17.38	.000	.54

	Opponent Directed				
Essay	53%(16)	47%(14)			
Dialog	93%(28)	7%(2)	χ^2 (1, 60) = 12.27	.000	.45
	Integrative Evaluations				
Essay	73%(22)	27%(8)			
Dialog	93%(28)	7%(2)	χ^2 (1, 60) = 4.32	.038	.27

Note. Numbers in parentheses represent number of subjects. Effect size = Phi value

Individual differences in the use of each evaluation type in the TV script. Individual differences were examined to determine if the percentage of participants using each evaluation type produced in their TV scripts differed by group. Students were categorized into one of two categories; they either used the evaluation type or not in their TV script. Non-evaluative, Positive single evaluations, Integrate own, Integrate own/other, Opponent Directed and Integrative Evaluations were tested using 2 x 2 Chi-square Tests of Independence. Critical single evaluations, Integrate other and Compare were tested using 2 x 2 Fisher's Exact Tests as there were 2 cells that had an expected count less than 5. Table 7 contains the Chi-square and Fisher's Exact Test results. The percentage for the use of Non-evaluative was less for the dialog group, but the difference was not significant. Moreover, the percentage for the use of Positive single evaluations, Critical single evaluations, Integrate own, Integrate other, Compare, Opponent Directed and Integrative Evaluations was greater for the dialog group; however the difference was not significant. The percentage use of Integrate own/other was the same for each group.

Table 7

Percentage of Participants Using Each Evaluation Type in the TV Script

Group	Evaluation Used	Evaluation Not Used	Test statistic	*p*	Effect Size
	Non-evaluative				
Essay	53%(16)	47%(14)			
Dialog	43%(13)	57%(14)	χ^2 (1, 60) = .60	.438	-.10
	Positive single evaluation				
Essay	30%(9)	70%(21)			
Dialog	53%(16)	47%(14)	χ^2 (1, 60) = 3.36	.067	.24
	Critical single evaluation				
Essay	7%(2)	93%(28)			
Dialog	10%(3)	90%(27)	Fisher's	1.000	1.56
	Integrate own				
Essay	43%(13)	57%(17)			
Dialog	50%(15)	50%(15)	χ^2 (1, 60) = .26	.605	.07
	Integrate other				
Essay	0%(0)	100%(30)			
Dialog	3%(1)	97%(1)	Fisher's	1.000	1.03
	Compare				
Essay	10%(3)	90%(27)			
Dialog	17%(5)	83%(25)	Fisher's	.706	1.80
	Integrate own/other				
Essay	33%(10)	67%(20)			
Dialog	33%(10)	67%(20)	χ^2 (1, 60) = .00	1.000	.00

	Opponent Directed				
Essay	37%(11)	63%(19)			
Dialog	53%(16)	47%(14)	χ^2 (1, 60) = 1.68	.194	.17
	Integrative Evaluations				
Essay	57%(17)	43%(13)			
Dialog	70%(21)	30%(9)	χ^2 (1, 60) = 1.15	.284	.14

Note. Numbers in parentheses represent number of subjects. Effect size for Chi-square Tests of Independence = Phi value. Effect size for Fisher's Exact Tests = Odds Ratio value.

Epistemological Understanding and Argumentive Writing

Effect of constructing a dialog on epistemological understanding. Descriptive statistics showed that the number of absolutists, multiplists and evaluativists in the essay group was 8, 9, and 13, respectively and 0, 12, and 18, respectively in the dialog group. A 2 x 3 Fisher's Exact Test was conducted to determine if epistemological understanding was significantly affected by the dialogic task. The results revealed that epistemological levels differed significantly by essay vs. dialog group (p = .007), as depicted in figure 2. Twenty-seven percent of the essay group performed at the absolutist level, whereas no one in the dialog group did so. Only 30% of the essay group was multiplist, vs. 40% of the dialog group. Lastly, 43% of the essay group was evaluativist vs. 60% of the dialog group. This strongly suggests that epistemological understanding was sensitive to the effect of the dialogic task, leading some participants in the dialog group to exhibit a more advanced level than they would have had they not participated in the dialog task. This finding is suggestive but of course the effect cannot be proven as no pre-test assessment of epistemological understanding was given.

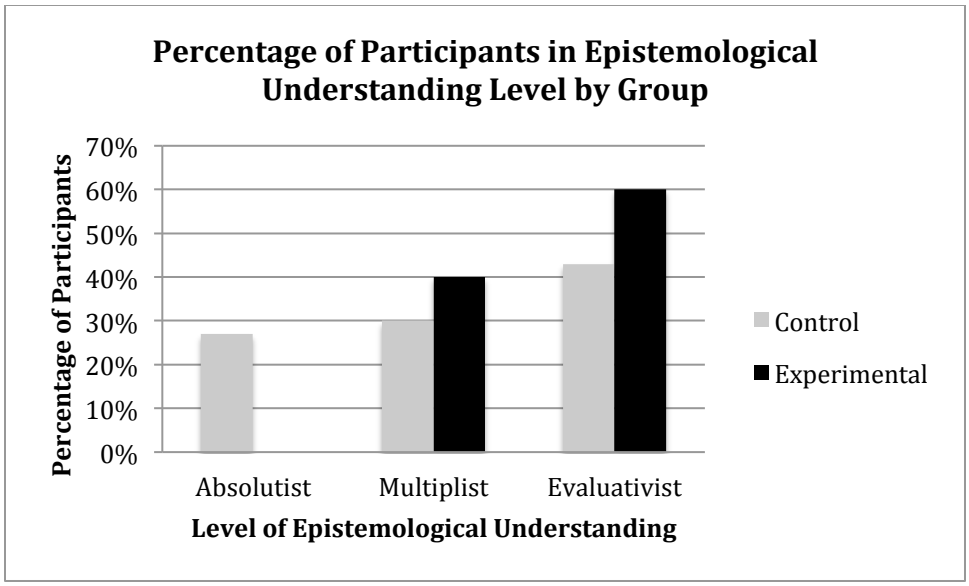


Figure 2. Percentage of participants in each epistemological understanding level by group.

Relationship between GPA and epistemological understanding. Research shows that educational level predicts epistemological understanding (Kuhn, 1992; Weinstock et al., 2006) and epistemological understanding predicts achievement (Muis & Duffy, 2013). A One-way ANOVA was conducted to test whether epistemological level was related to academic performance as measured by GPA. The mean reported GPA for the absolutists was the lowest (M=2.62, SD=.53), highest for the multiplists (M=3.27, SD=.40), and the mean GPA for the evaluativists was slightly less than the multiplists' (M=3.25, SD=.38). GPA did significantly differ by epistemological level, $F(2, 45) = 7.28$, $p = .002$, $\eta^2 = .24$. Tukey's HSD Post Hoc Tests revealed that GPA significantly differed between the absolutists and the multiplists (p = .003) and between the absolutists and the evaluativists (p = .002). GPA did not significantly differ between the multiplists and the evaluativists (p = .987). Figure 3 depicts the differences in mean GPA by level of epistemological understanding.

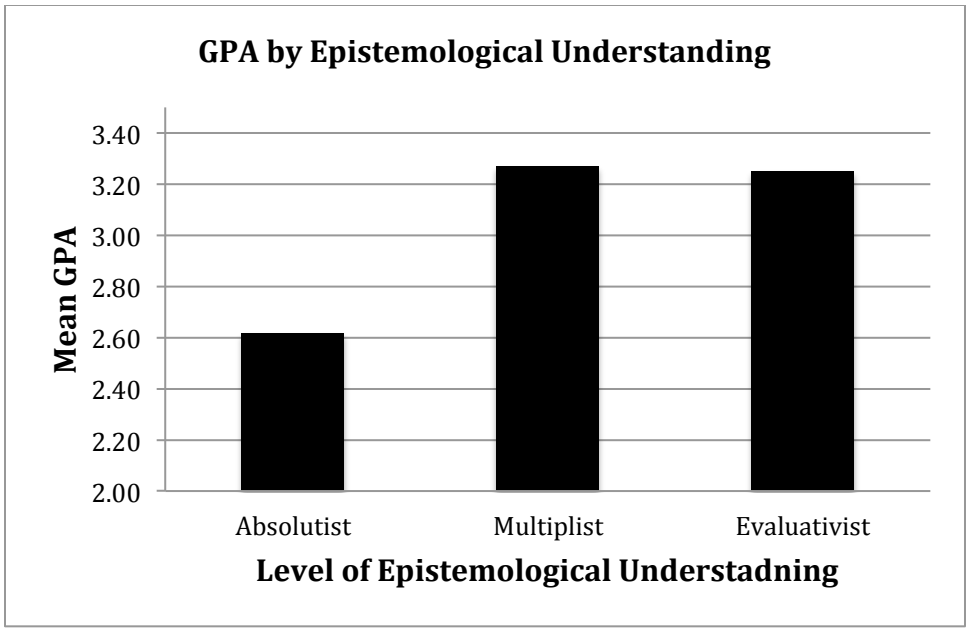


Figure 3. Differences in GPA across level of epistemological understanding.

Relationship between gender and epistemological understanding. Research shows that level of epistemological understanding does not differ by gender (Weinstock, 1999). To determine if this effect held in the present study a 2 x 3 Fisher's Exact Test was conducted. The results did not reveal a significant difference in epistemological level by gender ($p = .566$).

CHAPTER V:

Discussion

Purpose of Research

Two goals shaped the present research. The first was to determine if a dialogic writing task enhances argumentive writing skill. Typically, students predominately focus on support of their own position and do not consider and address opposing positions. Dialog with peers has been found to enhance students' argumentive writing with respect to addressing their own position in a framework of alternatives (Kuhn & Crowell, 2011; Reznitskaya et al., 2001). The present study asks whether a solitary dialogic writing task produces similar benefits.

A second goal was to determine if a dialogic writing task affected students' level of epistemological understanding. Specifically, whether the dialog task had an effect with respect to transition to an evaluativist level in which alternatives are recognized and compared by means of evaluative judgments.

Summary of Findings

Argumentive writing skills produced within the dialogs and essays were compared across the two groups - those who constructed a dialog and those who wrote an essay. The dialog group used significantly more statements that were integrative, comparative and addressed the opponent.

Argument skills produced in the TV script writing task were also compared across the two groups. Here, group differences largely disappeared, with both groups similarly focused on own-side arguments.

Comparisons were made to determine if the percentage of participants ever displaying a particular evaluation type differed between groups. A significantly greater percentage of participants in the dialog group used Compare evaluations (57% vs. 13%), Integrate own/other evaluations (83% vs. 30%), Opponent directed evaluations (93% vs. 53%) and Integrative Evaluations (93% vs. 73%). Thus, the writing of individuals in the dialog group was both more comparative and more integrative, relative to the essay group. In essence the dialog task elicited skilled argumentive writing that focused on addressing and weakening the opponent rather than only acknowledging the own-side position. The dialog task also elicited writing that to a greater extent integrated multiple pieces of information to support claims. Again, however, these differences disappeared in the case of the TV script.

The next goal was to examine the effect of the dialogic writing task on students' level of epistemological understanding. The results showed that epistemological levels differed significantly as a function of essay versus dialogic writing experience—60% of those who constructed a dialog were assessed to be at the evaluativist level after the writing task, versus 43% of those who wrote an essay. This suggests that the dialog activity scaffolded students in the direction of an evaluativist understanding involving evaluation of claims in a framework of alternatives.

Overall, then, constructing a dialog was beneficial in encouraging writing that addressed the opponent and that integrated multiple pieces of information. It was also beneficial in enhancing epistemological understanding toward the evaluativist level.

Limitations and Future Directions

Transfer of argumentive skill. Argument skills used in the dialogs did not transfer to the TV scripts. This lack of transfer suggests that students may have lacked a strong argument

schema. Such transfer has been claimed to be indicative of an argument schema having been developed (Reznitskaya et al., 2001). Argument schemas allow students to understand and appreciate the components of a skilled argument, such as acknowledging and weakening the opponent and integrating information to support or weaken a claim. An argument schema allows students to retrieve information from such a schema and to incorporate it in their writing. Reznitskaya et al. (2001) found that the argument schema students developed when engaging in peer-to-peer discourse allowed for the transfer of argument skills used during discourse to their writing.

The amount of time students had in the present study to construct a dialog may have been too short, as it requires practice for argument skills to develop (Kuhn et al., 2013). The interventions that fostered the transfer of argument skills used during peer-to-peer discourse to argumentive writing spanned from 5-weeks (Reznitskaya et al., 2001) to 3-years (Kuhn & Crowell, 2011). Students had plenty of time to develop the argument skills they applied to their writing. However, in the present study students were asked to construct a dialog just one time, which did not result in transfer of argument skills to their TV scripts. For adult students to develop more generalized argument skills, perhaps extending the construction of a dialog task over a short period of time (e.g., every day for one week) would be beneficial. Adults do not have as much time to engage in an extensive curriculum as younger students do. Also, adults demonstrate argument skills superior to those shown by adolescents (Felton & Kuhn, 2001) indicating they may possess a more developed argument schema than adolescents and may not need as much time to develop argument skills as adolescents do. Therefore, a short intervention may be sufficient for adult students.

In sum, giving students more time to practice the skills and develop an argument schema may result in students' transfer of those skills to the TV script.

Change in epistemological understanding. Only the effect of group (dialog vs. essay) on epistemological understanding was directly assessed; change over time was not. The results showed that the percentage of participants in level of epistemological understanding did significantly differ by group. Sixty percent of those who constructed the dialog vs. 43% who wrote an essay were assessed as at the evaluativist level, 40% vs. 30% at the multiplist level and zero vs. 27% at the absolutist level. These differences suggest that the dialog experience supported a transition to a more mature level of epistemological understanding. Specifically, those who constructed the dialog and were at the absolutist level o f epistemological understanding prior to the writing task may have progressed from this level in the direction of a multiplist or even evaluativist level. However, epistemological understanding was not assessed prior to the writing task, therefore, this conclusion can be no more than tentative, until further research is conducted in which epistemological understanding is assessed both prior to and after the writing task.

If change does occur, it would be important to determine when it happens and the sustainability of that change. If genuine change did occur in the present study, it happened rather quickly, as students engaged in the dialogic writing task just once. However, research results vary with respect to the timing and speed of change. Some have claimed that it occurs quickly (Ferguson & Braten, 2013; Kienhues, Bromme, & Stahl, 2008; Kienhues, Stadtler, & Bromme, 2011; Porsch & Bromme, 2011) and others over long periods of time (Muis & Duffy, 2013). Muis & Duffy (2013) found it to occur weeks after the intervention. Their participants were graduate students enrolled in a statistics course. In a control class of 32, instruction was lecture

based. In an intervention class of 31, students were also given a lecture, but in addition required to engage in inquiry and dialog about the content. The course spanned 12 weeks, and students' epistemological understanding with respect to statistics was assessed five times through the intervention and three weeks post intervention. The results showed that epistemological understanding changed as early as week 4 from a naïve level to a mature level for the students who were enrolled in the intervention class, whereas this effect did not occur for the students who were enrolled in the control group class. Furthermore, change in epistemological understanding was sustained 3-weeks beyond the intervention.

Others have identified change in students' epistemological understanding after they were exposed to opposing views in short-term interventions. Kienhues, Bromme, and Stahl (2008) conducted a study of 58 students who held either a mature or naïve epistemological understanding. They were randomly assigned to read a text consisting of opposing views (intervention group) or consistent views (comparison group) on DNA fingerprinting. Those who were in the intervention group and had a naïve epistemological understanding shifted to a more mature epistemological understanding post intervention. Those who were in the comparison group and had a mature epistemological understanding shifted to a more naïve level post intervention. In another study by Kienhues, Staftler, & Bromme (2011) 100 undergraduate students were recruited and randomly assigned to one of two intervention groups (where students either read medical information on the Web that was conflicting or consistent) or a control group (where students did not read information from the Web). Students in the intervention group were given 30 minutes to review the Web content. Fewer participants who studied conflicting information than consistent information believed in a clear solution regarding the medical problem they were reading about. Moreover, those in the intervention groups, compared to the

control group, showed a move in epistemological understanding toward a more mature level. Porsch and Bromme (2011) further confirmed the studies by Kienhues and colleagues and found that change occurred after a 30-minute intervention where students either read a conflicting or consistent text about a scientific topic. Ferguson and Braten (2013) conducted a study with 65 tenth graders in which everyone engaged in a 60-minute intervention in which they read five different views on sun exposure and health. Those who had low knowledge and naïve epistemological beliefs about the topic gained more knowledge and relied less on personal justification post intervention. However, 21% based their knowledge on authority (reflecting the absolutist level) and 29% based their knowledge on multiple views (reflecting a multiplist or evaluativist level). Those with moderate knowledge but low personal justification gained knowledge and maintained their epistemological belief. However, 50% based their knowledge on authority and 42% on multiple views.

These shifts indicate that short-term interventions can be effective in demonstrating that conflicting information leads to doubt, which creates cognitive conflict and thus may lead students to change their views. However, the sustainability of this change was not measured in these short-term interventions (except for Porsch & Bromme, 2011, where the effect of the change was still evident four weeks after the intervention).

Based on these short-term interventions, future research may be successful in confirming change in epistemological understanding as a function of the dialogic task employed in the present work. What also needs to be measured is how long that change lasts. Thus, a follow-up study is indicated in order to assess the sustainability of change observed in students' epistemological understanding.

Another limitation of the current study is that there was no true control group to compare the effect of the dialog versus essay on epistemological understanding. Another group who does no writing or only writes the TV script is needed. This addition would provide information on whether the essay has an influence on epistemological understanding.

Effect of epistemological understanding on argumentive writing. As explained earlier in the introduction and literature review, research has shown that epistemological understanding is related to argument skill, with those who are at the evaluativist level demonstrating better command and usage of argument skills (Kuhn, 1991; Kuhn & Park, 2005; Kuhn, Wang, & Li., 2011; Mason & Boscolo, 2004; Mason & Scirica, 2006; Weinstock, 2010; Weinstock and Cronin, 2003). This effect could not be analyzed in the present study because epistemological understanding was not assessed prior to the writing task. Again, it would be necessary in future research to assess epistemological understanding prior to the writing tasks to determine if the writing tasks were truly the causal factor.

Implications and Conclusions

Research suggests that students who engage in interventions with a large emphasis on peer-to-peer discourse benefit in many ways. They gain skills necessary to become a successful arguer (Kuhn & Crowell, 2014; Kuhn et al., 2008; Kuhn & Udell, 2003; Kuhn et al., 2013; Moore, 2013; Shaenfield, 2009), they gain the ability to write better argumentive essays (Khait, 2014; Kuhn & Crowell, 2011; Reznitskaya et al., 2001), they move in the direction of higher levels of epistemological understanding, and they develop a better understanding of topic specific content (Muis & Duffy, 2013).

Peer-to-peer dialog may not always be feasible in a classroom. Many teachers may feel they do not have the time to set aside for students to engage in peer discourse. Also, adult

students may not have the time to engage in long-term interventions that are not part of a course curriculum. Therefore, the so-called "solitary" dialog examined in the present study may be an alternative. Teachers can have students of many ages construct dialogs about many kinds of issues. These are skills they can then apply outside of the classroom as well -- for example when writing papers, studying, or wishing to explore an idea further and do not have a peer to engage with.

The benefits of dialog are broad and great, and the present study shows that so-called solitary dialog has the capacity to enhance argumentive writing and epistemological understanding. Specifically, college students who constructed a dialog produced more skilled argumentive writing. They made more comparative, opponent directed and integrative evaluations. Dialog construction also appeared to enhance students' epistemological understanding – a crucial achievement given more advanced epistemological understanding is related to argumentation skill (Kuhn, 1991; Kuhn & Park, 2005; Kuhn, Wang, & Li., 2011; Mason & Boscolo, 2004; Mason & Scirica, 2006; Weinstock, 2010; Weinstock and Cronin, 2003). The present work suggests the possibility of a bidirectional effect, that is one in which dialog enhances students' epistemological understanding, which in turn may further positively impacts students' argument skills produced in dialogs and perhaps other writing tasks as well.

Although college students participated in the present study, establishing that dialog can be effective in enhancing both argument skills and epistemological understanding makes it likely that primary and secondary school students could benefit from a dialogic writing task like the one used here as well. Since the Common Core (2010) has developed standards for students' argumentive writing, it would be worth researching whether this effect occurs with younger students who have not engaged in any argument intervention. If the tool employed here is just as

effective with younger students as it is with adults, it suggests a valuable means to better prepare them for the demands of college, the workplace, and citizenship.

References

Alvermann, D. E., Hynd, C. E. , & Qian, G. (1995). Effects of Interactive Discussion and Text Type on Learning Counterintuitive Science Concepts. *The Journal of Educational Research, 88*(3), 146-154.

Assessing 21st Century Skills: Summary of a Workshop. (2011). The National Academies Press.

Baron, J. (1988). Thinking and deciding. New York: Cambridge University Press.

Barzilai, S. & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36*, 86-103.

Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science, 43*, 737-766.

Benesch, S. (1999). Thinking critically, thinking dialogically. *Teachers of English to Speakers of Other Languages, Inc. (TESOL), 33*(3), 573-580.

Brookfield, S. D. (2013). *Powerful techniques for teaching adults.* San Francisco: Jossey-Bass.

Byrnes, J. P. & Dunbar, K. N. (2014). The nature and development of critical-analytic thinking. *Educational Psychology Review*, *26*(4), 477-493.

Chinn, C.A., & Brewer, W. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.

Chinn, C. & Brewer, W. F. (1996). Mental models in data interpretation. *Philosophy of Science, 63*, S211-S219.

Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching, 35*, 623-654.

CollegeBoard. (2015). Writing section. Retrieved from https://professionals.collegeboard.com/testing/sat-reasoning/about/sections/writing

Common Core State Standards Initiative. (2010). Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects. Retrieved from http://ww.corestandards.org/assets/CCSSI_ELA standards.pdf

Cooper, C. R., Cherry, R., Copley, B., Fleischer, S., Pollard, R., & Sartisky, M. (1984). Studying the writing abilities of a university freshman class: Strategies from a case study. In R. Beach & L. S. Bridwell (Eds.), New directions in composition research (pp. 19–52). New York: Guilford Press.

Crowell, Amanda & Kuhn, D. (2014). Developing dialogic argumentation skills: A three-year intervention study.

Educational Testing Service. (2015). Preparing for the argument task. Retrieved from https://www.ets.org/gre/revised_general/prepare/analytical_writing/argument/prepare_task

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron & R. Sternberg (Eds.), Teaching for thinking (pp. 9–26). New York: Freeman.

Ennis, R. H. (1995). *Critical thinking*. New York: Prentice Hall.

Felton, M. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development, 19*(1), 35-52.

Felton, M. & Kuhn, D. (2001). The development of argumentive discourse skills. *Discourse Process, 32*(2&3), 132-153.

Ferguson, L. & Braten, I. (2013). Students profiles of knowledge and epistemic beliefs: Changes and relations to multiple-text comprehension. *Learning and Instruction, 25*, 49-61.

Ferretti, R., MacArthur, C., & Dowdy, N. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92*(4), 694-702.

Graff, G. (2003). *Clueless in academe: How schooling obscures the life of the mind.* New Haven: Yale University Press.

Hays, J. N., & Brandt, K. S. (1992). Socio-cognitive development and student performance on audience-centered argumentative writing. In M. Secor & D. Charney (Eds.), Constructing rhetorical education (pp. 202–229). Carbondale: Southern Illinois University Press.

Hileman, T. (2006, October). *Democracy and our civic responsibilities.* Address to the New York Society for Ethical Culture, New York, NY.

Hofer, B. K. (2000). On dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*, 378–405.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*, 88-140.

Iordanou, K. (2012). Developing face-to-face argumentation skills: Does arguing on the computer help? *Journal of Cognition and Development, 14*(2), 292-320.

Khait, V. (2014). *Making use of the dual functions of evidence in adolecents' argumentation*. Unpublished doctoral dissertation, Teachers College Columbia University, New York.

Kienhues, D., Bromme, R., & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology, 78,* 545-565.

Kienhues, D., Stadtler, M., & Bromme, R. (2011). Dealing with conflicting or consistent medical information on the web: When expert information breeds laypersons' doubts about experts. *Learning and Instruction, 21,* 193-204.

Klaczynski, P.A., & Gordon, D.H. (1996). Self-serving influences on adolescents' evaluations of belief-relevant evidence. *Journal of Experimental Child Psychology, 62*(3), 317-339.

Knight, F. H. (1959). Sources Of Responsibility In Democracy. *American Behavioral Scientist*, *2*(4), 10–13.

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

Kuhn, D. (1992). Thinking as Argument. *Harvard Educational Review*, *62*(2), 155–178.

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

Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of childhood cognitive development.* Oxford: Blackwell. (2nd ed.)

Kuhn, D., Cheney, R. & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*, 309-328.

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, *22*(4), 545–52.

Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the Computer : A Microgenetic Study of Developing Argument Skills in a Computer-Supported Environment, *79*(5), 1310–1328.

Kuhn, D. Hemberger, L., & Khait, V. (2014). *Argue with me: Argument as a path to developing students' thinking and writing.* Bronxville, NY: Wessex Inc.

Kuhn, D., Hemberger, L., & Khait, V. (2016). *Argue with me: Argument as a path to developing students' thinking and writing.* New York: Routledge. (2nd ed.)

Kuhn, D., Iordanou, K., Pease, M., & Wirkala, C. (2008). Beyond control variables: What needs to develop to achieve skilled scientific thinking? *Cognitive Development, 23,* 435-451.

Kuhn, D. & Park, S. (2005). Epistemological understanding and the development of intellectual values. *International Journal of Education Research, 43*, 111-124.

Kuhn, D., Pennington, N., & Leadbeater, B. (1983). Adult thinking in developmental perspective. In *Lifespan Development and Behavior*, Baltes, P., Brim, O. (eds). Academic Press: New York; 157–195.

Kuhn, D., & Udell, W. (2003). The Development of Argument Skills, *74*(5), 1245–1260.

Kuhn, D., Wang, Y. & Li, H. (2011). Why argue? Developing understanding of the purpose and values of argumentive disocurse. *Discourse Processes, 48*, 26-49.

Kuhn, D. & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In *Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing*, Hofer, B., Pintrich, P. (eds). Lawrence Erlbaum: New Jersey; 121–144.

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.

Larson, A.A., Britt M.A., & Kurby, C.A. (2009). Improving students' evaluation of informal arguments. *The Journal of Experimental Education, 77*(4), 339.

Leadbeater, B. & Kuhn, D. (1989). Interpreting discrepant narratives: Hermeneutics and adult cognition. In *Everyday Problem Solving*, Sinnott, J. (ed.). Praeger: New York; 175–190.

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

Macago, F. & Konstantini, A. (2013). What students' arguments can tell us: Using argumentation schemes in science education. *Argumentation, 27*, 225-243.

Mason, L. & Boscolo, P. (2004). Role of epistemological understanding and interest in interpreting a controversy and in topic-specific belief change. *Contemporary Educational Psychology, 29*, 103-128.

Mason, L. & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction, 16*, 492-509.

Mercier, H. & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *The Behavioral and Brain Sciences*, *34*(2), 57–74; discussion 74–111.

Moore, W. (2013). *The Use of Evidence in Young Adolescents' Argumentation.* Unpublished doctoral dissertation, Teachers College Columbia University, New York.

Moshman, D. (2014). *Epistemic cognition and development: The psychology of justification and truth*. New York: Psychology Press.

Muis, K. & Duffy, M. (2013). Epistemic climate and epistemic change: Instruction designed to change students' beliefs and learning strategies and improve achievement. *Journal of Educational Psychology, 105*, 213-225.

National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2011* (NCES 2012–457). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

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

Nussbaum, M. & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education, 76*(1), 59-92.

O'Keefe, D. J. (1999). How to handle opposing arguments in persuasive messages: A meta-analytic review of the effects of one-sided and two-sided messages. In M. E. Roloff (Ed.), Communication yearbook (Vol. 22, pp. 209-249). Thousand Oaks, CA: Sage.

Osborne, J. F. (2009, December-2010, January). An Argument for Arguments in Science Classes. *The Phi Delta Kappan, 91*(4), 62-65.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), Informal reasoning and education (pp. 83–106). Hillsdale, NJ: Erlbaum.

Porsch, T. & Bromme, R. (2011). Effects of epistemological sensitization on source choices. *Instructional Science, 39,* 805-819.

Reznitskaya, A., Anderson, R., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of Oral Discussion on Written Argument. *Discourse Processes*, *32*(2), 155–175.

Reznitskaya, A., Glina, M., Carolan, B., Michaud, O., Rogers, J., & Sequeira, L. (2012). Examining transfer effects from dialogic discussions to new tasks and contexts. *Contemporary Educational Psychology*, *37*(4), 288–306.

Santos, C. M. M., & Santos, S. L. (1999). Good argument, content and contextual dimensions. In G. Rijlaarsdam & E. Esperet (Series Eds.) & J. Andriessen & P. Coirier (Vol. Eds.), Studies in writing: Vol. 5. Foundations of argumentative text processing (pp. 75–95). Amsterdam: Amsterdam University Press.

Scheurman, G. (2006, July). America as metaphor: Using argument to teach about American identity. *Magazine of History, 20*(4), 27-35.

Scriven, M. (1976). *Reasoning.* New York: McGraw-Hill.

Shaenfield, D. (2009). *The role of meta-level regulation in developing argumentive discourse skills.* Unpublished doctoral dissertation, Teachers College Columbia University, New York.

Stahl, E., & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, *17*, 773-785.

Toulmin, S. (2003). *The uses of argument* (updated ed.). Cambridge, United Kingdom: Cambridge University Press.

Walton, D. (1988). Burden of proof. *Argumentation, 2*, 233-254.

Walton, D. N. (1989). Dialogue theory for critical thinking. *Argumentation*, *3*(2), 169–184.

Walton, D. N. (1996). *Arguments from ignorance.* University Park, PA: The Pensylvania State University Press.

Walton, D. N. (2013). *Methods of argumentation.* Cambridge, MA: Cambridge University Press.

Weinstock, M. (1999). Epistemological understanding and argumentive competence as foundations of juror reasoning skill. Doctoral dissertation, Columbia University, *Dissertation Abstracts International,* 60: AAT 9916928.

Weinstock, M. (2006). Psychological research and epistemological approach to argumentation. *Informal Logic, 26*(1), 103-120.

Weinstock, M. (2010). Epistemically understanding and sound reasoning skills that underlie effective democratic engagement. *Journal of Peace Education and Social Justice, 4*(1), 56-77.

Weinstock, M. & Cronin, M. (2003). The everyday production of knowledge: Individual differences in epistemological understanding and juror-reasoning skill. *Applied Cognitive Psychology, 17*, 161-181.

Weinstock, M., Neuman, Y. & Glassner, A. (2006). Identification of informal reasoning fallacies as a function of epistemological level, grade level, and cognitive ability. *Journal of Educational Psychology, 89*(2), 327-341.

Wolfe, C. (2012). Individual differneces in the "Myside Bias" in reasoning and written argumentation. *Written Argumentation, 29*(4), 477-501.

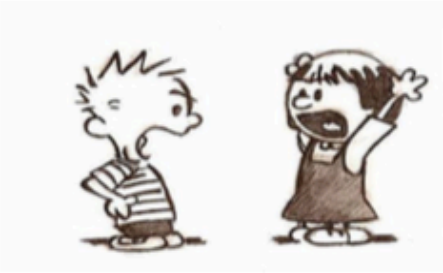
Wolfe, C. & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking and Reasoning, 14*(1), 1-27.

Wolfe, C., Britt, M. A., & Butler, J. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication, 26*(2), 183-209.

Zavala, J. (unpublished). *Influence of constructed dialog on argumentive writing.* Unpublished doctoral empirical, Teachers College Columbia University, New York.

Appendix A: Recruitment Flier

Argument Research Study



Argument is deeply rooted in politics and I would like to invite you to give your opinion on a couple of political issues. This research could contribute to the improvement of argumentive writing in educational settings.

Who is eligible?

- **Undergraduate students attending Mercy College**

What will you be asked to do?

- **Write your opinion on a political issue**
- **Listen to a story and answer a few questions**
- **Evaluate a short argument**

Compensation

- **A $10 Dunkin Donuts gift card will be given for your time (about 1 hour).**

If you are interested or have any questions, please contact me via email:
Julia Zavala, M.A.
Adjunct Instructor, Department of Social and Behavioral Sciences
Mercy College
jzavala1@mercymavericks.edu

This study has been approved by the Mercy College IRB to use this flyer for recruiting participants.

Teachers College, Columbia University
Institutional Review Board
Protocol Number: 15-061
Consent Form Approved Until: 12/02/2015
Signature:

Appendix B: In-person Recruitment Script

Hi. My name is Julia Zavala. I am an instructor here at Mercy College and a doctoral student at Teachers College Columbia University. I am conducting a study to fulfill the doctoral requirements for my dissertation. I would like to get your opinion on a couple of political issues. I am looking for participants and would be very grateful if you would be willing to take part in my study. A $10 Dunkin Donuts gift card will be given for your time (about an hour). Do you have any further questions? If you are interested, please send me an email and we can set up an appointment. Here is a flyer with my contact information on it. Thanks for your time.

Appendix C: Consent Form

Participant #_____

Julia Zavala
Social and Behavioral Sciences
Mercy College
555 Broadway
Dobbs Ferry, NY 10552

Informed Consent for Participants in Research

Project Title: Argumentive Writing

Dear Potential Participant:

You are being asked to participate in a study that will examine college students' writing. You will be given a story about two mayors running for office in a troubled city. There will be a list of the problems in the city and how each candidate will address the problems. You will write a short script in defense of who you think is the better candidate. You will be read short accounts of a war and answer questions in regard to it. Answers will be audio recorded. Afterward, you will evaluate an argument. The entire study will take approximately 60 minutes to complete. This study is being done by Julia Zavala, an instructor at Mercy College, who is completing this project as part of a doctorate degree at Teachers College, Columbia University in the Human Development Department.

The tasks that will be given to participants are of minimal risk and will not exceed what is encountered in everyday life.

To minimize the risk of a breach of confidentiality, all of the study information will be kept private by being locked in a file cabinet or on a secure laptop computer in the office of Julia Zavala at Teacher's College, Columbia University. Also, all tasks will only include a participant number and no names. Confidentiality will be maintained, as only the principal investigator and doctorate advisor, Dr. Deanna Kuhn, will have access to the raw data for coding and analysis. Audio recordings will be permanently deleted a year and a half after completion of data collection. You will not be identified in any publication or presentation of the study findings. Only group scores will be reported.

Your decision whether or not to participate will not affect your current or future relations with Mercy College. If you decide to participate, you are free to withdraw at any time without affecting this relationship. You may also choose not to complete the writing task or answer the questions. The researcher also may choose to withdraw you from this study if it is in your best interest.

You are encouraged to ask any questions regarding the research to Julia Zavala via email at jzavala1@mercymavericks.edu or 917-544-4886. The Institutional Review Board (IRB) of Mercy College has approved recruitment of participants for this research study. The IRB chairperson, Brian Baker, can be contacted to answer questions about the rights of research participants via email at mcirb@mercy.edu or 718-678-8812.

Teachers College, Columbia University
Institutional Review Board
Protocol Number: 15-061
Consent Form Approved Until: 12/02/2015
Signature: SBH

Participant #_____

If you wish to find out the results of this study you may contact me at about three months after participation.

SIGNED AGREEMENT

I understand that:

(a) My signature indicates that I voluntarily agree to be a part of this research study.
(b) Any questions that I have regarding this research study have been fully answered.
(c) By signing this form I do not waive any legal rights.
(d) I will receive a copy of this form.

Signature of Subject Date

I believe that the subject fully understands my explanation and has freely given informed consent.

Signature of Investigator Date

Teachers College, Columbia University
Institutional Review Board
Protocol Number: 15-061
Consent Form Approved Until: 12/02/2015
Signature:

Appendix D: Demographic Questionnaire

1. How old are you? _____

2. Are you a
    a. Freshman?
    b. Sophomore?
    c. Junior?
    d. Senior?

3. What program are you majoring in? _____

4. What is your current GPA? _____

5. What is your gender?
    a. Male
    b. Female

6. Which one category of the following best applies to you?
    a. Asian (includes East Asian, South Asian, Middle Eastern)
    b. Black or African American
    c. Hispanic or Latino
    d. Native American
    e. Native Hawaiian or other Pacific Islander
    f. White
    g. Other

7. Have you ever been part of a debate or argumentation team?
    a. Yes
    b. No

8. If you answered yes to question 7, please explain how you and your team prepared for debates.

Appendix E: Writing Task Coding Process

Step 1:

Students' hand-written dialogs, essays, and TV scripts were transcribed in a Word document. Here we will use the TV script from Subject 25 as an example.

Subject 25:

I'm Olivia and I think that the best candidate for mayor is Cruz. Cruz will address many problems that our city is facing; specifically she plans to set a teen curfew, which Diaz does not even mention in her plan. We are in an economic recession right now so it's not the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay. Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program. Overall, Cruz will be a better mayor because she addresses the problems.

Step 2:

Each coder segmented students' essays into idea units using forward slashes in red.

Subject 25:

I'm Olivia and I think that the best candidate for mayor is Cruz. /Cruz will address many problems that our city is facing; / specifically she plans to set a teen curfew, which Diaz does not even mention in her plan. /We are in an economic recession right now so it's not the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay. /Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program. /Overall, Cruz will be a better mayor because she addresses the problems.

Step 3:

Reliability for idea units was calculated by coder one (JZ). Coder one was used as the baseline, therefore, the number of idea units from the second coder that were in agreement with the first coder's were divided into the first coder's total amount of idea units. In the example below, both the first and second coder identified the same 6 idea units. Therefore, there was 100% agreement in identifying idea units→ 6/6=1.

Coder One's Segments:

Subject 25:

I'm Olivia and I think that the best candidate for mayor is Cruz. /Cruz will address many problems that our city is facing; / specifically she plans to set a teen curfew, which Diaz does not even mention in her plan. /We are in an economic recession right now so it's not

the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay. /Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program. /Overall, Cruz will be a better mayor because she addresses the problems.

Coder Two's segments:

Subject 25:

I'm Olivia and I think that the best candidate for mayor is Cruz. /Cruz will address many problems that our city is facing; / specifically she plans to set a teen curfew, which Diaz does not even mention in her plan. /We are in an economic recession right now so it's not the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay. /Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program. /Overall, Cruz will be a better mayor because she addresses the problems.

Step 4:

The first coder copied and pasted students' idea units into a table to code. This table was subsequently given to the second coder to code.

Segment #	Statement	Substantive Code
1	I'm Olivia and I think that the best candidate for mayor is Cruz.	
2	Cruz will address many problems that our city is facing;	
3	specifically she plans to set a teen curfew, which Diaz does not even mention in her plan.	
4	We are in an economic recession right now so it's not the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay.	
5	Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program.	
6	Overall, Cruz will be a better mayor because she addresses the problems.	

Step 5:

After both coders finished, reliability was established using the first coder as a baseline. The number of codes from the second coder that were in agreement with the first coder's was divided into the first coder's total amount of codes. In the example below, both the first and second

coder coded all 6 idea units with the same code. Therefore, there was 100% agreement in coding→ 6/6=1.

Coder One's Codes:

Segment #	Statement	Substantive Code
1	I'm Olivia and I think that the best candidate for mayor is Cruz.	NO CODE
2	Cruz will address many problems that our city is facing;	INTEGR-OWN
3	specifically she plans to set a teen curfew, which Diaz does not even mention in her plan.	COMPARE
4	We are in an economic recession right now so it's not the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay.	INTEGR-OWN/OTH
5	Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program.	COMPARE
6	Overall, Cruz will be a better mayor because she addresses the problems.	NO CODE-RECAP

Coder Two's Codes:

Segment #	Statement	Substantive Code
1	I'm Olivia and I think that the best candidate for mayor is Cruz.	"setting up"
2	Cruz will address many problems that our city is facing;	Integrate Own
3	specifically she plans to set a teen curfew, which Diaz does not even mention in her plan.	Compare
4	We are in an economic recession right now so it's not the time to be building a new athletic stadium. The money could go to something much more important like raising teachers' pay.	Integrate Own/Other
5	Diaz does not have any plans for unemployment problems while Cruz proposes to create a job training program.	Compare
6	Overall, Cruz will be a better mayor because she addresses the problems.	"recap"

Appendix F: Livia Problem Coding Scheme

Epistemological Understanding Coding Scheme (Adapted from Weinstock (1999), p. 178-181)

1. What is subject's core explanation of what differs?
   __No difference (0)
   __Difference noted (1-5)
      __Content of accounts of events differ (1,2)
      __Historians' perceptions of events differ (3-5)
         __Historians have different opinions (3)
         __Conditions under which historians conclude what happened differ (4,5)
         __Emphasis on different facts (4,5)

2. Structural reasons for difference
   __No difference (0)
   __Different content available to different people (1,2)
   __Logical contradiction, only one set of facts or conclusions can be right (2)
   __Different people—each has own ideas (2,3)
   __Different context in which historians operate, different perspectives on content (4,5)

3. Both right
   __Yes (0,1,3,5)
      __No real differences (0,1)
      __Both equally right or wrong (may read: No, both biased and wrong) (3)
      __Yes, from own perspective (3,5)
   __No (2,3,4)
      __One closer to truth (2,4)

4. One account more true?
   __No (0,1,3)
      __They are both accurate reports of the events (0,1)
      __They are both equally biased, or there is no way to judge (3)
   __Yes (2,4,5)
      __One account is a better fit and/or more probable (2,4,5)
   __Maybe, not without overall perspective, but interpretation still must fit evidence (5)

5. Reconcilability
   __Stories/facts not contradictory (0,1)
   __Stories/some facts are contradictory (2-5)
      __one right, one wrong (2)
      __stories, the framework for facts, are basically incompatible (3,5)
      __some parts of the story/facts can be agreed upon or brought into agreement(4,5)

6. # of accounts
   __One (0,1)
   __Two (2)
   __Three (with the 3rd Historian) (2,3)
   __Multiple or indeterminate number more than 2 (3-5)

7. Explanation of how two accounts emerge
   __Both people report what happened as far as the subject knows (0)
   __One person collected some facts; the other collected other facts—some overlap (1)
   __They use different facts to tell the story with no judgment about facts (1,2)
   __One person had access to correct facts; other in error or bias distorts one story (2)
   __Each tells story based on biased opinion or perception (3)
   __Each has different access to facts, different facts emphasized due to perspective (4)
   __Each approaches question of what happened from different subjective perspective (5)

8. Certainty
   __Yes, with access to facts (0-2)
   __No (2-5)
      __Not from the given accounts (2,3)
      __All views equal without possibility of adjudication (3)
      __That is the condition of knowing—but degree of confidence may be possible (4,5)

9. Determining factor in account
   __What happened (0,1,2)
   __Belief, or biased perspective of teller (2,3)
   __Both evidence and perspective (4,5)

10. Account as:
   __Report of what happened (0-2)
   __Opinion of teller (3)
   __Construction from evidence from a particular perspective (4,5)

11. Bias
   __Not considered (0,1)
   __Responsible for distortion, or way evidence reported, by at least one historian (2,3)
   __One possible factor in construction of account that may affect interpretation (4,5)

12. Person who was there
   __Would know what happened (0-2)
   __Would provide yet another, no more definitive perspective (3,5)
   __Would not know whole story, would provide firsthand account as evidence (4,5)
   __No mention

13. What mediates knowledge
    __Nothing (0)
    __Available information (1,2)
    __Bias, opinion (2,3)
    __Perspective; incomplete/unreliable evidence; methods of inquiry; interpretation (4,5)

14. Is at least one account reliable source?
    __Yes (0-2)
        __Both (0)
        __Insofar as the accounts are accurate but incomplete (1)
        __But only one is (2)
    __No (3-5)
        __Neither could be (3)
        __Neither completely reliable (4,5)
            __One could be found to be more reliable (4)
            __Each could be more or less reliable within perspective (5)

15. Means of assessing accuracy of facts from account
    __At least one account reports facts, which are accepted at face value (0-2)
    __One account is more persuasive, convincing, credible (2)
    __Isn't really one (3,5)
    __The accounts agree on facts (4,5)
    __Qualified corroboration of by another historian (4,5)

16. Accuracy possible
    __Yes (0-2, 4,5)
        __Without qualification (0-2)
        __To a degree (4,5)
    __No (3,5)

17. Best indicator of accuracy
    __It's appearance in accounts (0)
    __Most facts (1)
    __Story that makes most sense (2)
    __No indicator (3,5)
    __Good evaluation of evidence (4,5)

18. Assumption of underlying actuality
    __Yes (0,1,2,4)
        __Knowable (0-2,4)
            __Completely (0-2)
            __Only in part (4)
    __No (3,5)
        __There is no underlying, knowable reality (3,5)
        __Not a single actuality (5)

19. Fact and interpretation
    __No distinction made between fact and interpretation (0-1)
    __Distinction made between fact and interpretation (2-5)
        __Interpretation separate from fact, but facts determine correct account (2)
        __Interpretation determines construction of accounts (3-5)
            __Interpretation predominates to degree that facts are basically irrelevant (3)
            __Interpretation depends on evidence and is necessary (4,5)

20. Subjective element in interpretation
    __No (0,1)
    __Yes (2-5)
        __Objective predominates (2,4)
        __Subjective predominates (3,5)

21. Subjective emphasis
    __None (0,1)
    __Mistaken on facts (2)
    __Distortion of facts, or highly selective, slanted choice of facts (3)
    __Facts interpreted, receive different emphasis (4)
    __Facts selected and interpreted coinciding with perspective (5)

22. Information presented
    __Information presented as fact (although possibly incorrect) (0-2)
    __Information is opinion (3)
    __Information is interpretation and fact (4,5)

Appendix G: Livia Problem Coding Process

Step 1:

Participants' interviews were transcribed in an excel sheet. They answered 4 questions, and their answers were separated into 4 different columns. Below is a visual of one participant's answers.

Participant	Question 1 answer	Question 2 answer	Question 3 answer	Question 4 answer
7	They're two separate locations and from a northern perspective they were having a celebration and all of a sudden they were just ambushed. So out of being attacked, they responded with an attack. From a southern perspective, I guess the north was dominating in a sense and they came in to attack. They were winning at first but then got overpowered by the north because they were actually ready to attack. In the midst of it all a bigger guy came in and said this is not going to work and everyone said ok fine you'll do you're thing and I'll do mine. The north I feel has more of a complex whereas the south is like ok we're gonna make peace, but we're gonna ask before we go do something.	Yes, b/c the north I feel is that their perspective is right when they're under attack. There's no history on this is how things are going, this is how we were treating the south. The south was more or less fed up and they were gonna take charge more form the north already had started and the south was like here's some of our history and this is what happened.	Sure, but I feel that the attacked always has a different perspective than the attacker. Especially if they're starting the domino effect. If they're ultimately the reason why everyone got pissy. I always feel that the attacker doesn't always see everything as a full picture, they just see little chunks. I do feel they're both right but they don't start at the same point. And if they reference back into the past and what got them into this, then I think that we'll be able to tie down their perspective on what happened would be more keyed in. But I think we have to be in their shoes to see how right they are. I feel that the south as more truth in. It really all depends b/c the north neglected to tell us what was really going on and what happened prior to [the war]. So with that being said, they're not lying, they're withholding information, which could've explained why they were under attack to begin with. So I do feel that the north wasn't lying, they literally stated this is what happened. I definitely feel the north wasn't lying but neglected. But they were more true.	No it's all perspective because each side is going to have a different opinion and they're going to have a different side to that story. Even if being there, you still don't know the background history of what's going on. So I definitely feel that you can write but you weren't there and even if your getting it word of mouth you're still getting a fudged idea of what the story ultimately could be. We can get a ballpark idea.

Step 2:

Each answer was read first then the coders referred to the Livia Problem Coding Scheme. The first dimension was read, then each answer in the excel sheet was looked over again to see if any of the items could be applied. For example dimension 1 asked:

1. What is subject's core explanation of what differs?
   __No difference (0)
   __Difference noted (1-5)
      __Content of accounts of events differ (1,2)
      __Historians' perceptions of events differ (3-5)
         __Historians have different opinions (3)
         __Conditions under which historians conclude what happened differ (4,5)
         __Emphasis on different facts (4,5)

Given the answers provided from Participant 7, above, the 4th item-- Historians' perceptions of events differ (3-5)-- was applied because this person emphasized perspective in their answer. The numbers in the parentheses represent the epistemological levels for someone who receives this item. Therefore, epistemological levels 3, 4, and 5 were assigned to this participant for dimension 1 as seen below. This was repeated until all 22 dimensions were examined. If an item could not be assigned to a dimension, or if a dimension was not acknowledged in the participant's response, N/A was typed in each cell pertaining to that particular dimension.

Subject	**Dimension 1 Item**	**Level 0**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**Reason for code**
7	Historians' perceptions of events differ				3	4	5	Emphasizes perspective

Step 3:

All assigned levels were summed across the 22 dimensions. For instance, the sum of level 0 was calculated, level 1, level 2, level 3, level, 4, and level 5, separately. Participant 7 had a sum of 0 for level 0, 0 for level 1, 1 for level 2, 6 for level 3, 15 for level 4, and 19 for level 5. The

predominant level in this case was level 5, therefore, this participant was assigned to the evaluativist level as someone assigned to levels 0-2 are absolutists, 3 are multiplists, and 4 & 5 are evaluativists.

Step 4:

To establish inter-rater reliability, the second coder followed these same steps, blind to the first coder's responses. Subsequently, each dimension was discussed and identified as either agree (if both chose the same item for the given dimension) or disagree (if the coders chose different items for the given dimension). The percentage of agrees was calculated across dimensions by dividing the frequency of times the coders agreed by 22. In the case of Participant 7 the coders agreed on 18 of the 22 dimensions. Thus, the percentage agreement for this participant was 82% across dimensions.

Next, the 2 coders' predominant codes were compared. For participant 7 both coders' assigned level 5 as the predominant epistemological level. Therefore, the percentage agreement for this participant's predominant epistemological level was 100%.

Lastly, to establish overall agreement across dimensions and predominant epistemological levels, all participants' percentage agreement for dimensions and epistemological levels were averaged and the values were 83% and 96%, respectively.