

**Assessment of a Three-Year Argument Skill Development Curriculum**

Amanda Crowell

Submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
in the Graduate School of Arts and Sciences  
COLUMBIA UNIVERSITY  
2011

© 2011  
Amanda Crowell  
All Rights Reserved

## **Abstract**

### Assessment of a Three-Year Argument Skill Development Curriculum

Amanda Crowell

This study examines whether middle-school students' dense, extended engagement in an argumentation curriculum promoted development of argument skills, specifically increased use of direct counterargument and improved argument evaluation skill. A total of 56 students in two classes participated twice a week for three years (grades 6, 7, and 8) as part of their regular school curriculum. Students attended an urban middle school affiliated with a large university and were predominantly Hispanic and African-American and from lower and lower-middle socioeconomic backgrounds; 20% were from middle-class Caucasian families. In addition to its central element – electronically conducted pair dialogs on social issues - the curriculum encompassed a range of activities including small group preparation of arguments and reflective activities. A third class of 23 served as a comparison group; they also met twice a week over the same time period. They addressed similar social issues in more traditional whole-class discussion and wrote essays.

Assessments of dialogic argumentation skill and argument evaluation skill initially and at the end of each of the three years indicated that that the curriculum promoted the use of counterargument generally and the direct counterargument skill specifically. Performance of the experimental group increased over time in both respects and exceeded that of the comparison group. Students participating in the intervention also engaged in more sustained direct counterargument sequences than did students in the comparison group at the final assessment. Parallel improvements in argument evaluation

skill of the experimental group relative to the comparison group suggest that evaluation skill responds to practice much the same way as does argumentation performance. Theoretical implications for our understanding of developmental mechanisms are considered, as well as educational implications.

## Table of Contents

Introduction and Literature Review	1
The Present Study: Rationale	15
Method	18
Participants	18
Initial, Continuing and Final Assessments	20
Individual Argumentation Skill	20
Argument Evaluation Test	21
Intervention	22
Opinion Assessment	24
Year One Procedure	24
Team method	24
Electronic method	28
The Showdown and Debrief	30
Phase Four	32
Year Two Procedure	33
The Pre-game	34
The Game	35
The End Game	35
Year Three Procedure	36
Comparison Group Curriculum	39
Results	41
Transcript Coding	41
Counterargument Use	42

Sustained Direct Counterargument	45
Argument Evaluation	47
Discussion	49
Research Question: How does dense and extended practice in dialogic argumentation influence counterargument skill?	49
The Ceiling Effect	50
The Qualitative Case Study	54
Research Question: Do students in the intervention group sustain direct counterargument with an opponent for longer than students in a comparison group?	61
Research Question: Do students in the intervention group become more skilled in the evaluation of arguments relative to students in a comparison group?	61
Theoretical implications—Mechanisms	62
Practical and Educational Implications	64
Limitations and next steps	66
Conclusions	68
References	71
Appendix A: Experimental Group Equivalence	77
Appendix B: Capital Punishment Opinion Poll	78
Appendix C: Adaptive Response Schematic	79
Appendix D: Complete Description of Intervention Topics	80
Appendix E: Reflection Sheets	83
Appendix F: Showdown Preparation Sheets	85
Appendix G: Example of Year One Electronic Showdown Transcript	87
Appendix H: Complete List of Questions Provided and Answered	88
Appendix I: Example of Year Two and Three Showdown Transcript	95
Appendix J: Argument Game Script	100
Appendix K: Coding Scheme	101
Appendix L: Year-to-Year Change for Counterargument and Direct Counterargument	102
Appendix M: Argument Test 9-item Results	103
Appendix N: Indirect Counterargument Use	104
Appendix O: Full dialogs for students 32 and 62	105

## List of Tables

Table 1: Sample details	19
Table 2: Individual Argumentation Assessment Timeline	21
Table 3: Summary of Phase Topics	23
Table 4: Year One Session Activities by Group	33
Table 5: Example of the Argument game	38
Table 6: Frequencies and Proportion Use for All Codes Accounting for at Least 1% of Total Codes	42
Table 7: Verbatim examples of counter-alternative and counter-critique	43
Table 8: Mean of proportion counterargument use by group and time (SD)	43
Table 9: Mean of proportion counter-critique use by group and time (SD)	44
Table 10: Mean Rebuttal Sequence Length (Reduced N in parentheses)	46
Table 11: Mean of Score on 15 item test (SD)	47
Table 12: Ceiling Effectt	51
Table 13: Proportion Usage of Direct Counter Change Across Years and Subgroups	53

## List of Figures

figure 1: Year one schematic	31
figure 2: Super Card	35
figure 3: Mean of proportion counterargument use by group and time	44
figure 4: Mean of proportion counter-critique use by group and time	45
figure 5: Mean of Rebuttal Sequence Length by group and time	46
figure 6: Mean Score on 15 item argumexnt evaluation assessment	48
Figure 7: Ceiling Effect	51
figure 8: Case Study Student Trajectories	60
figure 9: Split Intervention Group Trajectories	60



## **Acknowledgements**

Thank you to my sponsor, Dr. Deanna Kuhn, whose expertise and guidance has shaped so much of my professional identity. Thank you to my committee members: Dr. John Black, Dr. Matthew Johnson, Dr. Jose Maldonado, Dr. Clark Chinn, and Dr. Michael Preston. Thank you to the members of the education for thinking lab (past and present!) without whom this project would have been a literal impossibility.

## **Dedication**

For Jane who loved, Floyd who believed, David who supports and Abigail who inspires.

## **Introduction and Literature Review**

The ability to generate and evaluate sound arguments has received increasing recognition as a fundamental aspect of good thinking and as an educational goal. The new Common Core Standards, adopted by roughly 75% of States in the U.S., dictate that students must be able to “justify their conclusions, communicate them to others, and respond to the arguments of others” as well as “compare the effectiveness of two plausible arguments (Common Core Standards, 2010, p. 6).” Similarly, the College Learning Assessment, the Graduate Record Examination and the Law School Admissions Test all explicitly assess these skills (Hersh, 2005; ETS, 2003; LSAC, 2008). The inclusion of argument generation and evaluation skills on such high profile indicators implies that they are necessary components for success in life in the 21<sup>st</sup> century. At the same time, the Nation’s Report Card on writing in 2007 indicates that 12th grade students’ performance on tasks involving argument are weak overall, and significantly weaker for African American and Hispanic students (NAEP, 2007). It is addressing this weakness, through the development and assessment of an innovative freestanding argument skill development curriculum for middle-school students, that is the goal of this work.

This study rests on a long history of inquiry that yields insight into what constitutes skilled argumentative reasoning and how to promote its development. In Ancient Greece, some of the most prominent thinkers in history argued whether skilled argumentative reasoning was about an argument’s decontextualized structural integrity or whether it was about its contextualized persuasive effectiveness. Billig (1996) notes that this tension is a recurring theme in the Socratic dialogues, with Socrates portrayed as a structural purist arguing against the rhetoricians of his time. Socrates argued that truth should be sought for its own sake and that if “truth” is revealed (by way of discussions aimed at revealing structural inconsistency within an opponent’s argument)

that stands contrary to one's position, one should let go of the position. For Socrates, then, a skillful arguer is one who can put together a logically defensible, explicit (syllogistic), and cohesive argument.

According to Billig, (1996) the rhetoricians (including Protagoras and Gorgias) took the position that truth is not a univocal, pure entity but rather a multivocal negotiation. In essence, they claimed that one can never escape rhetoric; there is no universal truth that can eliminate the need to argue and persuade. For the rhetoricians, then, the skillful arguer is the one who can flexibly respond to a range of disagreeing opinions.

This ancient tension is at the heart of the dual contemporary conceptions of argument. The first conceives argument as an individually constructed product. Under this conception, skilled argumentative reasoning is marked by constructing structurally variable and sound arguments. The second conceives argument as a process in which two or more individuals negotiate understanding in the presence of disagreement. Skilled argument under this conception is marked by one's ability to negotiate conversation directly and strategically. Researchers have utilized both of these conceptions of argument.

Contemporary structural approaches to evaluating argumentative reasoning follow from Stephen Toulmin's (1958) work, *The Uses of Argument* (1958). Toulmin, responding to the tradition of formal logic, stated that syllogisms (arguments such as "Socrates is a man [major premise], all men are mortal [minor premise], so Socrates is mortal [conclusion]") were not complex enough to encompass much thinking. Argument, according to Toulmin, requires that one not only present a structurally sound argument, but also justify one's claim in the context of

likely objections. To encompass this complexity, Toulmin reconceptualized argument structure as comprising more components than the premises and conclusions of the formal syllogism.

Specifically, Toulmin (1958) distinguished five components within an “argument pattern” which, taken together, state and justify a claim; the claim, the data, the warrant, the backing, the qualifier and the rebuttal. The claim is the statement one seeks to justify, the assertion to which one commits. The data is the evidence used to support the claim. The warrant is the link that connects the data and the claim. The backing is the general rule or law giving relevance and authority to the warrant. The qualifier puts limits on the certainty of the claim and the rebuttal acknowledges the circumstances under which the claim would no longer be true.

Toulmin’s argument patterns are applied to written or spoken transcripts of individually constructed arguments, but the social nature of argument is implicit in the pattern. Consider the following argument:

Chris will probably (Qualifier) not live near the University where she plans to work (Claim) because she needs childcare starting in two months (Data). All the childcare centers near the University have very long waiting lists. (Warrant) This is true unless she is willing to drive to a childcare center that is far away from where she both lives and works (Rebuttal).

The qualified claim regarding where Chris will live is supported by the data that she needs childcare soon. The author of this written argument anticipates that the reader may not see how childcare is related to where one lives (isn’t childcare available everywhere?). The author subsequently provides the warrant that available childcare options near her workplace are scarce. In acknowledgement of a likely counterargument, the author concedes that this argument is valid

unless Chris is willing to drive to a childcare center far away from where she lives and works. The fifth component, the backing, is left unsaid. Should the author anticipate that the reader may not see how waiting lists impact the likelihood of acquiring childcare, however, such a statement could be made explicit. Thus, while Toulmin's Argument Pattern evaluates arguments in a structural manner, it departs from formal logic in that the pattern evolves within a social context of justification.

Voss and Van Dyke (2001) note that Toulmin's Argument Pattern has been criticized as inadequate and unclear. Specifically, Voss and VanDyke argue that the pattern is not complex enough to apply meaningfully to larger pieces of written or spoken text. Efforts have been made to elaborate the model (i.e. Voss, Tyler, & Yengo, 1983), allowing for pieces of an argument to take multiple roles (that is, the data for one argument may also be the claim for another argument within the same piece), which allows for evaluation of longer, more complex pieces.

Such elaborations, however, may exacerbate the other criticism of Toulmin's argument model- that it can be difficult to designate a particular component of the model to a specific portion of an argument (Voss & Van Dyke, 2001). For example, it may not be clear whether a particular portion is the data in support of a claim or the warrant serving to explain how the data connect to the claim. Despite these criticisms, most researchers working within the tradition of argument as an individually constructed product have adapted or elaborated Toulmin's approach (Clark, Sampson, Weinberger, & Erkens, 2007).

Contemporary researchers working with adaptations of Toulmin's model generally identify a component that is underused or incorrectly used and then provide support to correct this deficit. For example, Wolfe and Britt have identified that counterarguments are often misunderstood or misused in expository writing (Wolfe & Britt, 2008). Given that arguments

containing counterarguments are given higher ratings of argument strength (Wolfe, Britt, & Butler, 2009), these researchers have sought to improve students' understanding of counterarguments. To do so they developed a counterargument tutor (Wolfe, Britt, Albrecht, Petrovic, & Kopp, 2009) that provides short pieces of didactic text that the program guides participants through, identifying counterarguments and other-side text. Students are then offered practice identifying counterarguments and feedback on their performance. Through such practice, college students became better able to identify counterarguments within text.

Researchers working within the tradition of argument as a process (of which the present work is a part) take the Vygotskian (1978) perspective that higher-order thinking skills first manifest through social interaction. According to Vygotsky, reflective thought develops first as argumentation occurring naturally among children at play and is then internalized. Thought, then, is an argument one has with oneself. Kuhn (1991) suggests that this view has powerful implications for education; if we are able to engage students in sophisticated argumentative discourse then we can foster sophisticated thought. Oaksford, Chater, and Hahn (2008) agree, claiming that argumentation is “the more general human process of which more specific forms of reasoning are a part” (p. 383).

Piaget's theory (1959) provides insight into how argumentation promotes cognitive development. When presented with new information, one must incorporate that information into one's existing cognitive schema. This is accomplished by way of assimilation or accommodation. Assimilation (appropriate if the new information is in concert with old ideas) incorporates new information into existing schemas. Accommodation (appropriate if the new information stands contrary to old ideas) requires that cognitive structures be modified to incorporate the new information. Until the new information is accommodated a state of

disequilibrium prevails, providing motivation for action. The more often a person engages in the process of accommodation the more sophisticated his or her cognitive schemas become, resulting in more sophisticated thought. Argumentation, particularly with peers, allows for genuine disagreement (Piaget, 1932), which creates disequilibrium and spurs accommodation.

Improvement in one's argumentation abilities, then, may also promote improved skill on a number of related reasoning tasks. Indeed, Means and Voss (1996) claim that engaging in argumentation improves storage and access to information in memory and that skilled argumentation promotes sophisticated problem solving and inference skills. Britt, Kurby, Dandotkar and Wolfe (2008) have found that precise memory for text is related to the person's ability to identify whether the text contains faulty arguments. Andriessen (2003) promotes argumentation as an effective learning tool well suited towards garnering conceptual change, promoting knowledge elaboration and increasing articulation.

Of interest to educators is the relationship between dialogic argumentation and individually constructed arguments like expository writing. Wolfe and Britt (2005, as stated in Wolfe, Britt & Butler, 2009) noted that while 89% of authentic written arguments (those written for genuine persuasive purposes) contained information that addressed an alternative perspective, only 50% of undergraduates addressed an alternative perspective in writing conducted under typical essay conditions (Wolfe & Britt, 2008). The authors claim that the reasons for such "myside bias" (Baron, 1995; Perkins, 1985) may be that students do not fully evoke an argumentation schema (or perhaps do not have a fully developed argumentation schema to evoke) and therefore do not see the relevance of the evidence not in support of their position. Chinn and Brewer (1993) posit a related, but more complex and interactive causal theory.

Chinn and Brewer's (1993) taxonomy of responses to anomalous data (data that does not



support one's prior beliefs) provides further insight into this deficiency. They suggest a continuum of responses one can take towards anomalous data and a number of interacting variables that would predict where one falls on that continuum. As an example, a person may hold the theoretical belief that a heavier object falls at a faster rate than a lighter object. When presented with data that all objects fall at a constant rate (such as watching a golf ball and a bowling ball fall from a common height and hit the ground at the same time) one must decide how to integrate this new information with the existing belief. One can 1) *ignore* the data without addressing any inconsistencies, 2) *reject* the data as collected in an erroneous manner, or 3) *exclude* the data claiming that it is not related to the theory in question. These responses allow one to ignore the relative validity of the data by declaring it irrelevant (either implicitly by ignoring it, or explicitly by rejecting or excluding it). Chinn and Malhorta (2002) argue that the impetus for many of these rejections and exclusions rest on breakdowns in the psychological process of observation. If one is willing (and able) to see the new data as relevant to the existing belief then he or she could 4) hold the information in *abeyance*, holding that the current theory will eventually explain the noted discrepancy 5) *reinterpret* the data, finding a way to see it in concert with the previously held belief, 6) make *peripheral changes* to the theory while leaving the core theoretical constructs intact or 7) make wholesale *theoretical changes* to the previously held belief.

Where one falls on this continuum is dependent on a number of interacting variables: the nature of one's prior knowledge, the nature of the new theory, the nature of the anomalous data, and how one approaches the process of integrating these different aspects. Deeply entrenched beliefs are intertwined with a number of other beliefs and are harder to change (with those beliefs about the fundamental nature of the world- ontological beliefs- as the hardest to change of all),

especially if one has a relatively undeveloped epistemological understanding of science. That is, if one does not understand the nature of science and scientific endeavors as constantly refining and evaluating theories in light of evidence, one would be more likely to hold fast to a theory despite disconfirming evidence. Additionally, people are more likely to deeply process claims that disconfirm their previously held beliefs if those claims are personally relevant. Deep processing promotes theory change in that it gives the inter-relationship between theory and evidence time to coalesce (coordinating and evaluating the alternative theories with all available evidence). Taken together we see one of the ways that these variables might interact to determine where a person falls on the continuum of responses: if the anomalous evidence is about an idea that is deeply entrenched then it would need to be particularly personally relevant to result in the kind of deep processing required for theory change, otherwise, we would expect to see a response that falls further down the continuum. It may be that such an interaction effect applies both to how one adjusts one's theory to account for anomalous data, and whether one chooses to include anomalous data (or data representing the alternative perspective) in one's writing.

Chinn and Brewer (1993) claim that the impact of anomalous data does not only depend on the more objective qualities of the components under consideration (such as the credibility of the data and the intelligibility of alternative theories), though these are important. The impact is also dependent on psychological qualities of the person making the distinction such as whether one sees the effort in which he or she is engaged (refinement of a theory in light of evidence) as an important part of science. This perspective echoes Kuhn's (1991, 2001, 2005) and others' (for review see Hofer & Pintrich, 1997, 2002; Chinn, Buckland & Samarapungavan, 2010) claim that effective argument requires both well-developed cognitive abilities and an advanced epistemology. Kuhn argues that one's beliefs about knowing can progress through three levels of

understanding: the absolutist level, the multiplist level, and the evaluativist level. The absolutist understanding of knowing, typical of children before adolescence, is characterized by a belief that “truth” is an available commodity; if one asks the right person, one can come to acquire the truth about any issue or concern. Pursuit of knowledge then, is a fact-finding mission; knowledge lies outside of the knower and has an intrinsic and objective value, in and of itself.

The multiplist perspective is quite the opposite. In adolescence, one often comes to the realization that “truth” is a highly contentious construct. Experts disagree, well-established theories are determined false, and one’s friends and family may hold widely different views with equal passion. In light of this recognition, one’s conception of knowledge shifts from an objective external entity to an entirely subjective and internally driven opinion. Pursuit of knowledge, then, under the multiplist perspective is a fool’s errand, with information destined to be debunked once acquired. This transition, Kuhn argues (2005), is often extreme and permanent; many do not transition into the evaluativist level, even as adults.

The evaluativist perceives knowledge as an entity reflecting both objective and subjective components. To construct knowledge one accesses relevant evidence and evaluates a theory as more or less sound in light of it. Though one must admit that the soundness determination is always somewhat subjective (as it is guided by one’s values and beliefs), one can still lay claim to a place in the conversation. Thus, while it is not possible to return to the simplistic notion of absolute truth, the purpose of knowledge seeking is restored. In fact, because knowledge is constructed by the knower, the relative quality of that knowledge is contingent on the efforts of the knower. Thus, knowledge seeking becomes more purposeful than it is for the absolutist because the seeker is entirely responsible for the quality of information available for decision-making. Being of an evaluativist perspective promotes more sophisticated reasoning and enables

construct change by providing the motivation necessary to seek out and accommodate alternative perspectives. In short, it gives argument a purpose.

Kuhn (2005) argues that students must not only have the skills to engage in counterargument and rebuttal but they must also see the point of utilizing them. This requires that students see their ideas as situated in a universe of alternative ideas that must be brought together and evaluated. Graff (2003) claims that engagement in dialogic argumentation has the potential to provoke that purpose in expository arguments. By fostering an internal interlocutor (through the internalization of the argumentative self developed through exposure to differing viewpoints in dialogic argumentation), writers are better able to anticipate and refute likely counterarguments, thereby improving the quality and complexity of their expository writing.

There is evidence that dialogic argumentation is, indeed, a promising path to the development of more sophisticated expository argument. Kuhn, Shaw and Felton (1999) found that students who engaged in a series of arguments about a topic of social importance wrote expository essays on the same topic that were more two-sided than those written prior to the intervention. That is, after engaging in dialogic argumentation students wrote essays that considered the perspective of the opposing position and sought to counter those arguments.

At this point, it is worth emphasizing the consensus in the literature regarding the relative importance of counterargument. The development of a willingness and ability to engage in counterargument is a constant across both structural and dialogic work on argument. Including counterargument in expository writing provides a necessary purpose that guides the writer's goals (Graff, 2003) and improves the perceived quality of expository writing (Wolfe, Britt, &

Butler, 2009). Counterargument engages a social context of alternatives and as such, promotes a shift from the multiplist to the evaluativist epistemology (Kuhn, 1999, 2001, 2005).

Given its intrinsic value and potential for broad transfer, researchers have attempted to promote counterargument development in educational settings. One such line of research is the work done by the Collaborative Reasoning lab at the University of Illinois (Reznitskaya, et al., 2009). Collaborative reasoning engages late elementary school (4<sup>th</sup>-5<sup>th</sup> grade) students in class discussions around texts that deal with “big questions” (such as animal rights and whether to turn in a friend who cheats). The process of Collaborative Reasoning is designed to allow for egalitarian dialogue within the classroom setting. Students do not volunteer to participate and receive permission to speak; rather students speak freely without explicit direction from the teacher. Indeed, teachers are trained to provide guidance as needed to support the argumentation taking place between students, and to lessen their involvement as the students become more skillful.

Collaborative Reasoning relies on previously explained notion of an argumentation schema that is developed and elaborated by engaging in peer-to-peer dialogs. As one’s schema develops, one’s argumentation should become more varied and sophisticated and one’s commitment to utilizing discourse to explore complex issues should strengthen (Reznitskaya, et al., 2009). The researchers report that the use of Collaborative Reasoning serves to increase the number of reasoned (if imperfect) arguments, the degree of student participation, and the number of open-ended questions posed by students and teachers (Reznitskaya, et al., 2009).

This argument schema can be broken into recurring patterns, or argument stratagems that represent strategic moves within the dialogic space. As students engage in Collaborative

Reasoning, their argument stratagems increase in complexity and become more varied (Reznitskaya, et al., 2009). Anderson, et al. (2001) claim that one process by which argument schema are elaborated is by trying out argument stratagems that one hears from one's conversational partners. Specifically, once a stratagem was used by a participant in a particular group, other members of that group would employ the stratagem with increasing frequency. This is referred to as the snowball phenomenon since the frequency of use of a particular strategy "snowballs" after its initial presentation.

Additionally, there is evidence that individual expository writing improves with engagement in Collaborative Reasoning. Basing evaluation on a rubric similar to Toulmin's model, researchers found that students who engaged in Collaborative Reasoning wrote essays that contained a greater number of arguments, counterarguments, and rebuttals than did the essays of students in the comparison groups (Reznitskaya, et al., 2009). There is also preliminary evidence that participating in Collaborative Reasoning may contribute to writing that considers multiple perspectives and solutions (Reznitskaya, et al., 2009).

The present work is situated within the framework suggested by Kuhn (Kuhn & Crowell, in press; Kuhn, Goh, Iordanou, & Shaenfield, 2008; Kuhn, Shaw & Franklin, 1999), which stresses the importance of providing practice with dialogic argumentation to improve the quality of the dialog itself as well as individual thought. By engaging in dialogic argumentation one is exposed to alternative viewpoints and gains experience countering them. Such exposure may begin to shift one's epistemology from the multiplist to the evaluativist, thereby further improving one's ability to engage and counter the alternative perspective. As this reciprocal cycle continues, one's reasoning becomes more sophisticated, one's epistemology advances, and one's individual argument improves in purpose and complexity.

Thus, as one's dialogic argumentation becomes more direct and evaluative one's argumentative thinking and writing should improve. Since it is the quality of the dialog that is promoting the improvement in thinking, Kuhn's work assesses argumentation in the tradition of Walton's Dialogue Theory, rather than Toulmin's Argument Pattern. According to Walton (1989), skilled argumentation has two goals. One is to secure commitments from the opponent that can be used to support one's own argument. The other is to undermine the opponent's position by identifying and challenging weaknesses in his or her argument. Both of these require that one pay careful attention to the opponent's position and seek to challenge it.

When one seeks to challenge a position put forth by one's opponent, one is said to use a counterargument strategy. We distinguish between two forms of counterargument: the direct and the indirect counterargument. The more effective strategy (in that it would be more likely to meet the dual goals of argument as laid out by Walton), the direct counterargument, responds to a statement with which one disagrees by identifying a weakness in the opponent's argument. The less effective strategy, the indirect counterargument also involves disagreeing with what the opponent has said but then providing a different reason to be in support of one's own position, leaving standing the opponent's argument.

Felton and Kuhn (2001) found that middle-school aged adolescents address an opponent's arguments 11% of the time during argumentation and adults attending a junior college address an opponent's arguments 24% of the time. Felton (unpublished manuscript) engaged a small sample of lawyers (presumably experts at argumentation) in argumentation on the topic of capital punishment and found that they performed well above these levels, with more than 50% of utterances directly addressing an opponent's argument. Taken together this evidence indicates that while maturation alone may provide the necessary foundation for skilled

argumentation development, it is by no means a sufficient condition. Rather it is experience and practice that promote its development (Kuhn, 2005).

An additional goal of this study is to determine whether argument evaluation develops in tandem with a propensity to utilize the sophisticated counterargument strategy in argumentation. Much argument evaluation research focuses on the role of evidence. Barchfield and Sodian (2009) state that only 49% of university students rated the quality of well evidenced arguments as superior to that of flawed arguments, and while 80% of adults rated well evidenced arguments higher, only those with a university education could articulate that it was the evidence that prompted the higher rating. Many studies substantiate and elaborate this claim, generally showing that adults and adolescents may not be able to distinguish between explanations and evidence.

When assessed, both adults and adolescents are usually able to generate explanations, but much less likely to be able to generate evidence (Kuhn, 1991; Glassner, Weinstock & Neuman, 2005). Although evidence generation performance improves when prompted (Brem & Rips, 2000; Kuhn & Udell, 2007), it is rare for people to use the evidence generated to inform arguments, even with very specific prompting (Bullock, Sodian, & Koerber, 2009). Additionally, students don't seem to value the contribution made to an argument by evidence, even when explicitly asked although with prompting they do exhibit some understanding of the different roles each play (Glassner, Weinstock & Neuman, 2005). Kuhn and Udell (2007) argue that students have difficulty accomplishing the cognitively demanding task of coordinating evidence with the theory across multiple perspectives (though they are capable of doing so when the cognitive demands are reduced) and that this is exacerbated by an immature epistemological understanding.



Given the potentially unstable nature of people's understanding of evidence, we opted for an argument evaluation measure that didn't rest on such an understanding. Rather, the measure rests on whether students value direct counterarguments above indirect counterarguments. This is the first such instrument having such a purpose to our knowledge, and preliminary results indicate that it may be more difficult to promote appreciation for direct counterarguments put forth by others than it is to promote the usage of direct counterargument (Goldstein, Crowell, & Kuhn, 2009). Evaluating the directness of a response made in reply to a prompt (the nature of the items on the assessment) regardless of whether one personally agrees with the claim requires counterfactual thinking in a way that promoting a position one believes in does not. The present study evaluates whether argument evaluation of this nature can be developed through the same principle of engaged practice that we hypothesize guides argument generation.

### ***The Present Study: Rationale***

The present study engaged paired young adolescents in electronically mediated argumentation and reflective activities over the course of three years, with the goal of improving the quality of the dialogs in which students engage by promoting use of counterargument. Kuhn (2005) argues that early adolescence is the ideal time to begin promoting the development of argumentation skills since variability in cognitive competency begins to widen at this time, indicating that environment is beginning to take a larger role (Kuhn, 2005). Additionally, Kuhn (2000) notes that metacognition (the ability to think about one's own thinking) is poorly developed until adolescence. Given that engaging in the counterargument strategy requires that one see one's position in a universe of alternatives, it would be difficult to promote improvement in argumentation in individuals without access to metacognition.

In the earliest studies done to develop students' counterargument skills, students engaged in a series of goal-directed small group activities (Kuhn, Shaw & Felton, 1997; Lao & Kuhn, 2002, De Fuccio, Kuhn, Udell & Callender, 2009), as well as mixed curricula involving small groups and verbal dialogs (Kuhn, Shaw & Felton, 1997; Kuhn & Udell, 2003). More recent versions of this curriculum have focused primarily on electronically conducted dialogs (Kuhn, Goh, Iordanou, & Shaenfield, 2008; Kuhn & Crowell, in press a). These studies show that engaging in dense practice with argumentation over time frames ranging from a few weeks to nine months leads to improved counterargument skill. The current study extended this line of research to an intervention that spanned the three years of middle school.

The present study utilized several practices to promote metacognitive awareness and reflection. We take the position that improvements in argumentation will occur as students become more skilled at evaluating their own thoughts (Kuhn, 2000). The practice of engaging students in electronically facilitated (via instant messaging software) dialogs provides an opportunity for students to review and reflect upon the arguments they have generated (Kuhn, Goh, Iordanou, & Shaenfield, 2008). This stands in striking contrast to face-to-face dialogs where the spoken word becomes open to misrepresentation by one's opponents and to the inevitable imperfections of one's own memory, as it disappears immediately after being spoken. To further encourage reflection, students were asked (as both in class work, and homework) to reflect upon their strategy use during the dialog. Such reflective activities have been shown to increase metastrategic awareness and increase the presence of counterargument (Kuhn, Goh, Iordanou, & Shaenfield, 2008).

Since same-side pairing essentially doubles the exposure to reasoned discourse (students must first agree with the same-side partner about what to submit to the dialog with an opponent

pair and then reason within the dialog itself), students worked in pairs during the dialogs. Shaenfield (submitted manuscript) has demonstrated that pairing students for dialogs significantly improves counterargument usage, compared to students who work alone.

The goal of the present study is to evaluate how dense and extended practice promotes skilled argumentation and argument evaluation over the course of the three years of middle school. A repeated measures design was used to assess students initially and at the end of each academic year. The present work addresses these questions:

- 1) How does dense and extended practice in dialogic argumentation influence counterargument skill?
- 2) Do students in the intervention group sustain direct counterargument with an opponent for longer than students in a comparison group?
- 3) Do students in the intervention group become more skilled in the evaluation of arguments relative to students in a comparison group?

## **Method**

The intervention engaged students in dense argumentation practice involving eleven topics over a three-year period. The research employed a pre-post repeated measures design and included a non-intervention comparison group to compare the effects of the intervention on dialogic argumentation skill and individual argument evaluation skill.

### ***Participants***

Participants were students attending an urban, public middle school in the Northeast United States. They had just entered sixth grade at the beginning of the study, and had completed eighth grade by the end of the study. The intervention began in the school's first year of operation, and these students were the first incoming class. Students came from an ethnically diverse low- to- middle-income surrounding neighborhood. Eighty percent of students were either Hispanic or African-American, and sixty percent qualified for free or reduced-price lunch. All students were eleven or twelve years old at the beginning of the intervention. The school was academically challenging and emphasized a math, science and engineering curriculum. The school boasted a 95 percent average attendance rate.

Students were assigned by the school administration to one of three "houses", each containing about 30 students. These class assignments were random, without consideration of academic or personal characteristics. It was subsequently verified that the three houses did not differ significantly by gender, ethnicity or incoming math and English standardized test scores. Two of the three houses were randomly selected to receive the experimental curriculum and the third served as the comparison group (the two experimental groups are treated as one group as students did not differ across the course of the intervention in terms of argumentation skill

[ $F=1.22$ ,  $p=.312$ , partial eta squared=.06], see Appendix A for details). Students remained in the houses in which they were initially placed, with the exception of three students who changed houses to address behavioral or administrative needs; one of these three changed from the comparison group to the experimental group in the middle of year one (Y1) and two (Y2) changed from one experimental group to the other experimental group in the middle of Y2. All students who changed houses have been maintained in the sample.

Given the length of the intervention (three academic years) and the moderate mobility of the population, not all students were present throughout the study, i.e., from the beginning of 6<sup>th</sup> grade to the end of 8<sup>th</sup> grade. Only students who attended the school for at least two years and were present for an initial assessment and the final assessment were included in the analysis, leaving a final sample of 79 (40 female) students, 56 (30 female) in the experimental group and 23 (10 female) in the comparison group (See Table 1 for further details of sample, including sample exclusions).

**Table 1: Sample details**

House	Included			Excluded*			Total
	Present 6 <sup>th</sup> -8 <sup>th</sup> Grade (3 years of data)	New in 7 <sup>th</sup> grade (2 years of data)	Total	Left School/ Retained	Missing Initial	Missing Final	
Experimental	49	7	56	13	1(1)		14
Comparison	21	2	23	6	4(3)	4(1)	14

\*Students new in 7<sup>th</sup> grade in parentheses

The sample for the argument evaluation test is smaller as it was not given to students who came into the school in year two (a reduction of eight students), and excludes seven more

students due to missing assessments. The final sample for the argument evaluation test is 63 students (44 experimental).

### ***Initial, Continuing and Final Assessments***

#### *Individual Argumentation Skill*

Students engaged in initial and continuing assessments of individual argumentation skill in September of year one (Y0) and then in December and May of year one (Y1) and December and May of year two (Y2), and May of year three (Y3) of the intervention (Table 2). The initial (Y0) assessment began with a survey soliciting the student's opinion on the use of capital punishment (CP) and the strength of that opinion (Appendix B). On this basis, students were matched with a partner whose opinion on CP opposed theirs (for or against) and whose strength of opinion matched theirs (this was done to ensure that a person with a strong opinion would not quickly overpower a person with a less strong opinion). For example, one who was "for" capital punishment and was "very sure" of his or her opinion was matched with one who was "against" capital punishment and was "very sure" of his or her opinion. These initial pairings provided the basis for all future pairings on this assessment. Of initial pairings, 22% remained the same across all reported times, while 73% occurred at least twice. Pairings were changed when students were absent on the day of the assessment (and were assessed during a makeup session) or when one of the partners had left the school.

At the assessment, students were informed that they were paired with a person who disagreed with them on the issue of CP. They were instructed to try to convince the other person that their own opinion was the better one, beginning with the pro-side student finishing the statement "I am for capital punishment because..." No other instructions were given. The

argument was conducted in writing. The two members of a pair passed between them a pad with multiple sheets of paper, blank except for the left margin, on which appeared (at intervals three inches apart) “A says,” “B says,” “A says” and so forth, and on which they conducted their dialog about CP. Given the electronic nature of the intervention, assessments were conducted in this low-tech “pass the pad” manner to avoid putting the comparison group at a disadvantage given the experimental group’s concentrated use of the electronic medium. These handwritten “pass the pad” dialogs were transcribed for analysis.

**Table 2: Individual Argumentation Assessment Timeline**

	Pre-test	Mid-test	Post-test
Year 1	September 07	December 07*	May 08
Year 2	September 08**	December 08*	May 09
Year 3	NA	NA	May 10

\*Excluded from analysis

\*\* Initial assessment for new students only

#### *Argument Evaluation Test*

Students engaged in an assessment of individual argument evaluation skill at the beginning of Y2, end of Y2, and end of Y3. To complete the assessment, students logged into an online assessment tool. Students read instructions stating:

“You are on the subway listening to your friend, Pat, argue with another friend (Lee) about why some kids fail at school. Pat is having trouble and asks for your help winning the argument against Lee. For each statement Lee makes, Pat must decide between two possible responses. To help Pat, choose the stronger of the two counterarguments. After you choose the better counter, click “next” to see how Lee responds. Remember, it does not matter what you believe; you just want Pat to win the argument against Lee.”

Lee begins with a claim about why some students fail in school. Upon reading Lee's argument, the student is presented with two options for how to respond. One choice is a direct counterargument (i.e. one serving to weaken the force of Lee's argument), and one is an indirect counterargument (i.e. one that offers an alternate argument but leaves the opponent's argument unaffected). Lee then rebuts (via a direct counterargument) the statement chosen by the student and the student is presented with another two response options. See Appendix C for the adaptive-response schematic on the topic of why students fail in school.

The first administration of this assessment presented three items on each of three topics (why students fail in school, why many criminals return to a life of crime after being released from jail, whether movie stars should make as much money as they do), for a total of nine items. To increase the range of the test, the second and third administration added two more topics (whether cigarettes should be outlawed, and whether technology is having a good effect on our lives) for a total of fifteen items.

### ***Intervention***

The two houses assigned to the experimental curriculum participated in an intervention designed and administered by our research team. Year one of the experimental curriculum included a comparison of two curricula, a team method and an electronic method. The two methods were used because the school had insufficient computers for both classes to use the electronic method at that time. Both methods had been used successfully in previous research, the team method by Kuhn and Udell (2007) and the electronic method by Kuhn, Goh, Iordanou & Shaenfield (2008) This comparison is described in *Year One Procedure*, below. Years two and three comprised a best-practice integration of the two methods compared in Y1; this method



is described in full in *Year Two Procedure*. *Year Three Procedure* describes additions to the Y2 curriculum made in Y3 as student skill development advanced. Over the course of three years students engaged in eleven phases. In each phase they addressed a different topic (See Table 3 for a summary of topics, and Appendix D for a complete description of topics). The two intervention groups and one comparison group each met as intact classes for a 50-min stand-alone class, identified as a class in philosophy, twice weekly throughout the three academic years (a total of roughly 145 sessions).

**Table 3: Summary of Phase Topics**

Phase #	Topic	Question
1	School Expulsion	Should students who disrupt the classroom be expelled after several warnings or not?
2	Homeschooling	Should a family have the right to homeschool their child or not?
3	USA Intervention	Should the US get involved in a hypothetical peacekeeping mission in South America or not?
4	Kidney Sales	Should a person be allowed to sell a kidney to another person who needs it or not?
5	Animal Testing	Should medical research labs use animals to test medicine for humans or not?
6	One Child Policy	Should China have a One Child Policy or not?
7	Presidential Priority	Should the new president focus on domestic or international issues first?
8	Intellectual Property	Should people be allowed to do what they wish with songs, books, etc that they buy or does the creator maintain legal control?
9	Juvenile Justice	Should teens be tried by 1) a jury of their peers, 2) in a juvenile justice system administered by adults, or 3) in the adult justice system?
10	Community Service	Should Congress enact a program that funds a year of community service 1) between high school and college, 2) a year after college, or 3) should congress refrain from funding such a program?
11	High School Curriculum	Should the high school curriculum allow students to 1) choose all their own classes, 2) engage in core classes in the morning and allow for approved activities of the students choice in the afternoon, or 3) should the curriculum remain as it currently is?

### *Opinion Assessment*

Before each phase began, an opinion poll was administered, asking students their opinion on the topic and the strength of that opinion. Students' opinions provided the basis for assigning students to positions and to pairs and groups within positions.

### *Year One Procedure*

Both intervention classes discussed the same four topics, one per phase of the intervention. For variety, during the third phase the assignment methods were reversed. That is, those students who used the electronic methodology for most phases followed the team method for the third phase, and those who followed the team methodology for most phases followed the electronic methodology for the third phase. This was done to give students exposure to both methods.

### *Team method*

The team method comprised nine sessions per phase, conducted in small groups. Students expressing the same opinion were assigned to the same team. Each team had two subteams of about 8 students who worked together during the subsequent sessions. Each subteam of eight students worked with an adult coach creating a ratio of 1:8, coaches to students. The procedure during each of the nine sessions is described in detail, below.

*Session 1. Generating reasons.* Students were introduced to the topic through a scenario sheet, which was an elaborated version of the opinion poll (see Appendix D). This first session began with ensuring that students understood the topic and the question at hand. Once all students showed understanding they were encouraged to discuss *why* they thought as they did on

the topic. Each student was given a large index card on which to write his or her “most persuasive reason” for why his or her side’s position was the better one. Students were provided with the rationale that it is important to get their reasons “out” so that they can, as a subteam, decide what they think of them. Once students had written their reasons on a card, each student took a turn reading his or her card aloud. After each student shared his or her reason, the coach asked the group, “What do you think of this reason? Does it make sense?” All reasons generated (and supported by even one member of the subteam after the discussion) were included in the set of reasons that would be used going forward.

*Session 2. Eliminating Duplicates.* In this session students were encouraged to interpret reasons (“What does this one mean?”) and to eliminate duplicates (“Is this new reason or one we already had?”). Students were encouraged to reflect on the reasons provided by their subteam mates and consider how their own reason was related. Students cycled through this process as often as was necessary to eliminate duplicates and ensure that they had generated all possible reasons. Coaches did not provide any reasons or make any decisions; rather, they only facilitated the students’ generation and evaluation process.

*Session 3. Elaborating reasons.* Students were asked to consider the relative worth of each of their generated reasons. This led to a discussion of what makes some reasons better than others, and then an evaluation of their reasons based on these criteria. Reasons were re-worded so that they would be as persuasive as possible. Students were asked to consider how the other team might respond to their reasons.

*Session 4. Evaluating reasons.* This session involved a more formal evaluation of the reasons generated in the previous two sessions. In pairs (or occasionally, triads), students were

provided with duplicate sets of the reason cards and asked to sort them into three categories (“best,” “good,” and “not so good”), discussing placement among themselves until they agreed. Once each pair had done this, the process was repeated as a whole group. A reason remained in the top category only if students could argue *why* it belonged there.

*Session 5. Developing reasons into an argument.* This session focused on building a supported argument, using a set of interconnected reasons. Using the metaphor of building a house of bricks, students were encouraged to build their reasons into an argument, with “best” reasons as the foundation, “good” reasons as the supports, and others as back-ups if needed. Initially in pairs and then all together, the subteams worked to construct and format an argument based on their strongest reasons on poster board. This activity fostered discussion of which reasons to use, the relationship between reasons and the possibility of combining reasons.

*Session 6. Examining and evaluating opposing side’s reasons.* At the beginning of this session students were asked what they thought the opposing team had been doing during this time. Students recognized that the other team was trying to come up with reasons of their own. Students were encouraged to try to anticipate these reasons, consider how valuable it would be to see the other teams reasons, and discuss what the students would do with the other team’s reasons, if they were to have them.

The coach then disclosed that an arrangement had been made with the coaches of the opposing team to exchange reasons. Duplicate sets of the opposing team’s reasons were distributed to the subteams. Students were then further divided into subgroups of 3-4 people to evaluate the opposing side’s reasons by the same process they had evaluated their own reasons, i.e. by sorting them into “best”, “good”, and “not so good” categories.

*Session 7. Generating counterarguments to others' reasons.* The generation of counterarguments (statements that directly critiqued the other side's reasons), which began spontaneously in the preceding activity, was formalized in this session. Students were divided into pairs and given colored "counter" cards in order to generate and record counterarguments to each of the other-side reasons. The coach encouraged students to consider what the opponents were trying to say and then construct a counterargument that would invalidate, weaken, or "shoot down" that reason. Subteams were then reassembled to deliberate about which of the generated counterarguments was the strongest counter to the other side's reason. Once decided, that counter was attached to the reason card.

*Session 8. Generating rebuttals to others' counterarguments.* In this session, the team's reason cards were returned to them, with counterargument cards produced by the opposing team attached. They were given several minutes to absorb and react to the opposing team's counterarguments. The subteam was then tasked with how to respond. Should they strengthen their reason to avoid the criticism, rebut the criticism, or drop the reason? Where appropriate, rebuttals ("comebacks") were generated.

*Session 9. Conducting & evaluating two-sided arguments.* Students devoted this session to rehearsing their own reason/counter/rebuttal sequences, as well as to reviewing the sequences of opposing team's reasons and counters. Team members volunteered to play the role of the opposing side in these practice exchanges. Team members took turns participating and critiquing one another's efforts. At the conclusion of the session, plans and guidelines were discussed for the Showdown (discussed in detail below) final debate to take place the following week.

### Electronic method

*Student pairings.* Students were paired for the duration of the topic with a like-minded peer (i.e., one who expressed the same opinion on the topic). When the opinion split in the class was not exactly 50/50, students whose answer to the question “How sure are you of your opinion?” was less than “sure” were assigned to the other side (Median number of students within one class moved from expressed side= 5, Range 3-8). As a pair, these two students argued electronically with an opposing-opinion pair. Pairings lasted throughout the phase. Two students were paired for only one topic each year.

*Electronic dialog.* During the first seven sessions of each phase students engaged in electronic dialog. Dialogs took place via Google-chat, an internet chatting program integrated into Gmail (the program the school uses for student and faculty email accounts). Pairs were given a specified email account used solely for these dialogs. A particular pair argued with seven different pairs from the opposing team, one during each of the seven dialogs. Students were directed as to who would begin the dialog and instructed to take turns and be civil; no other directions were provided. Each dialog lasted approximately 25 minutes.

*Reflection sheets.* While waiting for the other side to complete their turn, pairs were asked to complete a reflection sheet (Appendix E). The pair completed one reflection sheet per session. The sheets were of two forms – those focusing on “own argument” and those focusing on “other side’s argument.” On both sheets students were asked to reflect on how well they had responded to what the opposing pair posed. The first task was to choose the most important argument made either by their own side or the other side, depending on the sheet. On “the other side’s argument” reflection sheet, students were asked to reflect on how they had responded to

the other pair's best argument and how they could have responded more effectively; they were asked to "write a better, more effective counterargument." Similarly, when working on the "own argument" reflection sheet students were asked to note how the opposing pair had responded to their argument, state how they had rebutted (referred to as a "comeback") that response, reflect on how they could have rebutted more effectively and then "write a better, more effective comeback."

### Showdown preparation

The sequence for both groups culminated in a Showdown. In the team method, preparation for the Showdown was incorporated into session nine of the method described above. In the electronic method, preparation for the Showdown occurred over two sessions. Each member of a pair was placed into a different subteam. A total of four subteams were created for the Showdown preparation sessions (two for each side of the issue). On the first day of Showdown preparation, students completed one of two tasks, on the second day of Showdown preparation they completed the other task. One subteam started with task "A", one team started with task "B", and then they switched tasks for the second day.

Task "A" focused on the side's own arguments. Students were instructed to complete one Showdown preparation sheet for each of their arguments (Appendix F). The Showdown preparation sheets asked for a summary of an argument, the best counter that their opponents might pose, and how they could best rebut that counter. Subteams working on this task on the first day of Showdown preparation were asked to generate as many full sequences as they could (starting with what they considered to be their strongest argument). Subteams working on this task on the second day of Showdown preparation were asked to pick up where the other subteam

had left off, completing incomplete sequences and making any new sequences the other subteam may have forgotten. Students working on task “A” were given copies of all “own” reflection sheets completed during the dialogs to use for reference.

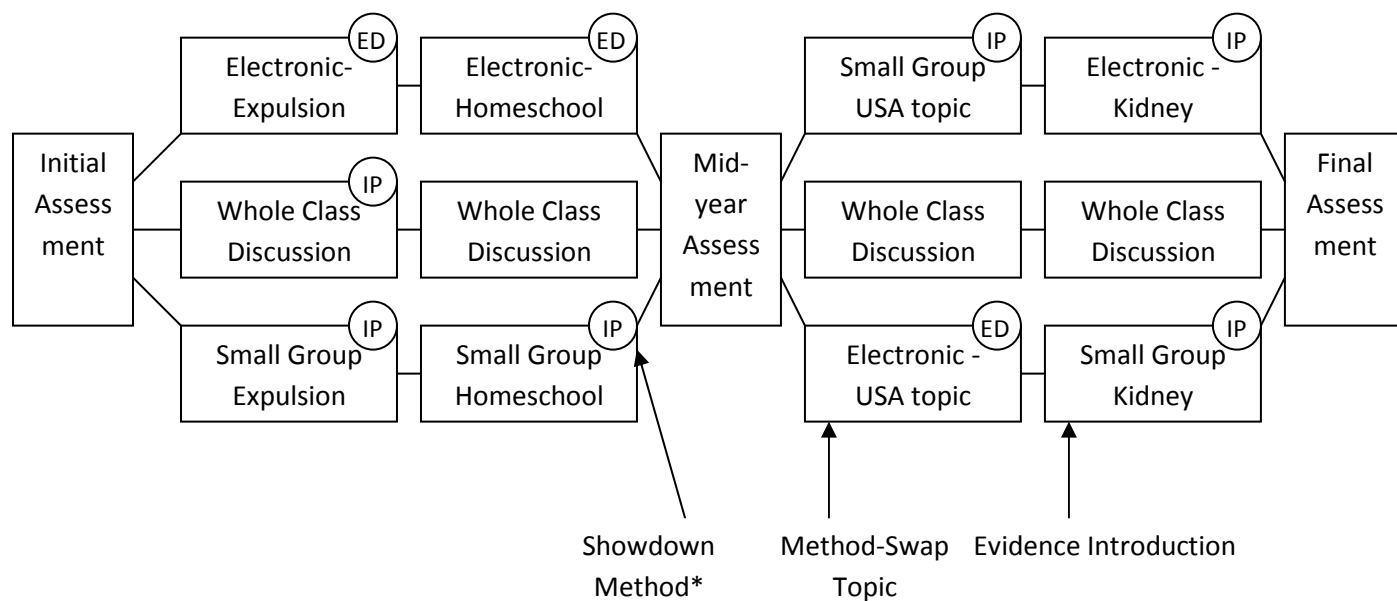
Task “B” focused on the other side’s argument. Students were instructed to fill out one showdown preparation sheet for each of the other side’s arguments (Appendix F). The Showdown preparation sheets asked for a summary of an argument the other side might make and the best counterargument they could make. Subteams working on this task on the first day of Showdown preparation were asked to generate as many as they could; students working on this task on the second day of Showdown preparation were asked to pick up where the other subteam had left off, completing incomplete sequences and constructing any new sequences their team mates had forgotten. Students working on task “B” were given copies of all “other” reflection sheets completed during the dialogs to use for reference.

### *The Showdown and Debrief*

The Showdown was a verbal or electronic whole-class debate. The debate structure alternated between electronic debates and in-person debates (See figure 1 for details). For both Showdown types, students remained in their subteams, with the first half of the Showdown engaging members from two subteams (one subteam per team), and the second half of the Showdown engaging members from the remaining two subteams. While students were engaged in the Showdown they were seated at the “hot table.” Those students not seated at the “hot table” were told that they could write suggestions that would be passed to the “hot table” for consideration.



**figure 1: Year one schematic**



\*IP= Showdown done in person, ED=Showdown done via electronic dialog

For Showdowns using the electronic dialog methodology, students at the “hot table” discussed and collaborated on a typed response, usually taking about five minutes per turn. The argument was then sent to the other team (which was in another room) and that team then collaboratively responded. Showdowns lasted for about 40 minutes, amounting to 20 minutes for each subteam at the hot table, and about 4 turns during that time (two per subteam).

For Showdowns using the in-person debate, students at the “hot table” were asked to volunteer to go into the “hot seat”. Both teams were in the same room and each active subteam put a member into the “hot seat” where they debated in person for two minutes at which time another member of the team took their place. Each subteam was allowed to call a one-minute

“huddle” to confer if it wished. On average about four people from each subteam took a turn in the “hot seat.”

A whole-class debrief occurred in the session immediately following the Showdown. For Showdowns done electronically, students were guided through an argument map, which was a model of their dialog transcript with points awarded for each effective counterargument, and points taken away for unwarranted assumptions and failure to respond to opponents arguments (Appendix G). Points were summed and a winner was declared.

For Showdowns done in person, students watched the video of the Showdown, and an “expert judge from Columbia” then led a discussion with the students about the strengths and weaknesses she saw in the debate, focusing on the same criteria identified above, and declared a winner.

#### Phase Four

The fourth phase in Y1 introduced a new aspect of argumentation, evidence, which became a focus of the Y2 and Y3 curriculum. To inform their arguments, students were presented with a set of nine pieces of evidence relevant to the topic (Kidney Sales) in the form of questions with answers. Four of the pieces of evidence supported the “yes” position, four supported the “no” position, and one was neutral (Appendix H). Students “checked out” a question on a sign-out sheet, reviewed the answer, and then returned the question to the coach. Students were told they could submit any additional questions they would like to have answered and the coaches would help answer them. None did so.

Also beginning in phase four, students were required to write an individual persuasive essay as homework. Students were instructed to “Please write a persuasive essay (just as you

would write in your ELA class) arguing for the position you currently believe with respect to [TOPIC]. You do not need to take the side you have taken during class activities.” The evidence questions and answers were available to the student for reference.

**Table 4: Year One Session Activities by Group**

Session #	Activities	
	Electronic Group	Team Group
1	Dialog 1: No Reflection Sheet	Generate reasons
2	Dialog 2: “Other” Reflection Sheet	Eliminate duplicates
3	Dialog 3: “Own” Reflection Sheet	Elaborate/improve reasons
4	Dialog 4: “Other” Reflection Sheet	Evaluate reasons
5	Dialog 5: “Own” Reflection Sheet	Develop reasons into an argument
6	Dialog 6: “Other” Reflection Sheet	Examine and evaluate opposing side’s reasons
7	Dialog 7: “Own” Reflection Sheet	Generate counterarguments to others’ reasons
8	Showdown prep 1	Generate rebuttals to others’ counterarguments
9	Showdown prep 2	Trial two-sided arguments
10	Showdown	
11	Showdown debrief & announcement of winner	

### *Year Two Procedure*

Based on the Y1 experience, a hybrid procedure was adopted for Y2 and Y3, incorporating both small group work and electronic dialogs with opposing pairs as both appeared

valuable and groups did not differ with respect to individual argumentation skill level at the end of Y1.

### The Pre-game

Working in teams of 8 same-side students, students engaged in three small-group activities (each occupying one session) during the first three sessions of each phase.

*Session 1.* During the “Our Reasons” activity, students generated and recorded the reasons they had for taking the position they did.

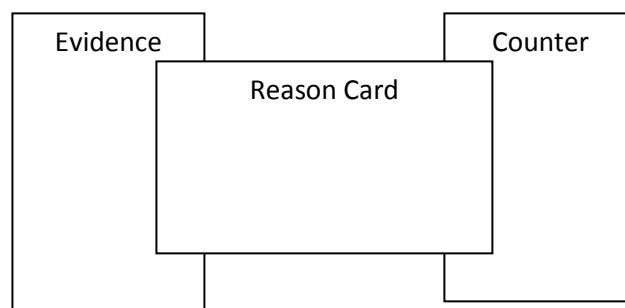
*Session 2.* In the “Evidence /Evaluation” activity students focused on two goals: first, to consider relevant evidence that might help to support their reasons and second, to rank their reasons by relative worth and effectiveness. During this activity students were presented with individual pieces of evidence (as described in *Year One Procedure*, above) each consisting of a question and brief answer. They were also encouraged to submit their own questions which, when answered, became additional evidence. By the fourth phase of Y2, students generated all questions in the evidence set; that is, the researchers did not generate any initial questions, they merely provided answers to questions asked by students (see Appendix H for a complete list of questions provided and answered for all topics).

*Session 3.* The third activity focused on a) anticipating what the other side’s reasons and evidence might be, b) anticipating how one’s own team would counter those reasons, and c) how they might rebut those counters.

These activities were facilitated by the use of super cards. Own-reason super cards (see figure 2) were created systematically throughout the small-group sessions, starting with just the

middle reason card (during the “Our Reasons” activity), adding the left/evidence card in the second session (the “Evidence/Evaluate” activity) and then adding the right/counter card at session three (“Their Reasons and Evidence”). Others-reason super cards were similarly constructed.

**figure 2: Super Card**



Note. Rebuttals were written on the back of the reason card.

### The Game

During the six sessions of the “Game” stage, students engaged in electronic dialogs, as described in *Year One Procedure*. Students could sign out pieces of evidence for reference during dialogs, as they saw fit.

### The End Game

*Showdown Preparation.* As described in *Year One Procedure*, each phase culminated in a Showdown. To prepare for the Showdown, students returned to the small groups formed in the pre-game for two class sessions; the first focused on the team’s own reasons, the second focused on the other side’s reasons.

The first class session focused on a team's own arguments. Students were instructed to fill out new super cards for each of their arguments (figure 2). Students were asked to generate as many full sequences as they could, starting with what they considered to be their strongest argument.

The second class session focused on the other side's argument. Students were instructed to fill out new super cards for each of their opponents' arguments (figure 2). Students were asked to generate as many full sequences as they could, starting with what they considered to be their opponents' strongest argument.

*The Showdown and Debrief.* The Showdown followed the procedure for the in-person debate described in *Year One Procedure*. Students were given an additional minute per turn in the hot-seat (for a total of three minutes per turn) starting in the third phase of the second year; otherwise all procedures remained the same.

During the Showdown debrief students watched a video of the Showdown and were guided through a portion of an argument map based on a transcript of the Showdown. In addition to the scoring described in *Year One Procedure*, points were awarded for using evidence in support of a reason, and points were taken away for using evidence incorrectly (Appendix I). Students were presented with a personal copy of the argument map and a full set of evidence to use when preparing their individual topic essays (topic essays are described in *Year One Procedure*).

### *Year Three Procedure*

The Y3 procedure was identical to the Y2 procedure with the following additions.

### *Student and Group pairings*

In previous years each topic offered a choice of two positions for students to take, typically for or against a particular idea or practice. In Y3, three positions were identified. For example, phase nine (focused on the juvenile justice system) identified the alternatives of teens being tried as adults, teens being tried as minors, or teens tried by a jury of their peers. This addition fostered a deeper understanding of the complexity of social issues, and provided experience identifying one's own position relative to two other positions. That is, students discovered that a position and relevant arguments might be framed differently depending on the position being argued against.

### *The Game*

In Y1 and Y2, students were asked to evaluate their efforts via an in-class reflection sheet (described in *Year One Procedure*, Appendix E). In Y3, student reflection took place outside of class, as homework. Students were asked to print a copy of their electronic dialogs, evaluate the argumentative quality of each statement made by either side by assigning a score ( a check, a check minus, or a check plus) and then providing a written rationale for that score. This addition allowed for more detailed and individual reflection, and freed class time for longer dialogs and activities.

After four of the electronic dialogs, students engaged in whole-class activities: the mini-showdown, the mini-showdown debrief and two sessions of the argument game. The mini-showdown (occurring after the third dialog) involved two student pairs (encompassing all three positions) engaging in a face-to-face debate similar to the Showdown. Turns were three minutes long, as they were in the Showdown, but huddles were not allowed. The mini-showdown debrief

(occurring after the final dialog) presented an un-scored transcript of the mini-showdown. Statements were evaluated through class discussion. The mini-showdown and mini-showdown debrief served to motivate and focus students for the Showdown preparation sessions beginning in the following session.

The argument game (occurring twice per topic) focused on developing the skill of argument evaluation. A script was created around a previously debated topic (with the exception of two “special editions:” a holiday edition and a springtime edition). The script provided a statement and two counterarguments; one direct counter and one indirect counter, as defined earlier (See Appendix J for a sample argument game script). Coaches acted as debaters and as the emcee. Students were provided with voting cards pre-marked with the initials of the coaches acting as debaters. The emcee read the first statement, and then each of the debater coaches read their response to the statement. Students then voted for which statement they felt was the stronger argument and students voting for each option were asked to explain their reasoning. The emcee then identified the correct answer and further discussion followed regarding its superiority (See Table 5 for an example [superior answer bolded], and Appendix J for Homeschool script).

**Table 5: Example of the Argument game**

MC: Aris only speaks Greek, so he should go to the town school so he'll learn English.
<b>Coach 1: Aris can learn English by living here and being exposed to people in the grocery store, for example.</b>
Coach 2: It will be easier for Aris to learn difficult academic subjects if he learns them in his own language.



Additionally, students were also asked to volunteer to contribute to the evidence set by assisting in researching questions posed by classmates. At each session, four students were chosen from volunteers to do research at home on a submitted question. These students returned to class and met with a coach to report their findings and finalize an answer that was then added to the class set.

As an additional element during the final topic of Y3, some work was done across the two intervention classes. Half of the electronic dialogs spanned classes; that is, for half of the electronic dialogs a particular pair would argue with pairs from the other intervention class. The Showdown was staged across classes, and the Showdown debrief included a showing of an excerpt of the cross-class video, discussion of the cross-class transcript and declaration of winners both between and within classes.

### *The Showdown and Debrief*

The debrief sessions in Y3 did not include a viewing of the full Showdown video (with the exception of the final topic, discussed above). This was done to free class time for a more detailed discussion of the argument map.

### *Comparison Group Curriculum*

The comparison group underwent all assessments described for the intervention group. They met for the same number of sessions over the same three-year time period. The class was taught by teachers from the school's staff, rather than our research team. This class covered a larger number of topics than the intervention classes, but all topics involved a philosophical or social issue, some of them the same ones addressed by the experimental classes. Students engaged in teacher-led, whole-class discussion of the issue, sometimes with additional activities

such as dramatizations. Students were assigned individual essays on one of the topics as homework at least once every other week. Hence, they received more practice in writing argumentative essays than did the intervention group.

An uncontrollable methodological limitation of the study arose during Y2 from the fact that one of the school staff who taught one half (16 students initially) of the comparison group observed our method, decided she liked it, and adopted aspects of it in her class. Thus, our experimental comparison was compromised, a fact we return to in examining the results.

## Results

### *Transcript Coding*

The analysis was based on 182 transcripts, 46 pretests (Y0), 47 from the end of year one (Y1), 46 from the end of year two (Y2) and 43 from the end of year three (Y3). Coding was based on the functional coding scheme used in previous research (Appendix K, Felton & Kuhn, 2001; Kuhn & Udell, 2003; Felton, 2004; Udell, 2007; Kuhn, Goh, Iordanou & Shaenfield, 2008). Each utterance (comprising one turn in the dialog) was segmented into idea units and each idea unit was assigned a single code from the scheme (Appendix J, see Table 6 for proportional use of each utilized code accounting for at least 1% of total codes). Codes are functional in nature; that is, each code represents the functional relation between the utterance being coded and the immediately preceding utterance. Two coders, blind to condition and time, coded a randomly chosen 20% of Y0 and Y1 transcripts and achieved 81% agreement (Cohen's  $Kappa = .762$ ). Two coders, blind to condition and time, coded a randomly chosen 20% of Y2 and Y3 transcripts and achieved 84% agreement (Cohen's  $Kappa = .757$ ). All remaining transcripts were coded by one coder.

**Table 6: Frequencies and Proportion Use for All Codes Accounting for at Least 1% of Total Codes**

Code	Y0Count	Y1Count	Y2Count	Y3Count	Proportion of Total Codes
CC	125	222	356	330	0.38
CA	84	194	122	139	0.20
Clarify	101	122	75	116	0.15
Meta	34	24	50	20	0.05
CaseQ	28	33	18	25	0.04
ClarifyQ	6	20	38	27	0.03
Agree	34	24	18	14	0.03
Disagree	15	19	16	8	0.02
JustifyQ	3	10	4	12	0.01
Null	3	4	12	7	0.01
Interpret	0	7	4	6	0.01
Add	3	0	4	8	0.01

*Note.* CC = Counter-critique; CA = Counter-alternative

### ***Counterargument Use***

The first research question asks whether dense practice with dialogic argumentation changes students' use of counterargument strategies over time. Counterarguments take two forms: the indirect counter, coded as counter-alternative and the direct counter, coded as the counter-critique (See Table 7 for verbatim examples). The counter-alternative strategy responds to an opponent's argument with which the student disagrees by providing an argument in support of his or her own position. The counter-critique strategy responds to an opponent's argument with which the student disagrees by seeking to weaken the opponent's argument. Both forms, however, are indicative of argument skill.

**Table 7: Verbatim examples of counter-alternative and counter-critique**

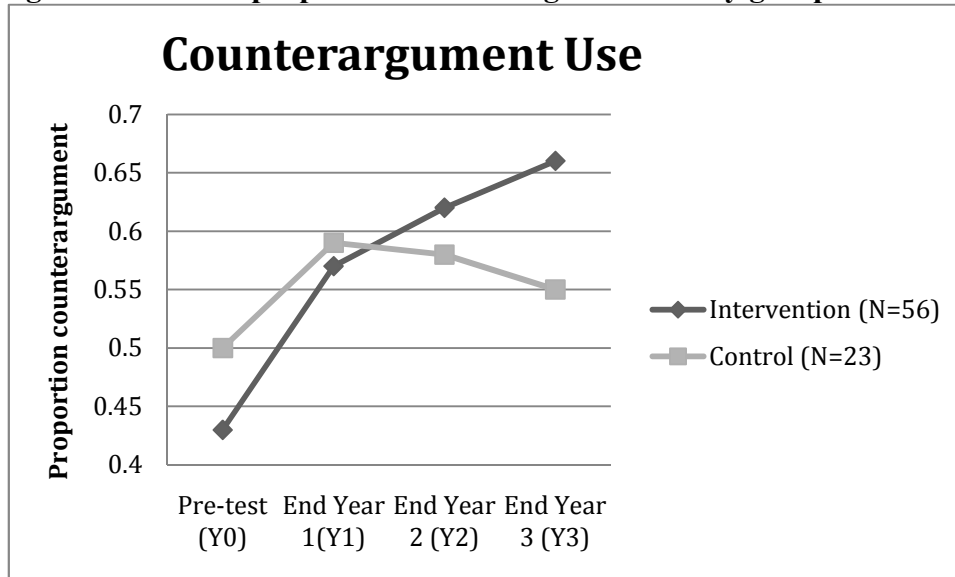
Person	Argument	Code
A	I just believe that a person who has committed a serious crime/killed someone does not deserve to live because they took someone else's life.	
B	There are other options for them, like a life term in jail.	Counter-Alternative
A	People can get out of jail and then they could kill again. It's not worth the risk.	Counter-Critique

Proportion use of counterarguments was calculated for each participant at each assessment time. Means across participants appear in Table 8 and figure 3. After an arcsine transformation normalized the proportions, a 2 by 4 (Group \* Time) repeated measure ANOVA showed a significant interaction between time and group  $F(3, 75)=3.41, p=.02$ , partial eta squared=.12. There is not enough statistical evidence to determine exactly where these differences originate, although the gains made in the last year (Y2 vs. Y3) is an almost moderately significant contributor ( $p=.14$ ) and the intervention group exceeded the comparison group at the final time point,  $t(77)=2.31, p=.02, g=.57$  (see Appendix L for further details of year by year change).

**Table 8: Mean of proportion counterargument use by group and time (SD)**

Group	Pre-test (Y0)	End Year 1 (Y1)	End Year 2 (Y2)	End Year 3 (Y3)
Intervention (N=56)	.43 (.25)	.57 (.19)	.62(.17)	.66 (.18)
Comparison (N=23)	.50(.16)	.59(.17)	.58(.23)	.55(.19)

**figure 3: Mean of proportion counterargument use by group and time**

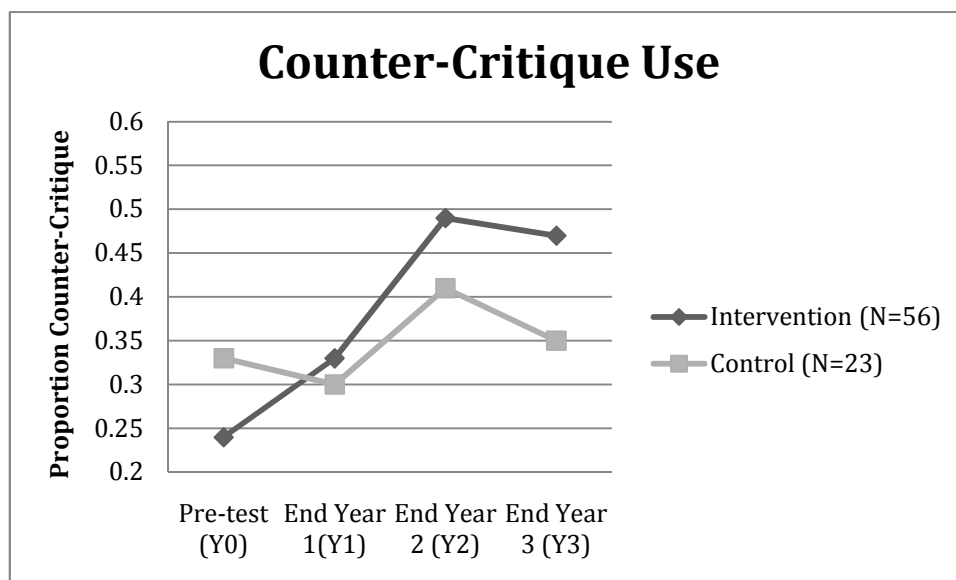


We were also interested in how use of the more effective and powerful counterargument, counter-critique, changed over time. A 2 by 4 (Group \* Time) repeated measure ANOVA of proportion counter-critique use showed a significant interaction between time and group  $F(3, 75)=3.57$ ,  $sig=.02$ , partial eta squared=.13. There is not enough statistical evidence to determine exactly where these differences originate, although the gains made in the first year (Y0 vs. Y1) were an almost significant contributor ( $p=.05$ ) and the intervention group exceeded the comparison group at the final time point,  $t(77)=2.68$ ,  $p=.02$ ,  $g=.57$  (see Appendix L for further details of year by year change) See Table 9 and figure 4 for means by group and time.

**Table 9: Mean of proportion counter-critique use by group and time (SD)**

Group	Pre-test (Y0)	End Year 1 (Y1)	End Year 2 (Y2)	End Year 3 (Y3)
Intervention (N=56)	.24 (.21)	.33 (.18)	.49(.18)	.47 (.17)
Comparison (N=23)	.33 (.25)	.30(.20)	.41(.22)	.35(.23)

**figure 4: Mean of proportion counter-critique use by group and time**



### ***Sustained Direct Counterargument***

The second research question, “Do students in the intervention group sustain direct counterargument with an opponent for longer than students in a comparison group?” is answered by analysis of rebuttal sequences. When a student responds to a counterargument with a counter-critique, he or she is said to rebut the previous counterargument. Rebuttal sequences reflect opponents’ arguing a particular point until one or the other changes the focus of the argument.

A measure of sustained direct engagement is the duration (how many uninterrupted utterances are involved) of the rebuttal sequence. In order to rebut an opponent’s argument one must be presented with an opportunity to do so; not all students were presented with such an opportunity. Only 69 students were offered the opportunity to engage in rebuttal at all time points. A repeated measure test across times would only include students who were offered this opportunity at all times and was not run since results based on this reduced sample would be difficult to interpret. All students were offered the opportunity to engage in rebuttal at the final two time points and the intervention group sustained engagement for longer than the comparison

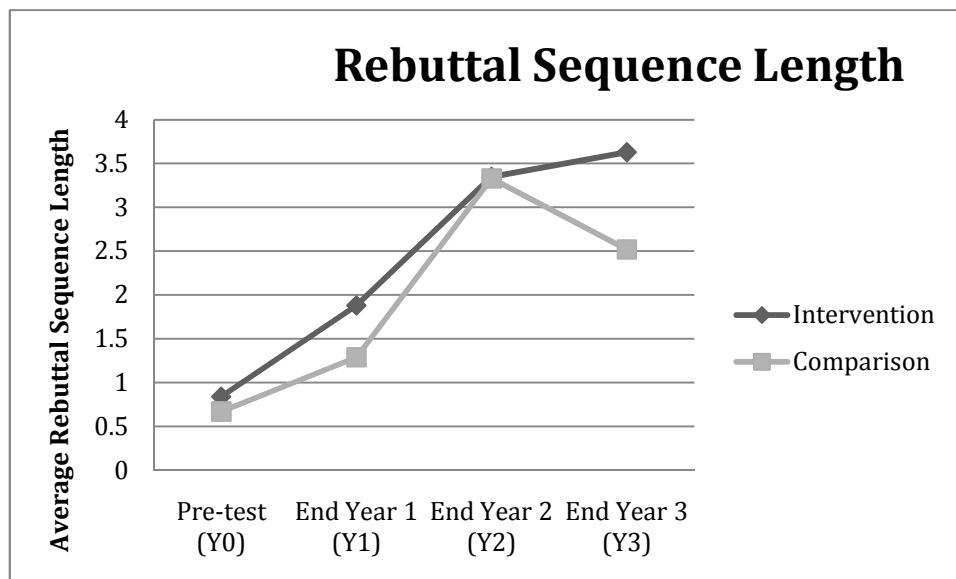
group at the final time point  $t(77)=2.38, p=.04, g=.59$ . See Table 10 and figure 5 for reduced sample means by group and time.

**Table 10: Mean Rebuttal Sequence Length (Reduced N in parentheses)**

Group	Pre-test (Y0)	End Year 1(Y1)	End Year 2 (Y2)	End Year 3 (Y3)
Intervention	.84 (49)	1.88 (49)	3.35 (56)	3.63 (56)
Comparison	.67 (21)	1.29 (21)	3.33 (23)	2.52 (23)

Note. Entries are mean number of turns.

**figure 5: Mean of Rebuttal Sequence Length by group and time**





### *Argument Evaluation*

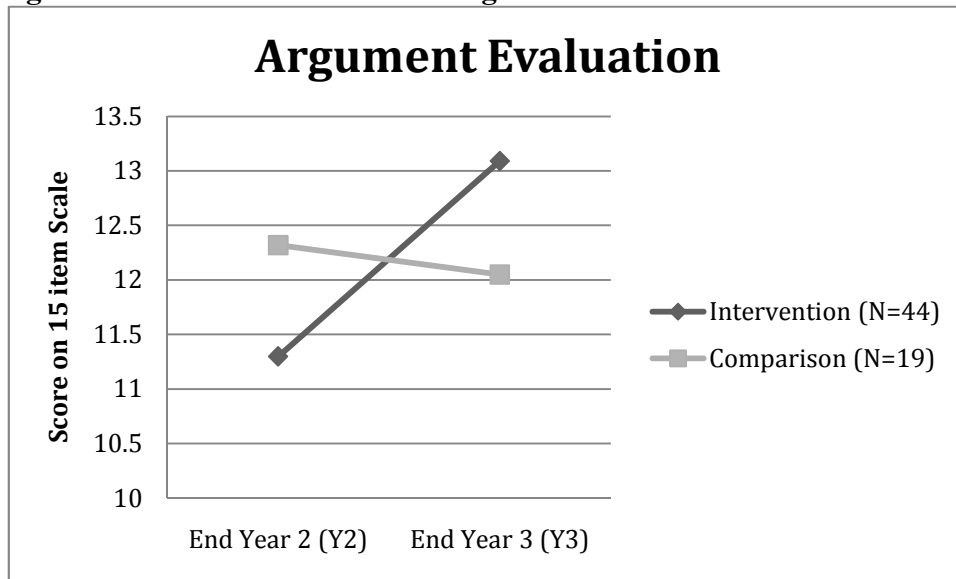
The third research question, “Do students in the intervention group become more skillful in the evaluation of arguments relative to students in the comparison group?” was assessed via an online argument evaluation assessment. For every item on the adaptive assessment on which students chose the direct counterargument over the indirect counterargument they received one point. Analysis was performed on the 15-item assessment presented at Y2 and Y3 (results for the 9-item assessment were equivalent, see Appendix M).

A 2 by 2 (Group \* Time) repeated measure ANOVA showed a significant interaction between time and group  $F(1, 61)=7.746$ ,  $sig=.007$ , partial eta squared=.11. Group difference was not significant at Y2 ( $p=.13$ ) but the intervention group outperformed the comparison group at the final time point,  $t(61)=2.045$ ,  $p=.04$ ,  $g=1.77$ . See Table 11 and figure 6 for means by group and time.

**Table 11: Mean of Score on 15 item test (SD)**

Group	End Year 2 (Y2)	End Year 3 (Y3)
Intervention (N=44)	11.30 (2.60)	13.09 (1.63)
Comparison (N=19)	12.32(1.91)	12.05 (2.30)

figure 6: Mean Score on 15 item argumexnt evaluation assessment



## Discussion

The findings of the present study indicate that argument skills can be effectively developed through dense engagement in dialogic argumentation for an extended period of time. On all measures the intervention group outperformed the comparison group at the final time point, indicating that at the end of the three-year intervention, those who participated exhibited better developed argument skills than those who did not.

***Research Question: How does dense and extended practice in dialogic argumentation influence counterargument skill?***

Counterargument skill, when engaged consistently, develops incrementally across the extended course of middle school. Students in the intervention group increased their likelihood to engage in counterargumentation across the three years. The comparison group, on the other hand, did not show a consistent pattern; they showed an increase in year one and then decreased across years two and three. Overall, between initial and final assessments, both groups increased their propensity to engage in counterargument strategy, but students who engaged in dialogic argumentation show a much greater increase -- 24% from the initial assessment to the final assessment, compared to a 5% increase from initial assessment to final assessment for the comparison group. Comparative year-to-year change is too incremental to point conclusively to a particular year as the source of the differences, providing support for the notion that change occurs gradually across the three years. The degree of improvement, however, decreases as time progresses; there is a 14% improvement between Y0 and Y1 for the experimental group, a 5% improvement between Y1 and Y2, and a 4% improvement between Y2 and Y3. This raises a question of diminishing returns, a question to which we return.

When examined by counterargument subtype, we see that students who engaged in

dialogic argumentation developed a greater propensity to utilize the more powerful direct counterargument strategy than did those in the comparison group by the end of the intervention. From initial assessment to the final assessment students who engaged in dialogic argumentation gained 23% likelihood to engage in direct counterargument, compared to a 2% increase from initial assessment to final assessment for the comparison group. Again, comparative year-to-year change is too incremental to identify the specific source of these differences (although the change from initial assessment to the end of the first year is almost a statistically significant contributor;  $p=.054$ ). In this case, however, experimental group gains are realized by the end of the second year and decrease marginally in the final year (a 2% reduction between Y2 and Y3), again raising the question of the value of the final year of the curriculum.

### ***The Ceiling Effect***

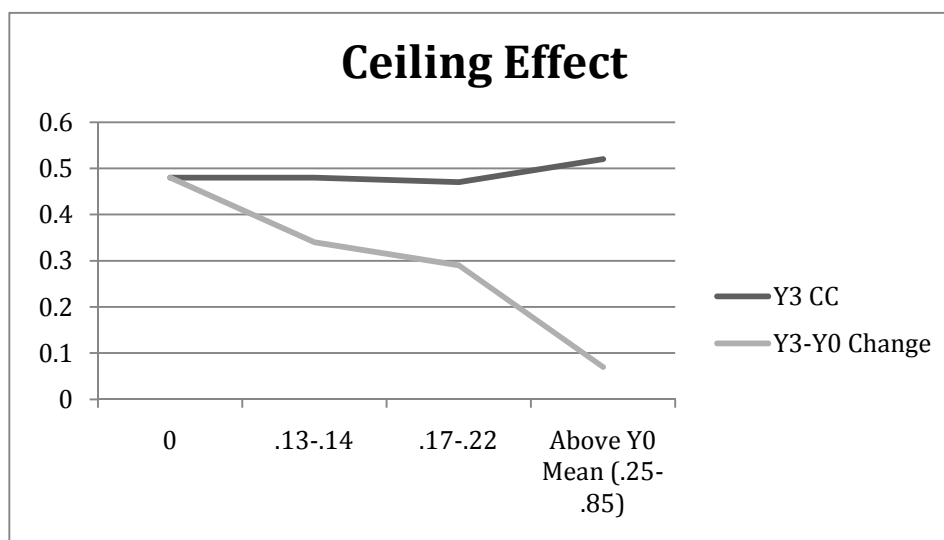
There is evidence that improvement in usage of the direct counterargument strategy is subject to a ceiling effect. As mentioned, the intervention group improved roughly 23% between initial and final assessment, but when the group is divided into those who began the intervention below the Y0 mean and those who began the intervention above the Y0 mean, different patterns emerge. Those who began the intervention above the group mean improved an average of 7% between initial and final assessment. However, those who began the assessment with no competency with direct counterargument (0% usage at Y0) gained 48%, those who began with minimal competency (13-14% usage at Y0) gained 34%, and those who began with moderate competency (17-22% usage at Y0) gained 29%. All four groups performed around the same level at the final assessment, indicating that for this age group competent usage of the direct counterargument strategy may reach of ceiling of around 50% (See Table 12 and figure 7 for details by these Y0 subgroup). This provides some insight into the slight decline observed for the

intervention group in year three; the group mean was at ceiling in year two and so, while those who began the intervention with less competence were still improving, the group mean hovered in the space around the ceiling at the assessment in year three.

**Table 12: Ceiling Effect**

Pretest Proportion Use of Direct Counter	Y3 Proportion Use of Direct Counter	Y3-Y0 (Change from Initial to Final)
Y0= 0%	0.48	0.48
Y0= .13-.14	0.48	0.34
Y0=.17-.22	0.47	0.29
Y0 Above Mean (.25-.85)	0.52	0.07

**Figure 7: Ceiling Effect**



Thus, the trajectories of those who began the intervention at the lower end of the spectrum versus those who begin higher on the spectrum is different; those who began the intervention with less skill improved more than those with more initial skill. Table 13 shows the number of students in each of the above referenced subgroups whose proportion use of the direct counter strategy increased, decreased, or remained the same over each of the three years. Eighty-five percent of students in the group that began the intervention with no exhibited competency

made improvements in the first year, compared with only 25% of students in the group that began the intervention above the mean. In year two, 77% the students who began with no competency improved compared with 67% of those who began above the mean. In year three, the trend continues with 62% of those with no initial competency improving into the third year while only 42% of those who began above the mean improving in the third year. Thus, across all three years students who began the intervention with no exhibited competency were more likely to show improvement in usage of the direct counterargument skill than were those who began the intervention above the mean. This is further evidence of the presence of a ceiling effect.

**Table 13: Proportion Usage of Direct Counter Change Across Years and Subgroups**

Range of Y0 Direct Counter Proportion Usage	Y0->Y1Change				Y1->Y2 Change				Y2->Y3 Change				Average Y0 to Y3 Change in proportion on direct counter usage.										
	#Up	%Up	#Dn	%Dn	#No	#Up	%Up	#Dn	%Dn	#No	#Up	%Up		#Dn	%Dn	#No							
Intervention 0-0	13	11	84.62%	0	0.00%	2	15.38%	0	0.00%	0	10	76.92%	3	23.08%	0	0.00%	8	61.54%	5	38.46%	0	0.00%	0.48
Intervention .13-.14	7	4	57.14%	2	28.57%	1	14.29%	0	0.00%	0	7	100.00%	0	0.00%	0	0.00%	2	28.57%	5	71.43%	0	0.00%	0.34
Intervention .17-.22	11	9	81.82%	0	0.00%	2	18.18%	0	0.00%	1	6	54.55%	4	36.36%	1	9.09%	5	45.45%	4	36.36%	2	18.18%	0.3
Intervention .25-.85	24	6	25.00%	15	62.50%	3	12.50%	0	0.00%	0	16	66.67%	8	33.33%	0	0.00%	10	41.67%	14	58.33%	0	0.00%	0.07
Comparison 0-.25	12	6	50.00%	0	0.00%	6	50.00%	0	0.00%	1	8	66.67%	3	25.00%	1	8.33%	3	25.00%	9	75.00%	0	0.00%	0.26
Comparison .34-.85	12	3	25.00%	7	58.33%	1	16.67%	0	0.00%	0	8	66.67%	4	33.33%	0	0.00%	4	33.33%	6	50.00%	1	16.67%	-0.05

Note. #Up means number of participants whose usage of the direct counter strategy increased that year. %Up means the percent of those whose usage of the direct counter strategy increased that year over the total N of that subgroup.

#Dn means number of participants whose usage of the direct counter strategy decreased that year. %Dn means the percent of those whose usage of the direct counter strategy decreased that year over the total N of that subgroup.

#No means number of participants whose usage of the direct counter strategy did not change that year. %No means the percent of those whose usage of the direct counter strategy did not change that year over the total N of that subgroup.

Usage of the indirect counterargument skill remains stable for students in the intervention group (19% at initial assessment, and 19% at final assessment, and within 5% points of this at the other assessment points, see Appendix N for more details), indicating that dialogic argumentation not only increases the propensity to engage in counterargument strategies broadly defined, but does so by increasing direct counterargument, not indirect counterargument.

Taken together, these trends show different pathways for change in counterargument strategy use for students who engage in dialogic argumentation and those who do not. Both counterargument broadly defined and direct counterargument specifically are better developed at the final time point for students in the intervention, showing that substantial improvements are generated by engagement in dialogic argumentation. Thus, students who engage in a dialogic argumentation curriculum like the one investigated here exhibit more fully developed argumentation skills than do students whose skill has developed via maturation and more general educational experience (even of the comparatively high quality provided at this school, with its emphasis on cognitive engagement and collaborative activity over rote learning).

### *Qualitative Case Study*

An example will help to elaborate how these differences play out in the assessment dialogs. Students 32 and 62 were paired together for all time points; student 32 is female and for capital punishment (CP) and student 62 is male and against CP. At each of the four time points these students had discourse about the necessity (or lack thereof) of second chances. Excerpts regarding this point help to illustrate how their treatment of the issue and their statements of position develop over time (see Appendix O for the complete dialogs between these two students).



Initial Assessment (Y0)

Test	Person	Utterance Number	Utterance
Y0	32	1	I am in favor of CP because if the murderer (if he is one) doesn't get killed they will do another crime so they should be punished.
Y0	62	2	The murderer should go to jail because <b>if he makes up for himself he can be free</b> . Killing would just cause problems.
Y0	32	3	So he should be punished for what he did because you wouldn't like it if the murderer killed your parents or something. You would want him to be killed.

Student 32 begins by making the claim that prisoners return to crime when released and so should be executed. Student 62 does not engage this idea; rather he makes the case for second chances (if he makes up for himself, he can be free). Rather than engaging what her partner says student 32 returns to her (rather vague) claim that the murderer should be punished and then provides a case with which she believes her partner must agree. Neither student truly engaged or grappled with the points made by their conversational partner.

End of Year One (Y1)

Test	Person	Utterance Number	Utterance
Y1	32	1	I am for CP because if the person committed murder that is a really big deal, then the person should get punished for their actions.
Y1	62	2	Ok, but sentencing them to death isn't a good idea. <b>You should give them another chance</b> , and violence doesn't really solve anything.
Y1	32	3	Well, first of all <b>why should they get another chance? The person that got murdered didn't get a chance at all.</b>
Y1	62	4	But murdering the murderer is just going to create problems, because then maybe their family will go after the police because of that. Also, even if the person is murdered they could just send the murderer to jail for a long period of time.

By the end of year one we see more of a willingness to engage the idea of the conversational partner, if only briefly. Student 32 engages the idea that prisoners deserve a second chance by stating that the prisoner did not give his or her victim a second chance. The

students move beyond the “second chance” idea quickly but goes on engage other ideas for two or three utterances (see appendix O) showing a good deal of progress towards really grappling with each others ideas.

End of Year Two (Y2)

Test	Person	Utterance Number	Utterance
Y2	32	1	Capital punishment should be allowed because if a person commits murder, or anything else that is very serious, then they should pay for what they did.
Y2	62	2	Yes, and they should pay by serving for a long period of time in jail and not being put to death.
Y2	32	3	<b>Why would you want to give them a second chance?</b> They didn't give the person that they killed (murdered) a chance. They took away a life. They don't deserve a long period of time in jail
Y2	62	4	Yes they do, because <b>what if the criminal understands that they did something wrong, and wants to live a better life? They need to get the second chance to do that!</b>
Y2	32	5	<b>It's too late for them to think that what they did was wrong.</b> They should've thought of that before they had killed someone. The person that was killed did not deserve to die because they didn't do anything. But the murderer does since they took a life away as well as doing something wrong. Imagine it was a close relative of yours, like your mom or dad, who was killed. Would you want the murderer to stay alive without knowing that there are consequences to their actions?
Y2	62	6	Three things first of all, you gave me no space [on this paper], second of all, who says that the person who was murdered didn't do anything wrong? <b>And anyways, imagine you're the criminal. After being sentenced to jail you would probably have realized your mistake and want a second chance.</b>
Y2	32	7	If the person [the victim] did do something wrong, then wouldn't the person [the murderer] just go to the police to get them instead of killing them and getting into trouble? That would've been something better to do. <b>Also, in the beginning you said that the murderer should serve jail time for a long period of time. Many other crimes can make you go to jail for a long time (i.e. that guy that is serving life in prison for tricking people for their money). It isn't the same thing as killing a person. That guy that is serving jail time for that long did not take away a life and that guy is learning his lesson in a way that is appropriate (serving jail time).</b>

Y2	62	8	Okay, well, in the argument you just stated, what if something the person killed (the victim) did was kill another person? That then puts the person killed in the same situation as the murderer. <b>They all deserve a second chance.</b>
----	----	---	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

This dialog looks different on a number of counts. The “second chance” idea is grappled with and returned to across the entirety of the argument rather than being summarily dismissed or quickly engaged and then disregarded. There is also evidence that students are attempting to engage many of the ideas that their conversational partner posits, resulting in longer and more complex utterances. We also see the students pointing back to earlier places in the dialog and directly countering those ideas while maintaining the current thread of the conversation. In essence, these students are really engaging with what their conversational partner means and attempting to draw connections and distinctions between the ideas posited by themselves and their partner.

*End of Year Three (Y3)*

Test	Person	Utterance Number	Utterance
6	32	1	I am for Capital punishment (death penalty) because if a person has been guilty of committing murder, then they should pay for what they've done since the crime was very serious.
6	62	2	Exactly! They should pay by spending time in prison.
6	32	3	What is time in prison going to do? Either way how is time in prison going to pay for what happened to the person that got murdered?
6	62	4	Time in prison can be spent doing courses of rehabilitation. Time will help the criminal pay because the victim has already been murdered and there is no reason to take another person's life away. In addition, you don't know the incentive of the murderer. Someone doing self-defense shouldn't be killed.
6	32	5	Then the crime would be different. We don't know what the court has chosen the case to be. It could have been a 1st degree murder, 2nd degree murder, manslaughter, or self-defense. But if the criminal is guilty of 1st or 2nd degree murder, their intentions

			were to kill the person, which shouldn't be forgivable at all.
6	62	6	Not true. <b>Everybody needs a second chance.</b> Long periods of time in prison can help the murderer. The murderer would obviously need help if it was 1st or 2nd degree murder.
6	32	7	Well, <b>did the criminal give the person that is now dead a second chance?</b> Did they stop themselves before killing the person? No, no they wouldn't.
6	62	8	Yes, <b>that was the murderer's mistake in not giving the person a second chance.</b> We shouldn't make the same mistake by killing them in return. <b>Placing him/her in prison will probably help them to stop themselves from committing another murder.</b>
6	32	9	<b>If they actually get the help, how are we sure that it has actually helped them?</b> What if they're released from jail and end up killing some other person? That wouldn't be helpful in any way! <b>And what gives them the right to be forgiven for taking someone's life away? I don't think you would like it so much if a person murdered someone very close to you...</b>
6	62	10	I wouldn't. If they killed again, then prison for life. If a person killed another person by accident, it wouldn't make any sense to kill them in return.
6	32	11	We don't even know if the crime was an accident or not. <b>And you're telling me that you would be totally fine with the idea that the person who murdered your close relative is still given a second chance?</b>
6	62	12	No, but it would be the right choice. <b>If I murdered I would want a second chance.</b>
6	32	13	<b>You would want a second chance, but does that necessarily mean that you deserve that second chance?</b> Also, was it the right choice to even murder someone in the first place?

Here we see further specificity and skill with direct counterargumentation. When her conversational partner attempts to sidestep an argument, student 32 comes back and reiterates the claim, asking again for her partner to deal with it. In line 9 student 32 states that student 62 would not “like it so much if a person murdered someone very close to you.” Student 62 quickly blows this off, “I wouldn’t...” but rather than allowing her point to be disregarded she reiterates “you're telling me that you would be totally fine with the idea that the person who murdered your

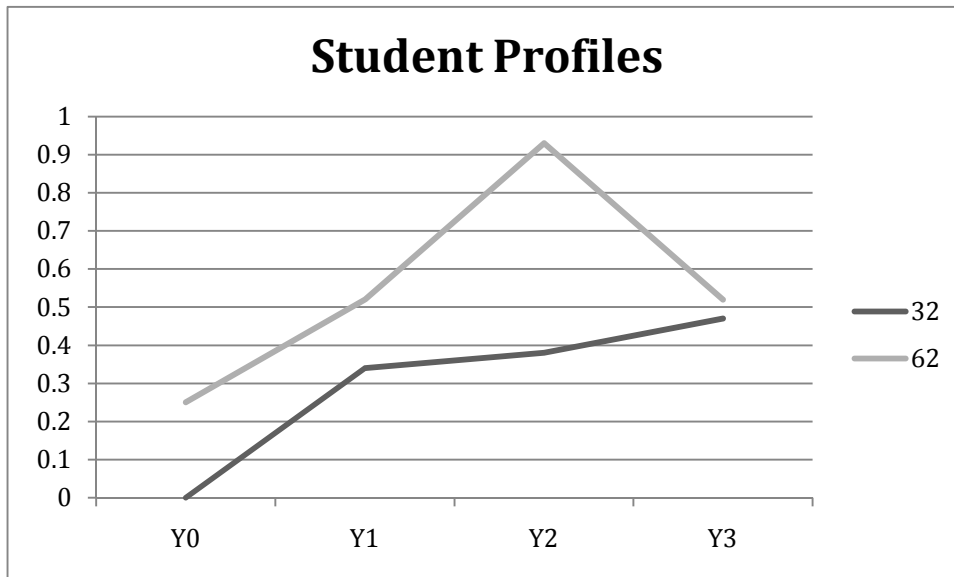
close relative is still given a second chance?” Student 62 is then forced to respond to the meat of her claim, stating “No, but it would be the right choice. If I murdered I would want a second chance” which allows student 32 to directly counter the prevailing idea: “You would want a second chance, but does that necessarily mean that you deserve that second chance?” This, coupled with the keen use of evidence in line 5 shows considerable improvement from the earlier arguments. The arguments of both partners are more nuanced, specific, and engaged than they were at all of the other time points.

A comparison of our case study students to the students in the intervention group shows them to be relatively representative of their peers when we consider them as representative of two contrasting subgroups: those students who began the intervention above the mean and those who began the intervention below the mean (see figure 8 for the case study students trajectories and figure 9 for information on the intervention group split into these groups). The students in our case study provide an interesting example of how a student who began the intervention below the mean (student 32, with 0% competency compared to the combined group mean of .24% at Y0) engaged with a student who began the intervention just above the mean (student 62, with 25% competency at Y0).

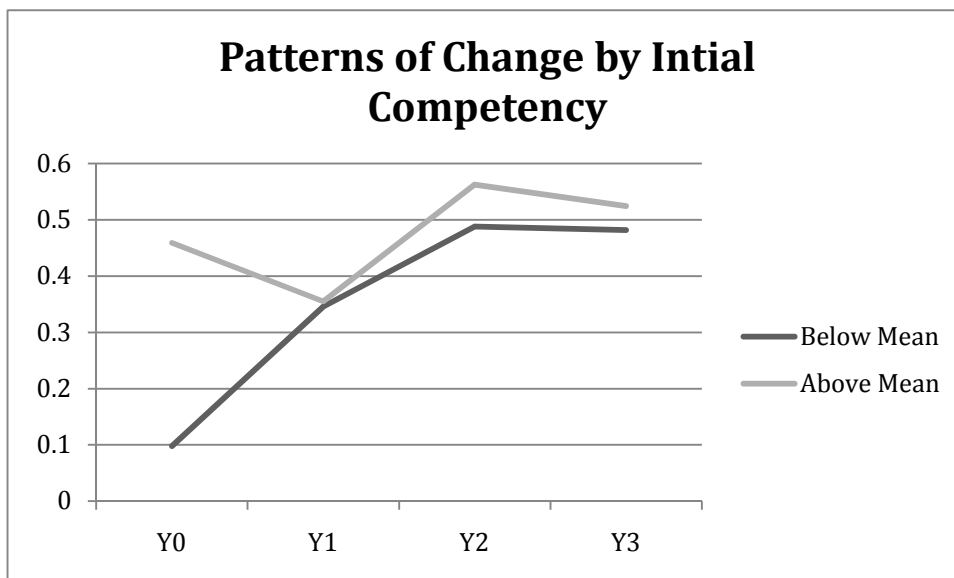
Our lower performing student has a very similar trajectory to those students who began the intervention below the mean; she makes her greatest gains in year one, continues to improve in year two, and ends the intervention at the hypothesized ceiling. Our higher performing student follows the trajectory of his sub group with the exception of the first year. While the higher performing subgroup declines in the first year, this student makes considerable gains, which is not surprising since our case study student is only barely above the mean at Y0 (.25 vs. the .24 combined group mean). His performance beyond Y1 is very similar to that of his

subgroup peers, however, with a peak above the ceiling in year two and a decline back towards the ceiling in year 3.

**figure 8: Case Study Student Trajectories**



**figure 9: Split Intervention Group Trajectories**



***Research Question: Do students in the intervention group sustain direct counterargument with an opponent for longer than students in a comparison group?***

Students who engaged in the three-year curriculum sustained direct counterargument for significantly longer at the final assessment than did students in the comparison group. Thus, participation in the curriculum promotes an ability to remain engaged on a particular argumentative point for a longer sequence of exchanges. Extended rebuttal sequences serve as an indicator of the depth of thinking that is required to fully explore an argumentative point. Engaging in extended rebuttal sequences during dialogic argumentation may well support improvement in expository writing, an issue we return to.

***Research Question: Do students in the intervention group become more skilled in the evaluation of arguments relative to students in a comparison group?***

Our measure of argument evaluation is based on the same distinction as our analysis of dialogic argumentation: recognizing the superiority of direct counterargument over indirect counterargument. The difference is that to perform argument evaluation one must recognize and value this superiority regardless of whether one agrees with its content. That is, when evaluating the two response options, the indirect counterargument (although the less effective argument) may be the one with which an individual actually agrees. To answer correctly, then, one must “bracket” one’s own position and consider the arguments distinct from one’s own beliefs. This requires counterfactual thinking and is difficult for adolescents (and some adults) to do (Goldstein, Crowell, & Kuhn, 2009).

Indeed, as suggested elsewhere (Goldstein, Crowell, & Kuhn, 2009), engagement in dialogic argumentation and associated reflective activities alone may not be sufficient to significantly improve argument evaluation skill. For this reason, in the third year of the

intervention direct practice with argument evaluation was added to the curriculum. We cannot say that this addition was critical, but argument evaluation performance improved from Y2 to Y3. Thus, the same principle of direct engagement and practice that was found effective in promoting argumentative production skills appears to promote argument evaluation skill as well.

Well-developed argument evaluation skills are arguably as important as production skills. Most of the arguments we encounter in life are not ones we personally construct. The ability to recognize when an arguer is directly responsive to a claim made by an opponent (whether that opponent is a live person as in a political debate or a theoretical critic as in the case of an op-ed) is central to argument evaluation. It also provides important information about the relative trust one should place in the ideas of the arguer. If an arguer promotes only his or her own ideas and sidesteps the ideas of an opponent, the arguer is unlikely to be as well informed on the issue as would be desirable and does not see his or her ideas in the larger context of opposing views.

### ***Theoretical implications—Mechanisms***

Dialogic argumentation, we have suggested, is developed through dense and extended practice with dialogic argumentation. This is an expensive (in both time and resources) option and it is reasonable to wonder if the possibly more efficient method of direct instruction would work as well as, or better, than the intensive intervention described here. That is, can we simply inform students that arguments that reply directly to the opposing side are better? Would such a simple directive result in similar gains in students' likelihood to choose the direct counter strategy? Resnick, O'Connor and Michaels (2010) argue that one's ability to engage in reasoned discourse develops through socialization. That is, only through sustained participation in activities and environments that "expect such behavior, support it, and reward it in overt and subtle ways" (p. 172) can the mind be built to privilege reasoned discourse over less effective



means of communication. Additionally, Reznitskaya, Anderson, & Kuo (2007) assessed whether the presence of direct instruction was an effective addition to an intervention focused primarily on practice. Although students who received direct instruction in addition to practice were better able to articulate the goals of argumentive discourse than students who only received practice, they were *less able* to demonstrate the key argumentive skills of counterargument and rebuttal. Thus, not only may direct instruction be less effective in developing counterargument generation skills, it may actually reduce the positive impact of practice.

Additional results are available from the same sample that has been reported on here regarding the impact that sustained engagement with dialogic argumentation over three years has on expository writing (Kuhn & Crowell, in press a). Students who participated in the intervention for at least two years were more likely to write essays that considered the weaknesses of an alternative perspective than were students in the comparison group. This was referred to as a dual perspective, as students who do this acknowledge that their ideas are situated in a framework that includes at least one alternative. Additionally, students who had participated in the intervention for three years began to exhibit an integrative perspective. That is, extended engagement in the intervention promoted essays that acknowledge that the alternative perspective has some favorable qualities and one's own position has some weaknesses. The integrative perspective reflects emergence of the notion that one's ideas exist within a rich context of alternatives; it shows students allowing for complexity within the issue and shows a willingness to take and justify an imperfect position. Most specifically, this perspective requires integration of ideas that lead in opposing directions.

Intervention students also showed greater understanding of the relevance of evidence to argument (Kuhn & Crowell, in press a). Students were asked: "Are there any questions you

would want to have answers to that would help you make your argument?” In addition to asking more questions ( $M = 4.18$  vs.  $.03$ ) intervention students asked general questions that were relevant to the larger issue (i.e. how many countries allow euthanasia?), rather than applicable only to the particular case in question (i.e. is the person in question on life support?) At the final time point, students in the intervention asked case-based questions 10% of the time, while students in the comparison group asked case-based questions 100% of the time.

Previous research has suggested that methods similar to those utilized here are effective at promoting metacognitive awareness and epistemological understanding with regard to argument, (Kuhn, Goh, Iordanou and Shaenfield, 2008) and the change across related skills (written argument and evidence seeking), indicates the presence of a relatively deep developmental change in these respects. However, because we did not directly measure metacognitive skill or epistemological understanding in this study we cannot say definitively whether increased capacities in these respects are the mechanisms underlying the improvements we did observe. This remains a question for further research.

### ***Practical and Educational Implications***

This work provides further support for and extends the scope of previous research done by Kuhn and colleagues. Previous interventions spanned time frames between a few weeks to an academic school year. The work reported here indicates that argument skills continue to develop during all three years of middle school, countering any notion that countereargumentation is a quickly mastered skill that does not benefit from long-term engagement. Thus, the present study provides support for the notion that dialogic argumentation should be engaged regularly and for a long time to support its development.

Without further research, in fact, no one time interval can be statistically credited over

others as the source of the improvements. A practical analysis of the results, however point to the potential for diminishing returns in the third year. Although continued improvement in counterargument broadly defined is seen through the third year, the slope of the improvement becomes less steep over time. Additionally, improvement in direct counterargument is entirely realized by the end of the second year (see p. 52 for discussion of a relevant ceiling effect). The curriculum is quite time and staff intensive, making whether to include a third year an important consideration.

Counterargument is an important skill and we focus this analysis on it because of its centrality to argumentation and its recognized deficit within the adolescent and adult populations. The proposed relationship between dialogic argumentation and the degree of sophistication of one's thinking and writing, however, encourages one to widen one's view to other argument skills that may be developing as a function of dialogic skill development. As reported earlier, it was not until the third year that the intervention group exhibited an integrative perspective on their individually constructed arguments. It was also in the third year that argument evaluation skills became noticeably better for the intervention students than the comparison students. While it is possible that these improvements could be reached in two years of dense practice if the methods and activities are tweaked, it may also be possible that they simply takes three years to develop. More sensitive assessment devices may be necessary to resolve this question.

Additionally, there is evidence that this student population began the intervention more advanced than a typical middle school population. Students in both groups utilized the direct counterargument strategy roughly 30% of the time at pretest. Previous research (Felton & Kuhn, 2001) found that adolescents in a non-competitive middle school utilized the direct

counterargument strategy only 11% of the time; adults in a community college utilized the strategy 24% of the time. Sixth grade students at another relatively competitive elementary school associated with the same university utilized the direct counter strategy 4% of time at pre-test and 24% of the time at the end of one year of an argument skill development curriculum similar to the first year e-dialog curriculum reported here (Kuhn, Goh, Iordanou, & Shaenfield, 2008). Thus, both the intervention and control groups on which we report began sixth grade with competency meeting or exceeding that of students at the end of 6<sup>th</sup> grade who had participated in a year of an argument skill development curriculum.

It is possible that the selective nature of the school and its focus on math, science and engineering makes the student population as a whole more initially competent than the general middle school population. The pattern of development for students in this sample who began the intervention with no competence is very similar to that of the sample in Kuhn, Goh, Iordanou, and Shaenfield's 2008 study, ending the first year with mean competence around 30%. Since the 2008 study spanned only one year, and the subgroup of students who began this intervention at a level similar to the 2008 sample performed similarly to that 2008 sample in the first year of this study, a replication of the three year curriculum would be critical to determine if the further and continued development across the three years of the intervention espoused by the lower performing students in this sample would be more representative of the gains expected from the general middle school population. This is an important question for further research.

### ***Limitations and next steps***

The present work is a first step towards determining how extended (multi-year) engagement in dialogic argumentation, together with reflective activities, promotes improvements in counterargumentation and argument evaluation. The curriculum utilizes many

powerful components: computers, reflective activities, small-group work, and repeated engagement in dialogic argumentation with multiple peers among other components (i.e. tactile modeling with index cards). The complexity and duration of the curriculum preclude one from determining which specific components contributed to the overall results observed. To more fully understand which aspects of the curriculum were responsible for the observed results one would need to evaluate carefully designed replications studies that eliminate single aspects of the curriculum in a comparison group.

Further, the results reported here were obtained from a selective sample of students attending a school focused on math science and engineering. As discussed above, students who applied to and were accepted at such a school may be more likely to value argumentation and may be better able to develop these skills than the general population of middle school students. The contamination that occurred in the comparison group<sup>1</sup> (see p. 42) also points to the need for replication. Also, an outside comparison group should be added to future studies to allow for between-school comparisons, since this school had many specific characteristics.

In addition to replication across more diverse sample, inclusion of other outcome variables would help to paint a more complete picture of the argument skills under development. Such indicators might be how participating in an argument skill development curriculum such as the one reported here affects the certainty with which one holds one's position on social issues and the degree to which one might change one's mind about his or her position on social issues over the duration of the phase. Measures of epistemology, metacognition, and valuing discourse

---

<sup>1</sup> Comparison of those who received the argument-focused curriculum in the comparison group with the those who did not receive such a curriculum does not yield a significant difference [ $t(8.19)=.648, p=.46$ ; equal variances not assumed]. Since pairings were made between those in the comparison group who did and those who did not receive the argument-focused curriculum, proof of gains in one group would show up across both groups.”

would help make the case for the mechanisms of change between and among the separate argumentative indicators.

The argument evaluation work reported here is in its infancy. Before firm conclusions are drawn, results from studies isolating argument evaluation activities would need to be brought to bear. Also, it would be interesting to see whether development in the realm of argument evaluation has a positive impact on engagement in counterargumentation in one's own arguments.

### ***Conclusions***

Education policy makers are currently grappling with the United States' only average performance on the Program for International Student Assessment (PISA) in Science Literacy (Fleischman, Hopstock, Pelczar & Shelley, 2011). As we look down the road at our future national workforce, it seems clear that we will not be able to compete with the 18 countries performing at significantly higher levels on science literacy (including Japan, Finland, Canada and the United Kingdom) as we vie for an economic foothold in these competitive and innovative fields. In response, there is a call for schools to focus on skills that prepare a person to contribute meaningfully to the work that is important in the 21<sup>st</sup> century, including but not limited to science, technology, engineering, and math (STEM).

The cycle of scientific investigation is three-fold, beginning with inquiry (seeking new information in answer to a question), continuing with inference and analysis (bringing information together and evaluating it as a whole) and concluding with argument (utilizing new information as one seeks to enact change and persuade others). Additionally, the process by which scientific information is generated is fundamentally an argumentative one. All scientific knowledge, once submitted to the field, is subjected to rigorous review and revision: that is,

scientists argue about it. Argumentation, then, is both the capstone and the language of the scientific process; it is an important 21<sup>st</sup> century skill. As policy makers and curriculum designers grapple with how to restore the United States to a competitive position within the STEM industries, argument development activities would be a good, even essential place to start.

Further, dialogic argumentation plays an essential role in the development of involved and capable democratic citizens. Argumentation has long been heralded as a cornerstone of a functioning democracy (Michaels, O'Connor & Resnick, 2008). The relationship between skill development and epistemological understanding discussed earlier suggests why. Since those subscribing to a multiplist epistemological positions see no point to knowledge seeking, they will not be inclined to invest the time it takes to generate high quality, well-supported arguments. Since high quality educational experiences promote movement from a multiplist to evaluativist perspective, under current circumstances those with socioeconomic privilege are more likely to progress to the highest epistemological level and, in turn to develop high-level argumentation skills.

Without well-developed arguments coming from across the spectrum of citizens, a democracy suffers in at least two ways. First the social and scientific problems of our time will not benefit from the full consideration of the diverse populace. As the problems of our society are multicultural and increasingly global, problem solving from the narrow perspective of the privileged elite is unlikely to be inclusive enough to truly address the root of those problems. Moreover, for a problem to enter into democratic debate at all it must be clearly articulated and well supported. If the only population capable of entering the debate is the educated elite, the concerns of the less privileged may not be considered at all. In this light, curricula that

incorporate practice with dialogic argumentation activities stand to promote equity within and access to the democratic debate.



## References

- Anderson, R. C, Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S.-y., Reznitskaya, A., (2001). The Snowball Phenomenon: Spread of Ways of Talking and Ways of Thinking Across Groups of Children. *Cognition and Instruction, 19*(1), 1-46.
- Andriessen, J. E. B., Baker, M., & Suthers, D. (2003). *Arguing to learn. Confronting cognitions in computer-supported collaborative learning environments*. Dordrecht: Kluwer.
- Billig, M. (1996). *Arguing and thinking: A rhetorical approach to social psychology* (2nd ed.). Cambridge: Cambridge University Press.
- Barchfield, P. & Sodian, B. (2009) Differentiating Theories from Evidence: The Development of Argument Evaluation Abilities in Adolescence and Early Adulthood. *Informal Logic, 29* (4), 396-416.
- Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning, 1*, 221-235.
- Britt, M. A., Kurby, C. A., Dandotkar, S., & Wolfe, C. R. (2008). I agreed with what? Memory for simple argument claims. *Discourse Processes, 45*, 52-84.
- Bullock, M., Sodian, B., & Koerber, S. (2009). Doing experiments and understanding science: Development of scientific reasoning from childhood to adulthood. In W. Schneider & M. Bullock (Eds.), *Human development from early childhood to early adulthood: Findings from a 20-year longitudinal study* (pp. 172-198). New York: Psychology Press.
- Brem, S.K., & Rips, L.J. (2000). Explanation and evidence in informal argument. *Cognitive Science, 24*, 573-604.
- Clark, D. B., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic frameworks for assessing dialogic argumentation in online learning environments. *Educational Research Review, 19*, 343-374.

- Chinn, C. A., & Brewer, W. F. (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., & Malhorta, B. (2002). Children's Response to Anomalous Data: How is conceptual change Impeded? *Journal of Experimental Psychology, 94* (2), 327-343.
- Chinn, C., Buckland, L., & Samarapungavan (2010). Expanding the Dimensions of Research on Epistemic Cognition: Applying Philosophy to Psychology and Education. Paper presented at the 2010 Annual Meeting of the American Educational Research Association, Denver, CO.
- Common Core State Standards Initiative. (2010). *Standards for mathematical practice*.  
[www.corestandards.org](http://www.corestandards.org).
- De Fuccio, M., Kuhn, D., Udell, W. & Callender, K. (2009) Developing Argument Skills in Severely Disadvantaged Adolescent Males in a Residential Correctional Setting, *Applied Developmental Science, 13* (1), 30 -41.
- Educational Testing Services (ETS) (2003). An Introduction to the Analytical Writing Section of the GRE General Test. Retrieved: December 13, 2008, from ETS website:  
<http://www.ets.org/gre>.
- Felton, M. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development, 19*, 35-52.
- Felton, M. (unpublished manuscript). Metacognitive Reflection and Strategy Development in Argumentative Discourse.
- Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skills. *Discourse Processes, 32*, 135-153.
- Fleischman, H.L., Hopstock, P.J., Pelczar M.P., & Shelley, B.E. (2011) Performance of U.S. 15

- Year Old Students in Reading, Mathematics, and Science Literacy in an International Context. National Center for Education Statistics.
- Goldstein, M., Crowell, A. and Kuhn, D. (2009). What Constitutes Skilled Argumentation and How Does it Develop? *Informal Logic*, 29 (4), 379-395.
- Graff, G. (2003). *Clueless in academe: How schooling obscures the life of the mind*. New Haven: Yale University Press.
- Glassner, A., Weinstock, M., & Neuman, Y. (2005). Pupils' evaluation and generation of evidence and explanation in argumentation. *British Journal of Educational Psychology*, 75, 105-118.
- Hersh, R. (2005). What does college teach? *Atlantic Monthly*, November.
- 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.
- Hofer, B. K., & Pintrich, P. R. (Eds.) (2002). *Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing*. Lawrence Erlbaum, Mahwah NJ.
- Kuhn, D. (1991). *The Skills of Argument*. New York: Cambridge University Press.
- Kuhn, D. (2000) Metacognitive Development. *Current Directions in Psychological Science*, 9 (5), 178-181.
- Kuhn, D. (2001). How do people know? *Psychological Science*, 12, 1-8.
- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuhn, D. & Crowell, A. (in press a) Dialogic Argumentation as a Vehicle for Developing Young Adolescents' Thinking. *Psychological Science*.

- Kuhn, D., & Crowell, A. (in press b) What are the cognitive skills adolescents need for life in the 21st century? In E. Amsel & J. Smetana, J. (Eds.), *Adolescent vulnerabilities and opportunities: Constructivist developmental perspectives*. New York: Cambridge University Press, in press.
- 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. *Child Development, 79* (5), 1310-1328.
- Kuhn, D., Shaw, V., & Felton, M. (1999). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*(3), 287-315.
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5), 1245-1260.
- Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking and Reasoning, 13*, 90-104.
- Lao, J., & Kuhn, D. (2002). Cognitive engagement and attitude development. *Cognitive Development, 17*, 1203–1217.
- Law School Admissions Council (LSAC) (2008). About the LSAT. Retrieved: December 15, 2008, from LSAC website: <http://www.lsac.org/LSAT/about-the-lsat.asp>.
- Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction, 14*, 139–178.
- NAEP (2007) The Nations Report Card: Writing 2007, National State and Tribal Urban Districts. Retrieved December 13, 2008 from NAEP website: <http://nces.ed.gov/nationsreportcard/pdf/main2007/2008468.pdf>

- Oaksford, M., Chater, N., & Hahn, U. (2008). Human reasoning and argumentation: The probabilistic approach. In J. Adler & L. Rips (Eds.), *Reasoning: Studies of human inference and its foundations* (pp. 383–413). New York, NY: Cambridge University Press.
- Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77*, 562-571.
- Piaget, J. (1932). *The moral judgment of the child*. New York: Free Press.
- Piaget, J. (1959). *The language and thought of the child* (3rd ed.). London: Routledge & Kegan Paul.
- Resnick, L.B., Michaels, S., & O'Connor, C. (2010). How (well structured) talk builds the mind. In R. Sternberg & D. Preiss (Eds.), *From Genes to Context: New Discoveries about Learning from Educational Research and Their Applications*. New York: Springer.
- Reznitskaya, A., Anderson, R. C., Kuo, L. - J. (2007). The acquisition and transfer of argumentative knowledge. *The Elementary School Journal, 107*, 449-472.
- Reznitskaya, A., Kuo, L., Clark, A.M., Miller, B., Jadallah, M., Anderson, R.C., Nguyen-Jahiel, K. (2009) Collaborative reasoning: a dialogic approach to group discussions. *Cambridge Cambridge Journal of Education 39* (1) 29-48.
- Shaenfield, D. (submitted manuscript). *Arguing with Peers: Examining Two Kinds of Discourse and Their Cognitive Benefits*.
- Toulmin, S.E., [1958]. *The Uses of Argument*. New York: Cambridge University Press.
- Udell, W. (2007). Enhancing adolescent girls' argument skills in reasoning about personal and non-personal decisions. *Cognitive Development, 22*(3), 341-352.
- Voss, J.F. and Van Dyke, J.A. (2001) *Argumentation in psychology: background*

Comments. *Discourse Processes*, 32, 89-112.

Voss, J. F., Tyler, S. W., & Yengo, L. A. (1983). Individual differences in the solving of social science problems. In R. F. Dillon & R. R. Schmeck (Eds). *Individual differences in cognition* (pp. 204-232). New York, Academic Press.

Vygotsky, L.S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

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

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

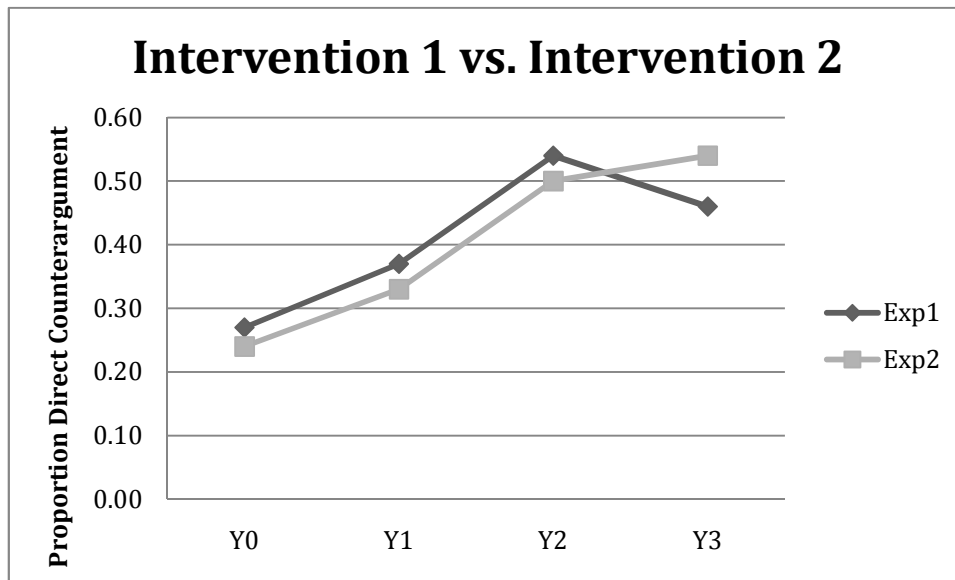
Wolfe, C.R., Britt, M.A., Butler, J.A. (2009) Argumentation Schema and the Myside Bias in Written Argumentation. *Written Communication*, 26 (2), 183-209.

Wolfe, C.R., Britt, M.A., Petrovic, M., Albrecht, M., & Kopp, K. (Accepted). The Efficacy of a Web-based Counterargument Tutor. *Behavior Research Methods*.

## Appendix A: Experimental Group Equivalence

**Mean of proportion direct counterargument use by group and time (SD)**

Group	Pre-test (Y0)	End Year 1(Y1)	End Year 2 (Y2)	End Year 3 (Y3)
Intervention 1 (N=29)	.27 (.19)	.37 (.20)	.54 (.21)	.46 (.19)
Intervention 2 (N=27)	.24 (.26)	.33 (.20)	.50 (.22)	.54 (.20)



**Appendix B: Capital Punishment Opinion Poll****Name:****House:**

Capital punishment, also called the death penalty, is the practice of putting someone to death for committing a serious crime, like murder. Being **FOR** Capital punishment means that you think that someone who has committed a serious crime like murder should be put to death. Being **AGAINST** capital punishment means that you think someone who committed a serious crime like murder should not be put to death.

**Question: Are you for or against capital punishment (Circle one)****FOR****AGAINST****How sure are you about your opinion? (Circle One)**

Certain

Very Sure

Sure

So-so

Not very sure

Not sure at all



## Appendix C: Adaptive Response Schematic

### Why Students Fail in School Schematic



Response items adapted from Kuhn, D. (1991). *The Skills of Argument*. Cambridge: Cambridge University Press.

## **Appendix D: Complete Description of Intervention Topics**

***School Expulsion:*** Doris and Roger at the ColumbiaTown School are misbehaving and disrupting the classroom. Even though they have been given warnings, their behavior does not improve. The school has told them and their parents that if they cannot follow the rules, they will be expelled. Expulsion is permanent and means that they can never come back to the ColumbiaTown school. They will have to be home-schooled or look for a school in another town.

**Question: Should the school board expel the misbehaving students or should they be allowed to stay?**

***Homeschooling:*** Another decision must be made about our new town and its new school! We ask you to consider the case of Nick. As you are all aware, we live in an undeveloped area, and we have a good school that the parents and students are happy with. All of the children in our town attend this school through high school. Since the houses are far apart, school gives children a chance to be together.

A problem has come up! The Costa family has moved to the edge of town from far away Greece with their 11-year-old son, Nick. Nick's parents are both teachers, and in Greece they were keeping him at home and teaching him themselves. Nick was a good student and good soccer player in Greece and his parents have decided that in ColumbiaTown, they want to continue to keep Nick at home with them, and not have him at the school with the other children.

The family speaks only Greek, and they think Nick will do better if he sticks to his family's language, and doesn't have to do his schoolwork in English. They say they can teach him everything he needs at home. As a town, we must decide what to allow.

**Question: Is it okay for the Costa family to live in the town but keep Nick at home, or should they send their son to the town school like all the other families do?**

***USA Intervention:*** A poor South American country is being invaded by a neighboring country which wants to take it over. The people in the invaded country are being injured and killed. The country has sought help from the United Nations (UN) and the UN members have voted to help. The UN has asked developed countries, including the US to provide peacekeeping forces for the invaded country. There are risks to the US troops who enter the country. Throughout its history, the United States has had to decide whether to be concerned about another country's problems. Some take the view that the US should help this South American country. Others take the view that we should focus on problems at home.

**Question: Should the US get involved or not get involved?**

***Kidney Sales:*** Many people need a new kidney. There aren't enough kidneys to give everyone who needs them. Other people are willing to sell one of their kidneys.

**Question: Should people be allowed to sell one of their body parts for money?**

NB: This topic is the first to incorporate the use of evidence, rather than a hypothetical scenario.

***Animal Testing:*** In medical research labs across the country animals are used to test new medications. This testing makes it possible to develop new medications that can save human lives.

**Question: Should companies be allowed to conduct this research upon animals?**

***One Child Policy:*** China is so overpopulated that they can longer take care of everyone. To try to solve this problem and improve everyone's access to goods and services, China has begun a one-child policy. This means that each couple is allowed to have only one child.

**Question: Should China have this one child policy?**

NB: This is the first topic that student's generated their own evidence questions.

***Presidential Priority:*** The United States has recently elected a new president with a lot of work to do! Some people think that the president should focus on domestic issues like the economy or health care, while others think that he should focus on international issues, like the war in Iraq or fixing foreign relationships. Obviously, the president must do both. However, he will have to decide where to focus his first efforts.

**Question: Do you think the president should first focus on domestic or international issues?**

***Intellectual Property:*** Some people think that when he or she downloads a song, tv show or movie he or she should be allowed to do whatever he or she wants with it- including posting it on the internet for others to download or otherwise sharing it with friends. Other people think that the music is owned by the artist/production house/movie studio and so they should be allowed to place harsh restrictions and penalties on people who distribute their work, even if it's just to his or her friends.

**Question: Should people who buy an item created by another person have the right to distribute that product to friends and acquaintances or should the artist have the right to put harsh penalties on people who do so?**

**Juvenile Justice:** Some people believe that teens who commit serious crimes should be tried under the adult justice system. Others think that teens should be tried in a separate system for juveniles. Still others think that teens should be tried in a “teen court” made up of their own teen-age peers.

**Question: Should teens A) be tried teens in the adult system, B) in a juvenile system or C) in a “teen court” by their peers**

**Community Service:** The president and congress are thinking about providing a program in which government funds would be available to support young people to do an optional year of public service (their choice of either a community service organization near home, the Peace Corps, or the military). There is disagreement about whether such a program should be designed for graduating high school students or that students should get a college education first and then do public service. Still others think that taxpayer dollars should not be used for such a program and should go to more urgent things.

**Question: Do you think that there should be A) a program before college, B) a program after college or C) no program funded by the government.**

**High School Curriculum:** What should teens do during high school?

- a. HS students are so busy with required activities that they don't have time to explore the things they are most interested in. By HS students know what they are interested in and good at. They should be allowed to choose what they will study and not have a required set of courses to take.
- b. HS students are so busy with required activities that they don't have time to explore the things they are most interested in. They should do regular HS work mornings but have most afternoons and evenings off to do what they wish.
- c. If HS students expect to do well in HS and go on to top colleges, they have to work hard and put in long hours to excel in the most important subjects: Math, Science and English. To do that they need to devote most of their time to those subjects and to the teachers who know what students need to learn.

## Appendix E: Reflection Sheets

Team Members: \_\_\_\_\_ Date: \_\_\_\_\_

Let's think... starting with the OTHER SIDE'S main argument:

One of the other side's main arguments was:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Our counterargument against this main argument was:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

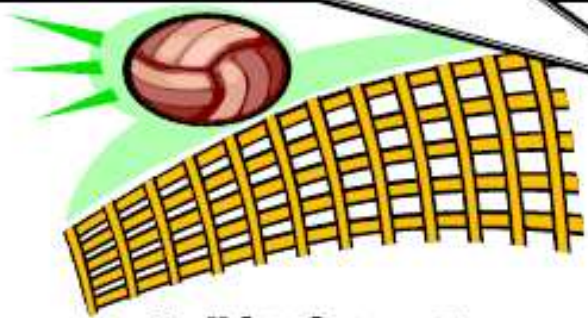
How could our counterargument be improved? Write a better, more effective counterargument here:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_



**Get the ball back over the net!**

Team Members: \_\_\_\_\_ Date: \_\_\_\_\_

Let's think... starting with OUR OWN main argument:

One of our main arguments was:

---

---

---

Their counterargument to our main argument was:

---

---

Our comeback was:

---

---

---

How could our comeback be improved? Write a better, more effective comeback here:

---

---

---

---

**Get the ball back over the net**



**Appendix F: Showdown Preparation Sheets**  
Showdown Preparation Sheets: Our-Side Arguments

Home School

Own Argument:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

...

↓

Counter from Opponent:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

...

↓

Comeback:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

...

Showdown Preparation Sheets: Other-Side Arguments

Home School

Other Argument:

---

---

---

---

...



Counter from Us:

---

---

---

---

...



## Appendix G: Example of Year One Electronic Showdown Transcript

**S** We are against expulsion because there are other punishments like doing cafeteria work, going to military school, community service, and charity work.

**E** They should be expelled because doing work is not going to make them better (unwarranted assumption).

**S** and if they go military school they will expelled from Columbia Town School because they are leaving one school and going to another.

**S** Military school can be done during the weekends or after school

**S** and how do you know that charity work and community service wont make them better? (challenging unwarranted assumption)

**E** You can't go to military school after school or just during the weekends because its like boarding school. You go there and you stay there.

**E** And they are way past charity work and community service because they have been warned dozens of times and they have crossed the invisible line.

**E** You can't go to military school after school or just during the weekends because its like boarding school. You go there and you stay there.

**S** Andres goes to military school only on Saturdays and

**S** they don't have to go to military school they could go to an after school program and be taught privately by someone that the school or parents hire.

**E** What if they don't want to go to tutoring? What if they lie to their parents and say that they are going but they don't?

**E** They might get embarrassed by going to military school and act up in military and Columbia Town school.

**E** What if they don't want to go to tutoring? What if they lie to their parents and say that they are going but they don't?

**E** They might get embarrassed by going to military school and act up in military and Columbia Town school.

**S** It's not their choice whether they go to tutoring or not

**S** and you can't act up in military school (unwarranted assumption)

**S** and if they do act up they can get bad grades and be suspended.

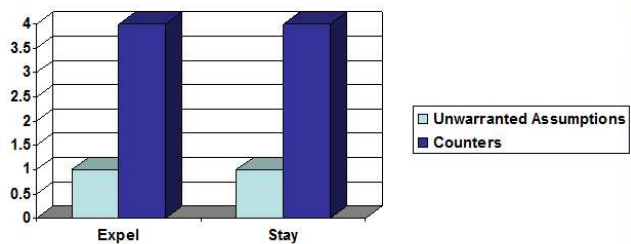
**S** They should be split up into different classes

**S** and they wouldn't tell people they are going to military school if they were embarrassed.

**E** Okay, they go to military school after school and on weekends. They would still disrupt the classes during the week. You never know if they go to military school and still behave badly during the week. And also if they really did care about grades then they would do well in the first place.

## FINAL SCORE:

EXPULSION: 4  
Stay: 4



## Appendix H: Complete List of Questions Provided and Answered

### Kidney Sales

---

1. How much would a kidney cost?
2. How many people need kidneys in the USA?
3. Do people die because they can't get a kidney in time?
4. What are the health risks to those who receive, and those who give a kidney?
5. How do people in the USA currently get a kidney?
6. Can research lessen the need for kidney transplants?
7. Why would someone want to sell one of their kidneys?
8. What percentages of people in the USA have agreed to donate their body organs?
9. What is a kidney and what does it do?

### Animal Testing

---

1. How humanely are animals treated in laboratories?
2. Has animal testing led to cures for any human diseases?
3. What are other ways to find treatments for diseases that do not involve animals?
4. Why are animals useful in medical research?
5. Can medical testing of animals be of any benefit to animals?
6. Does testing of animals for medical research do the animals any harm?
7. How many animals are involved in medical research each year in the USA?
8. What kinds of animals are involved in medical research?
9. Are treatments discovered through animal testing always effective when they are used with humans?

### One Child Policy

---

Coach Generated:

1. Have any countries tried other methods to control their population?
2. What has happened to the size of China's population since it instituted its one child policy in 1979?
3. How much did China's population grow in the years before the one child policy?
4. How have living conditions changed since the one child policy began?
5. What happens if a Chinese family has twins?

6. How many children did most Chinese families have before the one child policy?
7. What if a Chinese family does not agree with the one child policy?
8. Do people in China agree with the one child policy and accept it?
9. What are the effects of growing up as an only child?

#### Student Generated

1. What are the abortion policies in China?
2. Is adoption the result of people having more than one child?
3. What percentage of adoptions are of girls, versus boys?
4. What other methods are there for controlling population growth?
5. How many forced abortions are held in China annually?
6. What happens to unwanted children?
7. Can a Chinese couple adopt?
8. How well do orphanages care for the children?
9. What is the size of China's population?
10. How much room does each person have to live?
11. Is there social security in China?
12. Can adopted children from China find their birth parents?
13. What percent of abandoned children are adopted?
14. Are people forced to give their children up for adoption if they have another child?
15. How do students pay for college in China?
16. How much money does the United States owe to China?
17. Are marriages arranged in China?
18. Is abortion common in China?
19. How is global warming related to overpopulation?
20. Where do human rights come from?
21. Who does the One Child Policy apply to?
22. Are there age limit for adoption?
23. What are the odds of having a multiple birth?
24. What kind of government does China have?

#### Presidential Priority

---

#### Coach Generated questions

- 1a. What's wrong in Iraq and what has the US been doing to help?
- 1b. What could happen if the US doesn't continue to help in Iraq?
- 2a. What's the problem in Darfur?
- 2b. What could happen if we don't intervene in Darfur?

3. How do other countries view the US?
4. What's the problem with healthcare in the US?
5. What's the problem with the US economy?
6. Are Americans satisfied with the way things are going?

#### Student generated Questions

1. How much does the US spend on the Iraq war?
2. How many Americans have lost their jobs?
3. How many alliances does US have with other countries, what are those countries?
4. What will it take to develop helpful new alliances?
5. Will helping other countries help us gather alliances?
6. How can alliances help us to fix the problems of the nation?
7. How many people have gone homeless because of America's [economic] crisis?
8. How many people have lost the quality of life (water, food, shelter, etc) because of the [economic] crisis?
9. If any, what other countries are trying to help the US out of its economic crisis?
10. Have any countries tried to help us with our economic crisis? If we have gotten help, how much?
11. In the past how have economic crises been solved?
12. How have other presidents dealt with these issues?
13. How have we dealt with economic and international crises in the past?
14. Do we have a spare money back for things like this? Could we?
15. How many economic crises have we had before?
16. How many Americans have died in the Iraq war?
17. Which issue affects all of most of the people – domestic or international?
18. What is the US or Obama doing about the economy, for example, how are they making money literally?
19. Approximately how long will it take to bring back the military to America?
20. Why and how did the economy problem start?
21. What other serious international issues are there?
22. What else can we be spending the money that we are spending in the war?
23. How much debt are we in with all the countries?
24. From the 9% of Americans that are "happy" how many of those Americans are very wealthy?
25. What is considered to be an international problem? Is it problems that affect the whole world or certain countries or what?
26. How much money do we have? Do we have more than 3 trillion dollars?
27. What are some major domestic issues in the US?
28. Which problem is easier to solve- international or domestic?
29. How much money does the US spend each day/week?
30. How long was the great depression?

31. How did we get out the great depression?
32. What kinds of extreme things can happen when people lose their job?
33. Which issue is taking up the most money in our budget?
34. How is Barack Obama dealing with the economy?
35. What else is America spending a lot of money on? How much?
36. How long did it take for Great Britain to be our friend after the revolutionary war?

### **Intellectual Property**

---

1. About how many people use Lime Wire? Is it legal?
2. What does copyright do?
3. Have profits gone down/how much money is lost since people have been sharing music with others and the others did not buy it?
4. Why hasn't the government stopped free music downloading programs?
5. How much does the average musician or actor earn?
6. What are the reasons why people create a work of art?
7. How could sharing benefit the entertainment industry?
8. Why can some artists distribute their work for free, while others charge for it?
9. What does it mean to be a free country?
10. How much money does a college student earn? How much money are some of the programs and textbooks?
11. What are the rights that come with buying a song?
12. Is the point of making art to make money? Do artists create work for money?-
13. What happens when the person who creates the music dies?
14. What is DRM?
15. What can artists and singers do to stop copyright infringement?
16. Which ways of sharing are legal?
17. Are Dj's allowed to play music without buying it?
18. Is it legal to give people copies of CD's if you aren't selling it?
19. Where do artists make more money? Selling CD's, doing concerts, or doing appearances? or something else?
20. How many people use i-tunes v limewire?
21. Is Limewire national or international?
22. What is the difference between giving a gift and making a copy?
23. If you share a computer with someone and you both use the same file to add to your own ipods would that be illegal.
24. Is it illegal to listen to songs on YouTube?
25. Why doesn't the government just shut down LimeWire?

1. Can you get a life sentence in juvenile court?
2. Is it the same type of jury?
3. How many teens are tried in teen court; juvenile court; adult court?
4. What are teens convicted of?
5. What is the worse crime that a teenager has ever committed?
6. How many teenage crimes are tried in adult court?
7. How many teenage crimes are tied in juvenile court?
8. What is the lowest penalty a teenager has received when tried in an adult system? The greatest?
9. What happens to a teenager after juvie; what are they're choices; what are an adult's choices?
10. How effectively has each court system dealt with teenage crime?
11. Have teenagers who have been convicted of crime become suicidal in each system?
12. What is considered a serious crime?
13. How many teenage crimes were committed in 2008?
14. What age are you able to get a gun and how do they decide who to give a gun to?
15. Do the teenage courts have different penalties for different crimes?
16. Has a trial been the same but different outcomes due to a different way of trial – such as in a teen vs. adult trial. If so, what was the outcome?
17. What is the age range for teen court? What about those younger?
18. What types of trials are in teen court?
19. How many states trial teens in teen court?
20. How many teens have been put on death row?
21. What does the New York state do about teenage crimes?
22. What types of crimes are charged for the adult court?
23. How many teens in New York State were in court for the same action?
24. Which system has the highest rate of convicting innocent teens?
25. How many teenagers in New York State were tried in Teen Court?
26. Why do teens commit suicide when tried in Juvenile Court?
27. How many teen crimes were committed in New York State in 2008?
28. What type of crimes are tried in adult court?
29. What is the average age of teenagers tried as adults?
30. Do the teen judges go to a short version of law school before they are judges?
31. Has a teen ever been innocently convicted of a crime that got them on death row?
32. Is the death sentence only in the adult court?
33. Does teen court try all ages?
34. What are the benefits of juvy?

1. What are the benefits of doing public service?
2. What are the benefits of joining the military?
3. How many people do these community service/military/Peace Corps
4. What kinds of public services can young people do after college?
5. How many high school graduates decide to do public service compared to how many college graduates decide to do the community service?
6. How old do you have to be to enroll in the army?
7. How many people under 21 are in the army?
8. What is the Peace Corps? who are they? What do they do?
9. What percentage of teens need to get jobs after high school to pay for college?
10. Are college graduates developed more mentally than the high school students?
11. How many teens decide to go to college after high school?
12. What is the difference between the program after college and the program before college?
13. How many people take a gap year between high school and college?
14. What do people do during their gap year?
15. What percentage of people who do a community service year will go to college afterwards?
16. What is the length of time in which you do service?
17. Can people serve in the peace corps or the military without a college degree?
18. What is the average amount of time it takes to finish college?
19. About how many people in the US get loans for college?
20. Does community service help with your college applications or your job applications?
21. What percentage of the federal budget goes towards public service?
22. How many credits you can earn for doing community service depends on the college you attend.
23. How many credits do you need to graduate from college?
24. Would the program pay for college if people did it AFTER college?

## **High School Curriculum**

---

1. Does the average high school student know what they are going to do later in life?
2. Which costs more: everyone choosing their own classes or having the same classes for everyone?
3. Are there currently any high schools that let students choose all of their courses?
4. How many high schools supply career classes?
5. What does the courses high school curriculum look like for students in New York state?
6. Do high school students who know what they want to be have better luck finding colleges?
7. Do colleges look at classes other than the core classes?
8. What skills are required to do well on the SATs? How long is it recommended someone study to get a good SAT score?

9. How does stress affect high school performance? What are some factors that lead to stress?
10. What's the percentage of high school students that have jobs?
11. Do people do better in classes that their interested in, gradewise?
12. What percentage of students do community service after school?
13. Regarding option B, what would students do in the afternoons?
14. What classes do students usually take?
15. Do all high schools offer the same courses now
16. How many high school students drop out?
17. How much influence does the department of education have on courses teenagers take in high school?
18. Do required courses prepare HS students for lots of jobs? How?
19. How many high school students who take the SAT, score in the top 10%?
20. How does boredom affect performance of high school student? What causes boredom?
21. Does having community service and having interesting hobbies help students get into a better college?
22. Do HS freshman know what they want to do when they grow up?
23. What are the average earnings of a person that doesn't have a college education?
24. Can schools add on to the standard graduating courses that the state requires?
25. What are some reasons for dropping out?
26. What is the likely-hood that a HS will pursue ALL 3 of the most important subjects in college?
27. What is a regular high school day for a high school student? (their schedule?)
28. Are there any required courses to get into college (i.e. Math, Sci, ELA)
29. What is the minimum amount of courses a high school student can take? The maximum?
30. On average, how many hours of community service do high schools require students to do before they graduate?
31. Do all high schools require community service hours?
32. If not, then what percent in the US do?
33. How many teens are considered "responsible" or "mature" between the ages of 14-18?
34. What are some techniques of keeping high school students off the streets?
35. In college, do students usually have a full day of college classes?
36. Do students who focus on core classes get better grades than ones who take more electives?



## Appendix I: Example of Year Two and Three Showdown Transcript

Based on the Presidential Priority Topic (See Appendix D for description)

A= Domestic

B= International

- A) **According to evidence card K, the economic crisis comes from lending people money they can't pay back like houses debts and loans and stuff...** [+5, evidence]
- B) Like you said, about Domestic issues costing more than international issues , we need that money before we can even pay for the domestic issues that we have... so we have to stop the war now to get the money back.
- A) **According to question card J, the military will be back from Iraq in 2010, and by that time, our economy will be like, done.** [+5 evidence, -1, unwarranted assumption]
- B) Done?
- A) Well, already, in 2008 **2.6 million people lost their jobs and 12% increase on homelessness, we can't wait that long.** [+5 evidence, -1 unwarranted assumption]
- B) Well, wait, when you say wasting money, what do you mean? What are we wasting money on?
- A) I never said wasting money.
- B) Yes you did, 2 arguments ago you said we were wasting money, like, its not like we're throwing money into the garbage can.
- A) Well, we're at the beginning of this year, so let's just count 2009 and 2010...

TIME

- A) OK, we think Barack Obama should start focusing on domestic issues instead of international issues because **we are spending more money on domestic issues; we are only spending 130 B on the war in Iraq but we're spending over 2 trillion dollars on domestic issues** [+5 evidence]
- B) OK, **so we're spending 130 B annually, so far we've spent 600 Billion**, and the 2 trillion dollars is actually money we owe to other countries. So yes you're right, **we do need to fix that 2 trillion dollars that we're spending, but that's international.** [+1 direct counter, +5 evidence, -1 unwarranted assumption/incorrect evidence]
- A) **No, its domestic issues, per evidence card BB** [+5 evidence, +1 direct counter]
- B) Well, *we still owe 2 trillion dollars internationally*, plus 600B for the war in Iraq. [-5 misuse of evidence]

HUDDLE

B) OK, so one of the reasons that we don't have enough, see to fix the economy we had to go further in debt and *because of the Iraq war we have a continuous annual debt* and so even though it's going to cost a lot of money to bring the troops back from Iraq we'll have 130 Billion dollars more annually to fix the economy, [-1 unwarranted assumption]

A) **Yes but we'll have 2 trillion dollars more if fix the economy** [+1 direct counter, -1 unwarranted assumption]

B) Fixing the economy, OK, if you've even heard your parents listening to the radio in the past 6 months, they keep saying that we have no idea how long it will take the economy to get fixed, but **we know for a fact that to get the troops back from Iraq can be done before August 2010**, I think it was. [+1 direct counter, +.5 use of evidence]

A) OK

B) OK, would you like to respond?

A) **OK, well 2.6 million people have already lost their jobs**, what, OK [+5 evidence]

B) **Would you rather people continue losing their jobs in Iraq or would you rather people continue dying in Iraq?** [+1 direct counter question]

A) Is that the same? Are the same number of people dying as losing their jobs?

B) That's not the point, its not just American's dying. *We're actually, millions of other people have died in foreign countries because of US involvement. We've started so many wars, we need to help those people first, we owe it to them.* [f-1, unwarranted, unsubmitted]

A) *We don't owe anything to them.* And Barack Obama became president of the United States, we elected him president so he could help the United States. *He didn't sign up so he could help other countries.* [-1 unwarranted assumption]

B) **He's helping the United States rep** [+1 direct counter]

TIME

B) **Luis was saying that Barack Obama signed up to help America, but that doesn't mean that we have to help America in America. We can help America in any other country and thats still helping America.** [+1 direct counter]

A) Well, one of the problems with America is that **lots of people are losing their jobs** and jobs are what make America run *and if America doesn't run then lots of countries are going to think that American can't run on its own which is what we don't want to happen* [+5, evidence, -1 unwarranted assumption]

B) Yeah the thing is that **if we can stop the war in Iraq or other expenses internationally then America will have more money** then we can implement that in our *700B dollar bailout which is a stimulus package, so its trying to get America running again and if*

**we can get America running again then so many more jobs will come back** [+0.5 evidence, -.5, unsubmitted, +1 direct counter].

- A) **So basically you know how lots of people are dying in Iraq and lots of people are losing their jobs in America**. Well, the people who are losing their jobs in America *aren't doing anything to cause that* but *the people who are dying in Iraq signed up to do that* [+0.5, evidence -1 unwarranted assumption].
- B) **Are you saying that the innocent American's deserve to die? just because they said that they were willing to risk their lives, then we shouldn't care about them?** [+1 direct counter question]?
- A) No, I never said that, but **if they are willing to risk their lives they need to be ready to do that, but the people in America are like sandwiched, with no way out** [+1 direct counter].
- B) **That's not true, they do have a way out. That's the whole point of that stimulus package** [+1 direct counter]
- A) **But if we help our economy first then, they can be safe and since they didn't volunteer to go homeless and stuff then they can get fixed** and then we can move on to the people who did volunteer to get killed [+1 direct counter].
- B) **Or we could help the people who volunteered to get killed and then help the US because then less people will get killed**, because more people will end up dying [+1, direct counter].
- A) **Well, of course, that's because they knew they were going to die** [+1 direct counter, -1 unwarranted assumption]!
- B) **Not all people, some people were in the reserves and could have signed up before the war started and they could still be in the war or in a grave somewhere in Washington, dead, because they didn't know that the war was going to go on for 6 years and would be so bloody** [+1 direct counter].
- A) Well, **they should have been prepared for that** [+1 direct counter].
- B) Well some people were in the reserves [repeat].
- A) Yeah, but ..Well, let's not get off topic...

TIME

- A) We think Barack Obama should focus on domestic issues first because *we don't have enough money do both and if we don't start with the domestic issues first then they are just going to get worse and worse* [-1, unwarranted assumption]
- B) That's why we need to start with international issues first, **cause if we stop the war then money will start flowing back into America** and that can help us fix the economy [+1 direct counter].

HUDDLE

B) Oh yeah, so if we fix the war then that will bring the money back into the economy [repeat].

A) Yeah but if we are going to bring the people back from Iraq **it's not going to happen overnight, its going to take a lot of time** and if we fix the domestic issues first *and we bring them back we still have to support the ones that are injured* [+1 direct counter,-1 unsubmitted evidence].

B) *Yeah but bringing the troops back from Iraq is going to take less time than fixing the economy* [-1 unwarranted assumption].

A) Yeah, but if you bring them

HUDDLE

A) Well, first of all, that's **an unwarranted assumption because you don't know how long it will take to fix the economy. It could happen over one day, you never know.** [+1 direct counter]

B) You really think that's going to happen?

A) I said it could.

B) The problem is that the domestic issues are trying to implement **new things like rebuilding healthcare** and, for example, Hurricane Katrina there's still a lot of work to do there and that's going to take a lot of money but in order to do that we need to alliances with other countries and stop the... [+0.5 evidence, -1 unsubmitted evidence]

A) But **how are you so sure that other countries are going to want to be aligned with us?** *We are one of the most hated countries* [direct counter +1, -0.5 misuse of evidence].

B) Exactly, so **that's why we need ask them** [+1 direct counter]!

A) **That doesn't make sense, that's like having a best friend and saying "I want you to be my best friend, but they hate you."** [+1 direct counter]

B) Exactly! You have to ask them [repeat]

A) **Ok, there's no guarantee that they'll say yes. So in the time that you are taking to build alliances you could be using to fix domestic problems.** [+1 direct counter]

B) Yeah, but those domestic problems are costing a lot of money and if you think about it, the international issues by stopping we gain money but the domestic issues, stopping them costs money.

A) OK, but **you are saying that bringing the troops back from Iraq is not going to cost money, but they are both going to cost money** we need to focus on the one that people are struggling with. **In 2008 2.6 million people lost their jobs** (+1 direct counter, -1 unsubmitted evidence, +0.5 evidence]

TIME

- A) Ok, the international side is so focused on bringing the soldiers back and ending the war, right? **Well, one death of an American soldier when that soldier has health insurance costs about half a million dollars if that soldier has health insurance and if that person is injured it costs less but we still have to pay those injured soldiers, so even if we bring those people back we're still going to have pay all those people who got injured in that war [+1 direct counter, -1 unsubmitted evidence].**
- B) OK, so you're saying that the soldier that served our country for 6 years now., you're saying that because its costs half a million dollars per death that you are not giving the money for it.
- A) What?
- B) You said that you're not giving us the money for to pay the healthcare. How much would it cost to pay an injury?
- A) I don't know, I just know that when people die its half a million to the wife.
- B) And you're saying we can't afford it?
- A) Right
- B) **So, shouldn't we stop it? Before more deaths come? [+1 direct counter]**
- A) Yeah, but we should also help the United States of America...

HUDDLE

- A) **According to question 4 of the coach generated questions 45 million people under the age of 65 didn't have health care in 2007, and we need to help those people (+.5, evidence].**
- B) So how can we help those people if we don't have the money for it?

TIME

**Appendix J: Argument Game Scripts- Homeschooling**

**MC: I think that homeschooling should be allowed because it should be a parents choice to decide what's best for their kids.**

**Coach 1:** Homeschooling keeps kids from making friends and will stunt their social development.

**Coach 2:** Some parents may choose to homeschool their kids for reasons that aren't good for their kids.

**MC: Aris only speaks Greek, so he should go to the town school so he'll learn English.**

**Coach 1:** Aris can learn English by living here and being exposed to people in the grocery store, for example.

**Coach 2:** It will be easier for Aris to learn difficult academic subjects if he learns them in his own language.

**MC: Aris should be homeschooled because he will be bullied in the town school because he's different from the other kids and they won't like him.**

**Coach 1:** He may not be bullied since he's a good soccer player.

**Coach 2:** Students in the town school should be exposed to students who are different from themselves.

**MC: But Aris won't be able to play soccer because he does not go to the school.**

**Coach 1:** He may be able to play soccer if there is a town team, instead of a school team.

**Coach 2:** The kids at the town school wont give him a chance to make friends, because he is different from them.

## Appendix K: Coding Scheme (Felton and Kuhn, 2001).

### Transactive questions

<b>Agree-?</b>	A question that asks whether the partner will accept or agree with the speaker's claim
<b>Case-?</b>	A request for the partner to take a position on a particular case or scenario
<b>Clarify-?</b>	A request for the partner to clarify his or her preceding utterance
<b>Justify-?</b>	A request for the partner to support his or her preceding claim with evidence or further argument
<b>Meta-?</b>	A question regarding the dialogue itself (vs. its content)
<b>Position-?</b>	A request for the partner to state his or her position on an issue
<b>Respond-?</b>	A request for the partner to react to the speaker's utterance

### Transactive statements

<b>Agree</b>	A statement of agreement with the partner's preceding utterance
<b>Clarify</b>	A clarification of speaker's previous utterance in response to the partner's preceding utterance
<b>Coopt</b>	An assertion that the partner's immediately preceding utterance serves the speaker's opposing argument
<b>Counter-A</b>	A disagreement with the partner's preceding utterance, accompanied by an alternate argument in support of one's own side.
<b>Counter-C</b>	A disagreement with the partner's preceding utterance, accompanied by a critique of the content of the partner's utterance.
<b>Disagree</b>	A simple disagreement without further argument or elaboration
<b>Dismiss</b>	An assertion that the partner's immediately preceding utterance is irrelevant to the speaker's position
<b>Interpret</b>	A paraphrase of the partner's preceding utterance with or without further elaboration
<b>Null</b>	An unintelligible or off-task utterance

Felton, M., & Kuhn, D. (2001). The development of argumentive discourse skills. *Discourse Processes*, 32, 135-153.

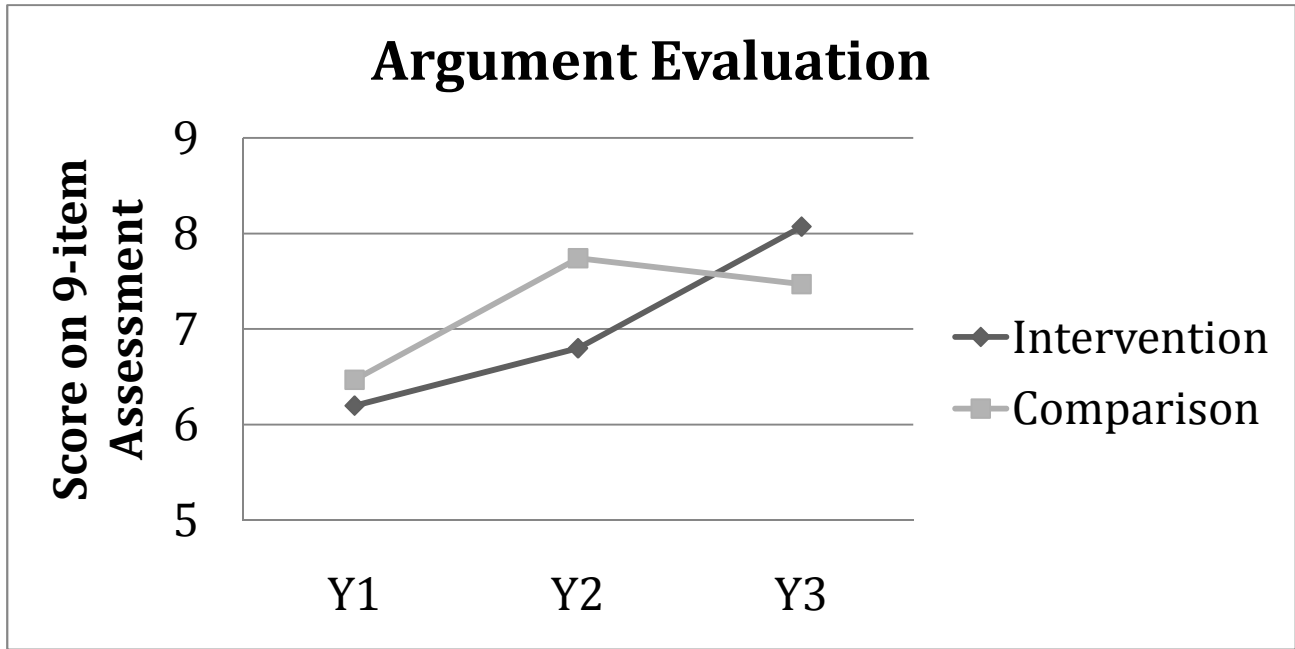
**Appendix L: Year-to-Year Change for Counterargument and Direct Counterargument (p)**

	Y0 v Y1	Y1 v Y2	Y2 v Y3
Counteargument	3.91 (.05)	.25 (.62)	.73 (.40)
Direct Counterargument	.56 (.46)	.34 (.54)	2.24 (.14)



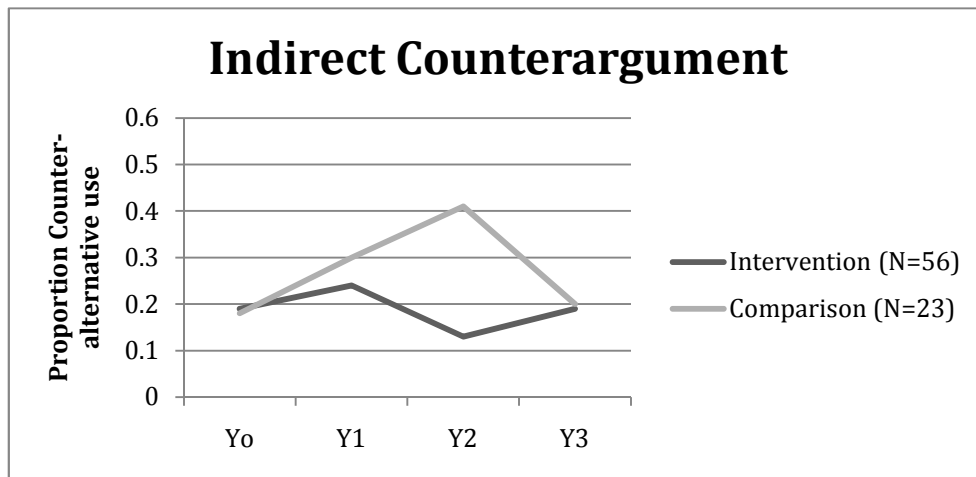
**Appendix M: Argument Test 9-Item Results**

F= 3.31, p= .043, partial eta squared = .10



**Appendix N: Proportion Indirect Counterargument Use**

Group	Pre-test (Y0)	End Year 1(Y1)	End Year 2 (Y2)	End Year 3 (Y3)
Intervention (N=56)	.19	.24	.13	.19
Comparison (N=23)	.18	.30	.41	.20



### Appendix O: Full dialogs for students 32 and 62

Test	Person	Utterance Number	Utterance
Y0	32	1	<b>I am in favor of CP because if the murderer (if he is one) doesn't get killed they will do another crime so they should be punished.</b>
Y0	62	2	The murderer should go to jail because if he makes up for himself he can be free. Killing would just cause problems.
Y0	32	3	So he should be punished for what he did because you wouldn't like it if the murderer killed your parents or something. You would want him to be killed.
Y0	62	4	True, but revenge is never really a good option, especially to kill
Y0	32	5	Okay, so he shouldn't get killed. But he should be punished for what he did, right?
Y0	62	6	Yes, I think if he kills once he should be in jai for the rest of their life.
Y0	32	7	Okay then. He should not be killed but he should be punished.
Y0	62	8	Ok, the case is solved.

Test	Person	Utterance Number	Utterance
Y1	32	1	<b>I am for CP because if the person committed murder that is a really big deal, then the person should get punished for their actions.</b>
Y1	62	2	Ok, but sentencing them to death isn't a good idea. You should give them another chance, and violence doesn't really solve anything.
Y1	32	3	Well, first of all why should they get another chance? The person that got murdered didn't get a chance at all.
Y1	62	4	But murdering the murderer is just going to create problems, because then maybe their family will go after the police because of that. Also, even if the person is murdered they could just send the murderer to jail for a long period of time.
Y1	32	5	What good is that going to do? You never know if when the murderer comes out prison he will murder another person. He should pay the consequences.
Y1	62	6	Exactly, he is paying the consequences by going to jail. And if he murders again, he goes to jail again for another 30 years.
Y1	32	7	he should pay the consequences by Capital punishment. If he keeps murdering people and just sent to jail it might not teach him anything and a lot of lives are going to be lost because of one person who isn't punished correctly
Y1	62	8	If he's going to "murder lots of people" just stick him in jail for life so he wont hurt anyone.

Y1	32	9	He should be sentenced to death. Think about the other people in jail. He could murder someone else there so he should be sentenced to death so no lives can be killed.
Y1	62	10	So make a special cell in jail! Or send him to an asylum.
Y1	32	11	Don't you think that other people are hurt (as in family and friends) by what he has done? He isn't innocent and he should feel what the people he murdered felt when they were murdered.
Y1	62	12	But what if he just stays in jail for a long time? How could he murder people then? And why is the murderer a he, not a she?
Y1	32	13	If someone murdered, for example, your Mom would you want them to feel the same way your mom felt if she was murdered. Shouldn't he suffer for what he did?
Y1	62	14	I would feel that way but I would know that murdering is not the right thing. Revenge is never a good thing.

Test	Person	Utterance Number	Utterance
Y2	32	1	<b>Capital punishment should be allowed because if a person commits murder, or anything else that is very serious, then they should pay for what they did</b>
Y2	62	2	Yes, and they should pay by serving for a long period of time in jail and not being put to death.
Y2	32	3	Why would you want to give them a second chance? They didn't give the person that they killed (murdered) a chance. They took away a life. They don't deserve a long period of time in jail
Y2	62	4	Yes they do, because what if the criminal understands that they did something wrong, and wants to live a better life? They need to get the second chance to do that!
Y2	32	5	It's too late for them to think that what they did was wrong. They should've though of that before they had killed someone. The person that was killed did not deserve to die because they didn't do anything. But the murderer does since they took a life away as well as doing something wrong. Imagine it was a close relative of yours, like your mom or Disagreed, who was killed. Would you want the murderer to stay alive without knowing that there are consequences to their actions?
Y2	62	6	Three things first of all, you gave me no space, second of all, who says that the person who was murdered didn't do anything wrong? And anyways, imagine you the criminal. After being sentenced to jail you would probably have realized your mistake and want a second chance.

Y2	32	7	If the person did do something wrong, then wouldn't the person just go to the police to get them instead of killing them and getting into trouble? That would've been something better to do. Also, in the beginning you said that the murderer should serve jail time for a long period of time. Many other crimes Can make you go to jail for a long time (i.e. that guy that is serving life in prison for tricking people for their money). It isn't the same thing as killing a person. That guy that is serving jail time for that long did not take away a life and that guy is learning his lesson in a way that is appropriate (serving jail time).
Y2	62	8	Okay, well, in the argument you just stated, what if something the person killed did was kill another person? /That then puts the person killed in the same situation as the murderer. /They all deserve a second chance.

Test	Person	Utterance Number	Utterance
6	32	1	<b>I am for Capital punishment (death penalty) because if a person has been guilty of committing murder, then they should pay for what they've done since the crime was very serious.</b>
6	62	2	Exactly! They should pay by spending time in prison.
6	32	3	What is time in prison going to do? Either way how is time in prison going to pay for what happened to the person that got murdered.
6	62	4	Time in prison can be spent doing courses of rehabilitation. Time will help the criminal pay because the Victim has already been murdered and there is no reason to take another person's life away. In addition, you don't know the incentive of the murderer. Someone doing self-defense shouldn't be killed.
6	32	5	Then the crime would be different. We don't know what the court has chosen the case to be. It could have been a 1st degree murder, 2nd degree murder, manslaughter, or self-defense. But if the criminal is guilty of 1st or 2nd degree murder, their intentions were to kill the person, which shouldn't be forgivable at all.
6	62	6	Not true. Everybody needs a second chance. Long periods of time in prison can help the murderer. The murderer would obviously need help if it was 1st or 2nd degree murder.
6	32	7	Well, did the criminal give the person that is now dead a second chance? Did they stop themselves before killing the person? No, no they wouldn't.

6	62	8	Yes, that was the murderer's mistake in not giving the person a second chance. We shouldn't make the same mistake by killing them in return. Placing him/her in prison will probably help them to stop themselves from committing another murder.
6	32	9	If they actually get the help, how are we aware that it has actually helped them? What if they're released from jail and end up killing some other person? That wouldn't be helpful in any way! And what gives them the right to be forgiven for taking someone's life away? I don't think you would like it so much if a person murdered someone very close to you...
6	62	10	I wouldn't. If they killed again, then prison for life. If a person killed another person by accident, it wouldn't make any sense to kill them in return.
6	32	11	We don't even know if the crime was an accident or not. And you're telling me that you would be totally fine with the idea that the person who murdered your close relative, is still given a second chance?
6	62	12	No, but it would be the right choice. If I murdered I would want a second chance.
6	32	13	You would want a second chance, but does that necessarily mean that you deserve that second chance? Also, was it the right choice to even murder someone in the first place?
6	62	14	No one deserves to die. Killing is a mistake. That is what the murderer did, and we would be making a mistake by killing them.
6	32	15	Why would it be a mistake? If it was an accidental killing, then it's a mistake, but if a person has intentionally killed someone, on purpose, then it is not a mistake. It's not like the person said "Oops, I killed him/her!". No, they wanted to kill the person, it was their intention to kill the person and they did.