International Journal of Environmental & Science Education, 2015, **10**(5), 623-647

Comparing Students' Individual Written and Collaborative Oral Socioscientific Arguments

Amanda M. Knight
Egan Maritime Institute, USA
Katherine L. McNeill
Boston College, USA

•*Received 23 March 2015* •*Revised 11 May 2015*•*Accepted 11 May 2015*

Constructing and critiquing scientific arguments has become an increasingly important goal for science education. Yet, the differences in the ways students construct collaborative oral and individual written socioscientific arguments are not well established. Our research with one middle school class in an urban New England school district addresses the following question: What are the similarities and differences between students' collaborative oral and individual written scientific arguments? Data sources consisted of transcripts from three videotaped lessons and associated student work. The sophistication of both the collaborative oral and individual written argument products were analyzed using a proposed learning progression. Results suggest that the students' collaborative oral arguments tended to be of lower sophistication whereas the individual written arguments tended to be of higher sophistication; however both modalities tended to include inappropriate justifications. Moreover, in the written arguments it was easier for students to include a rebuttal than limit their argument to using only appropriate justifications. These findings suggest that there are both commonalities and differences across the expressive modalities that can be targeted in an effort to strengthen the quality of students' arguments.

Keywords: argument, socioscientific, student learning, learning progression, middle school science

INTRODUCTION

The 2007 National Research Council report, *Taking Science to School: Learning and Teaching Science in Grades K–8*, provided a new framework for proficiency in science classrooms, which included a focus on students' ability to "generate and evaluate scientific evidence and explanations" and "participate productively in scientific practices and discourses" (Duschl, Schweingruber, & Shouse, 2007, p.2). This emphasis on disciplinary literacy again reverberates through the *Common Core English Language Arts Standards,* which calls for students to "write arguments to

Correspondence: Amanda M. Knight,
Egan Maritime Institute, United States PO Box 2923, Nantucket, MA 02554, USA
E-mail: aknight@eganmaritime.org
doi: 10.12973/ijese.2015.258a

ISSN: 1306-3065

support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence" (Common Core Standards Initiative, 2010, p.18). Moreover, the *Next Generation Science Standards* (NGSS Lead States, 2013) unprecedentedly seek to interweave scientific knowledge and practices within learning experiences, of which argumentation is one such example. Such opportunities promote communication and "communicating in written or spoken form … requires scientists to describe observations precisely, clarify their thinking, and justify their arguments" (Schweingruber, Keller, & Quinn, 2012, p.74). These policy changes reflect an expanded and more authentic perspective of science competence in which students are expected to construct and critique written and oral arguments using the rules of evidence and reasoning that are respected in scientific discourse.

While the policy changes reflect the view that argumentation is an important goal for science education, incorporating it into classroom practice is a challenging endeavor for both students and teachers (Osborne, Erduran, & Simon, 2004). Some have suggested that providing a socioscientific context supports students in developing stronger arguments as compared to those that are exclusively scientific because students can marshal informal justifications based on their own experiences or ethics (Osborne et al., 2004) and engagement is increased when personal relationships with the issue are reinforced (Sadler, 2004). Consequently, in this study we focused on examining students' arguments within a socioscientific context. Specifically, we focused on comparing students written and oral arguments. Argumentation is an authentic inquiry-based discourse that coordinates conceptual and epistemic goals across both writing and talking (Osborne et al., 2004). Furthermore, research indicates that scientific discourse across modalities is a key mediator to access knowledge in science learning (Kelly & Greene, 1998). Yet little research has compared the quality of students' arguments across the two modalities, which is noteworthy because if there are differences in students' abilities to construct and critique oral and written arguments, then we should be supporting students differently according to the modality. Consequently, our research addresses the following question:

> What are the similarities and differences between students' collaborative oral and individual written socioscientific arguments?

THEORETICAL FRAMEWORK

The quality of socioscientific arguments

Defining arguments and argumentation. Scientific argumentation is an authentic scientific process in which knowledge is socially constructed through evaluating scientific claims, weighing evidence, and critiquing alternative explanations (Driver, Newton, & Osborne, 2000; Schweingruber et al., 2012). For instance, in response to a purported claim, other scientists identify the claim's weaknesses and limitations in terms of how it is being justified (Schweingruber et al., 2012). In the classroom, the argumentation process becomes a beneficial pedagogical technique because it makes student scientific thinking visible when they articulate why they believe a claim to be true, which enables teachers to identify misconceptions and redirect teaching (Osborne et al., 2004, p. 995). Additionally, it is hoped that highlighting competing viewpoints will move students' views of science away from a set of discrete facts and towards a body of knowledge that is constructed by a community through discussion, discernment, and revision in light of both contradictory and confirmatory justifications. These arguments, which are the products that result from argumentation (Sampson & Clark, 2008), include a justification for a claim, can be either oral or written, and can be made about models,

design solutions, questions around measurement, explanations as well as socioscientific issues (NGSS Lead States, 2013). In our research we focus specifically on the quality of students' socioscientific argument products.

Socioscientific arguments. A socioscientific context occurs when social issues are conceptually, procedurally, or technologically related to science (Sadler & Donnelly, 2006). Moreover, socioscientific issues often involve open-ended problems with multiple solutions (King & Kitchener, 2004) that can provide an opportunity for students to learn how to argue (Dawson & Venville, 2009; Driver et al., 2000; Jiménez-Aleixandre, Rodriguez, & Duschl, 2000; Osborne et al., 2004; Zohar & Nemet, 2002) while also developing deeper science understandings (Rogers, Busch, & Berland, 2012; Zohar & Nemet, 2002). However, the additional societal complexities that are introduced in socioscientific issues, such as moral, ethical, and political influences (Cavagnetto, 2010) have different value systems. For instance, whereas empirical measurements and observations are more valued with the scientific community (Aikenhead, 2005), the societal knowledge domains (moral, ethical, and political) frequently rely on non-empirical evidence. Moreover, there may be different criteria for the moral, ethical, and political influences (Kolstø, 2001; Sadler & Zeidler, 2005). These complications could, in turn, make it more difficult for students to critique the appropriateness and quality of justifications.

Quality of Justifications. In an argument, a justification is used to support a claim (Kuhn, 1991). If a claim is not justified or not well justified, then either no argument exists or the argument is of low quality. As such, the quality of the justifications impacts the persuasiveness of an argument, or how it will communicatively influence people. As have other researchers (Sadler & Fowler, 2006), we consider this irrespective of form (moral, ethical, political, or scientific) because it can be methodologically difficult to limit a justification to a single category (Sampson & Clark, 2008). Instead, we measure the quality of justifications in terms of their relevancy, support, and conceptual accuracy. By relevancy, we mean that the justification addresses the topic of the claim. Relevant justifications have the potential to be of high quality if they also support the claim and are conceptually accurate. Whereas support refers to whether the justification exemplifies the relationship purported in the claim, conceptual accuracy refers to whether the justification is conceptually sound. When justifications are relevant, accurate, and supporting, then they are appropriate supports for an argument.

Previous research suggests that when students construct arguments they tend to include inappropriate justifications (McNeill & Krajcik, 2007; Schwarz, Neuman, Gil & Ilya, 2003) and/or struggle to coordinate available relevant justifications to their claims (Sandoval, 2003; Sandoval & Reiser, 1997). Researchers have attributed students' difficulties with evaluating the appropriateness of justifications to content knowledge (McNeill & Krajcik, 2007), inconsistent application of evaluative criteria, and not considering the source, content, or stakeholders' interest in the issue (Kolstø, 2001). While the reasons underlying why students struggle with appropriateness of justifications are important, it is not a focus of our research. Instead, we focus specifically on evaluating whether the justifications, regardless of form, are appropriate within the argument, and how the quality of justifications impacts the overall quality of an argument.

Rebuttals provide additional support for a claim by critiquing how or why an alternative explanation or component thereof is irrelevant, contradictory, or conceptually inaccurate (McNeill & Krajcik, 2012). While the research literature tends to agree that students' arguments with rebuttals are more sophisticated than those without (Kuhn, 1991; Kuhn & Udell, 2003; Osborne et al., 2004; Zohar & Nemet, 2002), it also suggests that there may be a direct relationship between inclusion of rebuttals and conceptual accuracy (Clark & Sampson, 2008). Similarly, we consider both the appropriateness of students' justifications (i.e., relevance,

support, and conceptual accuracy) as well as their use of rebuttals when determining the quality of students' arguments.

Arguments across modalities

Collaborative oral arguments. More than a mere set of concepts, science is a culturally bound discourse with respected ways of thinking, behaving, and reasoning that are learned through social interactions (Michaels, O'Connor, & Resnick, 2008). In the classroom, students learn the language and methodology of science (Osborne et al., 2004) as well as develop science understandings (Varelas, Pappas, Kane, & Arsenault, 2008) through participation in scientific inquiry because it requires them to do science while also communicating their ideas (Duschl, et al., 2007). While there has been some research on the process of developing collaborative arguments (e.g. Evagorou & Osborne, 2013; Kuhn et al., 2008; Ryu & Sandoval, 2008; Sampson & Clark, 2008, 2009), little is known about how this affects students' argument products (Evagorou & Osborne, 2013) or how the quality of students' collaborative oral argument products compare to their individual written products. It is the latter that we seek to explore in this study, which is important to consider because if there are differences in students' oral and written argument products then we should be supporting students differently within each modality. Specifically, by collaborative oral arguments we mean the final argument a group of students develop through discussion and debate.

The research on collaborative oral argumentation can, however, inform our research on collaborative oral arguments. Specifically, the research suggests that the interactions within collaborative argumentation are dependent on the group members and with the issue at hand (Evagorou & Osborne, 2013; Ryu & Sandoval, 2008; Sampson & Clark, 2009). Specifically, research suggests that while students can initially articulate their claims, they often struggle with providing justifications (Jiménez-Aleixandre et al, 2000). However, research also suggests that both the quantity (Sampson, Grooms, & Walker, 2010) and quality (Osborne et al., 2004) of students' justifications and evaluating criteria can increase over time. Regardless, students tend not to explore alternate explanations (Sampson et al., 2010) or tend not to be able to do so while at the same time trying to persuasively present their own argument (Berland & Reiser, 2011). Consequently, students' oral argument products may also tend not to include appropriate justifications or critiques of alternate explanations (e.g. rebuttals). In our research, we compare the sophistication of students' collaborative oral arguments and individual written arguments to explore similarities and differences between the two modalities.

Independent written arguments. Similar to the research on students' collaborative oral arguments, research on students' written arguments suggests that students often struggle with the quality of their justifications in terms of their appropriateness and sufficiency (McNeill, 2011; McNeill, Lizotte, Krajcik, & Marx, 2006; Sampson et al., 2010). While this is likely related to students' content understandings, it may also indicate students do not understand what counts as a justification (McNeill, 2011; McNeill et al., 2006; Sampson et al., 2010). The latter premise is further warranted when we also consider that students tend not to use observations or lack of data as evidence (Sandoval & Millwood, 2005). It is, however, promising that some research suggests, with a sustained focus on the structure of arguments, students may gain a better understanding of the persuasive genre as well as increase the quality of their individual written arguments (McNeill, 2011). In our research, we measure the quality of students' individual written as well as collaborative oral arguments, and investigate similarities and differences between the modalities.

Comparing collaborative oral and independent written arguments. While some researchers have made comparisons between the two expressive modalities (e.g. Halliday & Martin, 1993; Rivard and Straw, 2000), very little empirical research has done so within the argumentation literature. The closest example compared the structure of arguments that occurred during group discussions to individual students' written arguments on the same scientific topic (Berland & McNeill, 2010). Specifically, they used a learning progression, which placed students' understandings along a continuum, to measure the sophistication in both modalities. Their findings suggest that the structure of students' arguments within their discussions was more sophisticated than their written products, which they attributed to the need to be persuasive to a live audience as well as a higher frequency of rebuttals that resulted from hearing others' counterarguments.

Two additional studies made similar comparisons between the quality of students' collaborative oral argumentation and their individual written products, however the individual written products always occurred after the students participated in the oral process (e.g. Evagorou & Osborne, 2013; Sampson et al., 2010). One study did observe a cumulative trend across the two modalities within a scientific context. More specifically, they found a statistically significant positive relationship between a group's composite oral argument scores and individual written argument scores (Sampson et al., 2010). However, the other study did not find a relationship within a socioscientific context (Evagorou & Osborne, 2013). While this discrepancy could be due to the differing contexts (scientific versus socioscientific), it could also be related to the ways these students engaged with each other and/or the issue under debate.

Taken together these studies suggest that there are differences within the quality of students' oral argument discourse and their written argument products. However, inflated in each of their results are differences in the discourse form (e.g. process and product) as well as differences in the modalities (e.g. oral and writing) and the number of participating students (individual versus collaborative). To our knowledge, no research that has examined these factors separately. In our study, we hold the form of the discourse constant (e.g. product) and compare their sophistication across modalities (e.g. oral and writing). We were not, however, able to separate the number of students participating (e.g. individual versus collaborative) from the modality (e.g. oral and writing) because talking inherently involves more than one person and writing tends to be individual.

Research question

In this study, we situate the arguments within socioscientific issues and encourage the students' relationship to the issues in an effort to provide an authentic reason for the students to be persuasive. Additionally, because the students' socioscientific arguments traverse both talking and writing, we compare their quality across modalities. More specifically, the research question was: What are the similarities and differences between students' collaborative oral and individual written socioscientific arguments?

METHODOLOGY

Context of the study

This study took place in one middle school science classroom in a large New England urban school district. The teacher, Mr. Keiffer, previously participated in two series of professional development (PD) workshops focused on scientific argumentation, which were provided by our research team. The beginning and

advanced workshops were held in consecutive school years, and this study took place following the completion of the advanced workshops. The PDs provided strategies on how to integrate the claim, evidence, reasoning, and rebuttal (CERR) framework (for more details, see McNeill & Krajcik, 2012) for scientific argumentation into classroom practice, which was illustrated with video clips of teacher practice, student writing, and classroom transcripts. In small learning groups, participant teachers also designed learning tasks and reflected on their outcomes.

Three scientific argumentation lessons were observed. The teacher was requested to provide opportunities for oral argumentation products—whole class and/or small group—in addition to written products. While the written products were individually constructed, the oral arguments were collaboratively constructed and presented by small groups.

Participants

Mr. Keiffer was a 7th grade math and integrated science teacher. At the time of this study, he had six years teaching experience as well a bachelor degree in science and a master's degree in education. Mr. Keiffer was selected from teachers who had previously participated in a beginning level scientific argumentation PD series provided by our research team, and responded with intent to participate in an advanced level series of PDs. Teachers who met the two prior qualifications, taught at the middle school level, and previously developed good quality data-driven argumentation questions were solicited to determine interest.

Mr. Keiffer taught in an urban New England public school that emphasizes math and science. Of the 18 students in Mr. Keiffer's class, 17 of the students (9 females and 8 males) participated in this research. In a demographic survey, 53% (n=9) of the students identified as being Black/African American, 12% (n=2) identified as being both Black/African American and Native American or American Indian, 24% (n=4) identified as being Latino/Latina, 6% (n=1) identified as being white and other, and 6% (n=1) identified as being other. While only one student identified that he was born in a country other than the United States (Nigeria), 29% (n=5) of the students identified one parent as being born in a country other than the United States, and 12% (n=2) identified both parents as being born in countries other than the United States. Additionally, 18% (n=3) of the students identified speaking a language in addition to English—all three identified the other language as being Spanish. Moreover 12% (n=2) identified that their parents speak to them in another language (Spanish; Portuguese), but that they respond in English. This is an ethnically and linguistically diverse class, which is an increasingly more common phenomenon within our nation's urban public schools.

Data collection

Data were collected in regards to both oral and written argument products in three-videotaped lessons over two months. The lessons are summarized in Table 1. The videotaped sessions varied in length from 35 to 100 minutes, and each of the three lessons focused on a different thematic unit. The units spanned more class sessions than were videotaped, however we only recorded lessons in which products were completed. By product, we mean the argument artifacts that were collected from each lesson. More specifically, when the focus of the lesson was on written arguments, then the product was individual student arguments. In comparison, when the focus of the lesson was on oral arguments, then the product was in the form of a collaborative group argument. This is different from the dialogic form of oral argumentation that is often referenced within the research

literature (e.g. Jiménez-Aleixandre et al., 2000; McNeill & Pimentel, 2010; Osborne et al., 2004). Instead of looking at the process by which a group of students orally debated an argument, we analyze the sophistication of the collaborative oral argument product that was constructed by a group of students. Moreover, the collaborative group arguments were not scripted word for word. Rather, the groups came to a consensus on how they would respond to the argument and who was responsible for saying what when they presented their perspective to the class. Therefore, they are inherently different from the individual written arguments.

Table 1. Summary of lessons and data

Lesson	Modality	Length of Lesson (minutes)	Products	Topic
1	Talk	75	9 group arguments (5 groups of 3-4 students and 2 arguments/group; 1 missing argument)	Belo Monte Dam
	Writing	35	13 individual arguments	
2	Writing	65	12 student arguments	Tap Water vs. Bottled Water
3	Talk	100	6 group arguments (6 groups of ~3 students and 1 argument/group)	Community Fair

Individual written arguments were collected in lessons 1 and 2. Of the 17 students, artifacts were collected from 13 students from lesson 1 as well as 12 students from lesson 2. This represents 74% of the potential arguments that could have been collected. The missing student work is the result of student absences when the writing was either assigned or collected as well as incomplete work.

Collaborative oral arguments were collected in lessons 1 and 3, and the groups of students constructing the arguments changed across lessons. In lesson 1, five groups presented two arguments each, and there were three or four students per group. While ten oral arguments could have been collected from lesson 1, the video for one lesson was not captured due to technical difficulties. Therefore, there are nine argument products for lesson 1. In regards to lesson 3, six groups presented one argument each, and there were approximately three students in each group. It is also important to note that every student in each group participated in the preparation and presentation of the group arguments. While we can trace what each student contributed to the argument as a whole, we do not know how the interaction of the group affected what each individual presented. Consequently, the unit of analysis is at the group level (as opposed to the individual level).

Lesson 1. The first unit addressed the following question: Should the Belo Monte dam be built? The building of the Belo Monte dam on the Xingu River in Brazil —a tributary of the Amazon River—has been under debate for nearly 25 years. While some argue that the relatively clean and consistent source of hydroelectric power is needed for the country to develop, others argue that the cost associated with the destruction of the rain forest is too high. While the dam has not yet been built, current political leaders based recent campaigns on promises to do so.

The unit addressing the Belo Monte dam spanned eight class periods, and resulted in both collaborative oral and individual written argument products. The collaborative oral arguments were presented to the class on day 6 of the lesson, and the individual written argument products were collected at the end of the unit on day 8. In preparation for the presentations, the class spent one period learning about how dams generate electricity. In the second lesson the class was divided into

five groups—the power company, hydrologists, ecologists, climate scientists, and the Kayapo Tribe—and each group was directed to represent the perspective of their assigned group. While the first two groups represented the pro-dam perspective, the latter three groups were against building the dam. Each group read a different 1-2 page article that included a summary of the issue as well as justifications from the perspective of the group they were representing. After reading the article, the groups were instructed to identify a claim and justifications on a graphic organizer. The following day (i.e. day 3), each group predicted and summarized the argument for each of the other groups. On the fifth day of the unit, the students prepared their presentations, which constituted determining each student's role in the group presentation, the order in which the students would talk as well as numerous rehearsals. On day 6, Mr. Keiffer introduced his expectations for the presentations saying:

> We're going to give the power company, um, the first chance to speak and they'll also get the last chance to speak at the very end because they have a close relationship with the government, so they are really important. Um, and then we'll go through the ecologists, the hydrologist, the climate scientists. If we have time we'll even let the Kayapo Tribe speak. Um, and then your job up here, as you remember, is to start off with your claim. Tell, you say, we are the climate scientists and we want to build this dam because blah blah blah blah blah and then you list your evidence and reasoning.

By focusing on each stakeholder's perspective, Mr. Keiffer is setting an expectation to be persuasive by attending to the audience, which he again frames as he explains what he expects the students to be doing when they are listening to the other groups' arguments:

> Every group that goes up, you are going to have to write down either they're for or against the dam. You are going to write down at least 2 pieces of evidence that they give and then after they present I'm going to give you about 30 seconds to write a question to ask them.

Each initial oral argument presentation was followed by a question from each of the other four groups. While this was designed as an opportunity to critique the quality of the claims and justifications, the students tended to focus on sense making and asked clarifying questions. The session concluded with a final statement from each group. Again, Mr. Keiffer set persuasive expectations, but also asked the students to consider the quality of their justifications:

> It sounds like some people are really fired up and, sort of, want to have a last, last go at it. And so we're going to give everybody a final statement … You're just going to get one last final chance to argue … Alright, as you're preparing this, figuring out what you're going to say, what's most important, you might think about what facts do you think, now that we've heard all of the evidence, are most important? … Alright, think about which arguments you really want to drive home.

Because the students had the opportunity to prepare both their initial and final arguments, both were viewed as products, and, therefore, transcribed for analysis. In comparison, the question and answer sessions were more representative of students asking and responding to clarifying questions as opposed to debating parts of the argument. This was, therefore, outside our scope.

On the days after the debate (i.e. day 7 and 8), the students constructed individual arguments from their own perspective. The students used research from their oral arguments as well as justifications from any of the articles used for the oral arguments, Internet searches, and videos the class watched on the topic. In introducing the directions for their work, Mr. Keiffer again presented his persuasive expectation by highlighting the contrasting viewpoints:

> Can I just take a vote in here real quick? How many people decided: Yes, the dam should be built and these are the reasons why? So, Ben said that and Amy. And how many people said: No, this dam should not be built and here are the reasons why. Ok, so, a lot of people chose that one. Ok. That's fine as long as you provide evidence.

During this timeframe the students also completed a peer editing and revisions. As very few students finished, they were required to do so for homework. We collected 13 student responses for analysis.

Lesson 2. The second lesson was recorded in one session and addressed the following question: Should people drink tap water or bottled water? The question stemmed from the students' own experiences with tap and bottled water in their school. More specifically, the students' school was originally plumbed with lead pipes and provided bottled water so as to avoid lead poisoning. The class researched the sources of their schools' bottled and tap water as well as the merits and drawbacks of each source. Ironically, they learned that both originated from the same reservoir, and that there were differences not only in how they were distributed, but also in how they were tested and prepped for distribution. In introducing the written argument prompt, Mr. Keiffer structured his expectations in terms of taking a side:

> Today we have to sort of wrap up all the information that we learned, from, alright, we learned about bottle water and tap water right? And so probably we've got lot of evidence, and some of it's for bottled water and some of it's for tap water.

He again polled the class in terms of which claim the students supported. This time he was surprised to find all the students in support of drinking tap water. He then asks the students to consider the quality of their evidence:

> You guys need to, what's the most convincing or most important piece of evidence to say that we should drink tap water … And you need to find three pieces of evidence that support your claim, … and that are the best piece of evidence, the most important or the most convincing.

This again shows how he framed persuasive expectations.

Lesson 3. The last lesson was a culminating project for the school year in which the students prepared presentations for a community fair. The goal of this fair was for the students to inform and persuade the attendees—which included students in other grades, family members, and community members—on a variety of socioscientific issues they had researched throughout the year. Logistically, the class was divided into six groups, and each group researched and prepared an argument with supporting visuals around a different question. The questions are summarized in Table 2. In introducing the project, Mr. Keiffer said:

> So basically, people are coming to this fair, and you guys are the experts. You're going to know everything there is to know about your group. And by the end of their time at your station, you need to have convinced them to take action because of all the evidence you've shown them, and explained to them.

Again, we see that Mr. Keiffer introduced his expectations in terms of being persuasive or convincing. Each group presented their argument to the first author on the day prior to the fair.

Table 2. Summary of questions in lesson 3

Group	Question
1	What is the impact of dams on people and ecosystems?
2	How has overconsumption of resources affected people, the environment, and wild Atlantic salmon?
3	Why is it important to protect the rainforest?
4	How do water pollution and wastewater create problems for obtaining a clean water source?
5	What are the reasons why people do not have equal access to essential resources such as water?
6	Why is it important to protect our oceans and rivers?

Data analysis

A theoretical learning progression was used to analyze both the students' written and oral arguments in terms of an increasing sophistication of the structure. While all learning progressions indicate successively more sophisticated ways of thinking about a topic (Duschl et al., 2007), we employed an approach that progresses from students' naïve forms on the lower border to more scientifically accepted forms on the upper border (Furtak, Thompson, Braaten, & Windschitl, 2012). The employment of the proposed learning progression as a coding scheme afforded the opportunity to qualitatively analyze both oral and written arguments at the argument grain size as opposed to the components within the argument as is done routinely within the argumentation literature (Bell & Linn, 2000; Jiménez-Aleixandre et al., 2000; McNeill, 2011; Osborne et al., 2004). By argument grain size we mean that the entire argument was given a single code. For instance, instead of counting the frequency of the argument components—claims, evidence, reasoning, and rebuttals—we focused on how the structure of the argument and the appropriateness of the justifications impacted the sophistication of the argument.

The argumentation learning progression used for analysis was adapted from McNeill, Corrigan, Goss, and Knight's (2012) expression of the scientific argumentation construct map. An adaption was warranted because we were applying the scientific argumentation construct map to socioscientific arguments. In particular, the original construct map separated evidence and reasoning into separate constructs. However, socioscientific arguments can rely on other forms of justification (e.g. ethical, moral, or political), and these are sometimes stronger forms of justification (Sadler & Zeidler, 2005). Consequently, in our adaptation all forms of justifications were considered equal. Our proposed argumentation learning progression is presented in Table 3, which we used to code the sophistication of both the oral and written arguments. The increasing sophistication of the argument is dependent on the structure: 1) Claim – an answer to the question, 2) Justification – support for the claim, and 3) Rebuttal – a justification for how or why an alternative explanation is incorrect (McNeill & Krajcik, 2012). Additionally, the levels of sophistication are dependent on the appropriateness of the justifications measured in terms of relevancy, accuracy, and support. By relevancy, we mean that the justification addressed the topic of the claim. In comparison, accuracy addressed whether the justification was conceptually sound, and support addressed whether the justification exemplified the relationship purported in the claim. We do not dispute the literature that suggests arguments with rebuttals are of higher quality than arguments without rebuttals (Clark & Sampson, 2008; Kuhn, 1991; Kuhn & Udell, 2003; Osborne et al., 2004; Zohar & Nemet, 2002). This is, in fact, reflected in our learning progression. However, we concurrently measure the quality of the argument in terms of the presence of rebuttals and appropriate

justifications. While it is very well possible that rebuttals may be more difficult than limiting arguments to solely appropriate justifications, based on the frequency of rebuttals as compared to solely appropriate justifications within the arguments in our data, we believe that limiting arguments to relevant-accurate-supporting justifications may be more difficult. Consequently, this is reflected in how we compare the quality of students' independent written and collaborative oral socioscientific arguments. Example arguments that align with each level within the learning progression are presented in Table 3.

Two persons, with a 31.7% overlap, independently coded the students' written and oral arguments with 89% reliability. Prior to coding the oral arguments, one coder chunked the transcript into sections according to changes in classroom activity. Within the artifacts the coders looked for evidence of a claim, appropriate justifications (e.g. relevant-supporting), inappropriate justifications (irrelevant/inaccurate/non-supporting), and rebuttals. This evidence was used to determine the level on the learning progression, and only one code was assigned for each argument. For instance, an argument that includes a claim and inappropriate justification corresponds to level 2a. In comparison, an argument that includes a claim, appropriate justifications, inappropriate justifications, and a rebuttal corresponds to level 2c. Once reliability was achieved, we collapsed the data across the three lessons within each modality. Specifically, the frequencies of the oral and written arguments were tabulated and graphed. This method afforded the opportunity to visually represent trends between and within modalities (Miles & Huberman, 1994). Quotes from the coding charts were then compared and reflected upon in order to better understand the nature of the trends. The combination of the trends and quotes supporting them was used to articulate three emergent themes. The themes and supporting quotes will be discussed next.

Table 3. Proposed argumentation learning progression

Learning Progression			**Examples: Should the Belo Monte Dam be built?**
3. Appropriate Justification	3b	Student constructs an argument with only ***appropriate justifications*** for the claim as well as ***rebuttal*** that critiques alternative evidence and/or reasoning	*Level 3 +* Although the power company argues that most of this area is already deforested, history has shown that whenever you build roads to a new construction site in the Amazon rainforest people use these roads to sneak in and cut down more trees.
	3a	Student constructs an argument that includes ***only appropriate justifications*** for the claim.	The Belo Monte dam should not be built. The dams' reservoir causes methane pollution. When the reservoir is formed vegetation decomposes and releases a toxic gas called methane. Methane rises into the atmosphere and causes global warming. The reservoir can possibly release more than 100,000 pounds of methane.
2. Justification	2c	***2b*** as well as a ***rebuttal*** that critiques alternative evidence and/or reasoning.	*Level 2b +* Although the power company argues that most of this area is already deforested, history has shown that whenever you build roads to a new construction site in the Amazon rainforest people use these roads to sneak in and cut down more trees.
	2b	Student constructs an argument that includes both ***2a*** as well as ***appropriate justifications*** for the claim.	*Mixture of Level 2a and Level 3a*
	2a	Student constructs an argument that includes only inappropriate justifications for the claim (e.g. ***irrelevant, inaccurate***, ***or non-supporting)***.	The Belo Monte dam should not be built. 20% of the amazon rainforest has been cut down mostly because of dams [inaccurate justification].
1. Claim	1	Student constructs the ***claim*** of an argument.	The Belo Monte dam should not be built.
0. No Claim	0	Student ***does not provide a claim*** to argue.	The pro is that people are building the dam to get electricity in Brazil. But, the con is people that live there will be losing their homes.

RESULTS

The analyses address the following research question: What are the similarities and differences between students' collaborative oral and individual written scientific arguments? Because the number of participants was at the group level for talk and student level for writing, we will next present the findings for each modality separately. We will then cautiously proceed with a comparison across the two modalities with the caveat that they were constructed by differing number of students. We believe this to be a necessary concession when authentically comparing across writing and talk because of the nature of each discourse. More specifically, talk occurs between people and thus is inherently a group activity whereas writing is most often constructed as an individualistic activity. Therefore, an authentic comparison would include differences in the number of participants contributing to the argument product. The themes organized according to modality are presented in Table 4.

Table 4. Themes organized by modality

Theme	Argument Modality		
	Oral	Written	Comparison
Theme 1: While the students regularly provided justifications, they tended to use inappropriate supports (e.g. irrelevant, inaccurate, and/or non-supporting)	√	√	
Theme 2. Including a rebuttal was easier than including only appropriate justifications for individual written arguments		√	
Theme 3. The most sophisticated arguments were in individual's writing while the least arguments were in the collaborative talk			√

Students' collaborative oral arguments

Using the methodology previously discussed, we tabulated the number of oral arguments for each argument level, which is presented in Figure 1. We present the actual sample frequencies as opposed to percentages due to the small sample size. From Figure 1 our first theme quickly becomes apparent: While the students regularly provided justifications in their oral arguments, they tended to use inappropriate supports (i.e. irrelevant, inaccurate, and/or non-supporting). These low quality justifications were captured in level 2 (i.e. 2a, 2b, and 2c) of our learning progression. Specifically, 10 responses were at level 2b—justifications with mixed appropriateness. Additionally, there was 1 response at level 2c—justification with mixed appropriateness as well as a rebuttal. In total, inappropriate justifications were found in 73% of the students' oral arguments. However, it is important to note that these students were trying to justify their claims.

The use of inappropriate (i.e. irrelevant, inaccurate, and/or non-supporting) justifications is exemplified in the final statement from the hydrologist group during lesson 1, which addressed the question: Should the Belo Monte Dam be built? From a pro-dam perspective they argued:

> Alright, so, it's not only us. Not only dams can cut trees, kill fish, or have floods. You know, people, as we were talking about with salmon are being overfished. People, no people, loggers are cutting down trees and floods, like have you heard of [inaudible] that have floods? See, so it's not only us.

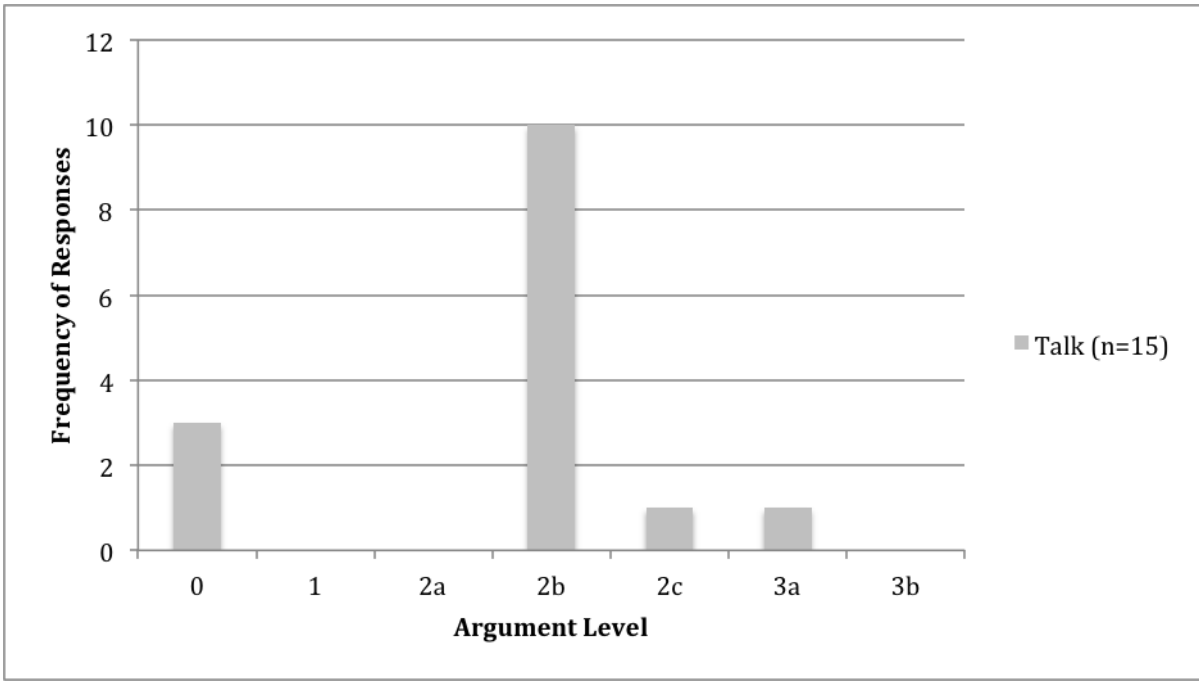

Figure 1. Frequency of collaborative oral arguments at each level of the argument learning progression.

In taking on the perspective of the hydrologists this group was trying to argue that the dam would provide energy and would not hurt the environment. However, they proceeded to admit that it would, in fact, hurt the environment with the caveat that they would not be the only ones doing so. This line of justification supported the claim and was, therefore, relevant; however it was contradictory to their claim and thus undermined their argument that the dam should be built.

While the previous example illustrated how a group inappropriately used a relevant-contradictory justification, we next turn to an example in which a group provided irrelevant justifications. More specifically, this occurred during the community fair arguments (i.e. lesson 3) in which each group addressed a different socioscientific issue they had previously studied during the academic year. This particular example comes from the group answering the following question: How has overconsumption of resources affected people, the environment and the wild Atlantic salmon? In their oral presentation of the argument they said:

> There are three major reasons why salmon are in danger. Natural causes: the soil covers the salmon nests and eggs die. This is a major reason because salmon need to live in cold water to survive. Predators: many animals eat salmon, or bears, even sea lions, killer whales and more. This is a major problem because there are many predators. Dams: Salmon also use up a lot of energy as they try to find their way around way past dams. Many dams are in the way of salmon so they use more than, more than half of their energy to get over.

While these justifications all support the declining population of the wild Atlantic salmon, none of them address how human's overconsumption has severely impacted their population. These justifications are, therefore, irrelevant to the

argument. Additionally, in addressing how overconsumption has affected people they said:

> There are several affects of growing levels of consumption on people. Too much shopping. The average American spends thousands of dollars each year on personal consumption. Many Americans waste money, create [inaudible]. People work to get money to see TV and to buy. This makes people unhappy.

As this was a science class, the students were supposed to be discussing how overconsumption of natural resources has affected people. For instance, a strong example might link deforestation to increased carbon dioxide concentrations to global warming to a rise in sea water, and then explain how people living in low lying coastal communities could be affected by rising sea waters. In fact, this particular group had the data to support this argument and, yet, they inappropriately focused on financial resources to construct an opinion argument based on the idea that spending money makes people unhappy. While the students' use of irrelevant justifications could be related to a lack of clarity in the question, we further argue that the use of opinions as a form of justification makes these justifications irrelevant. More specifically, the scientific community does not accept opinions as a source of evidence; therefore their use in this context is outside the scope of the project, and, thus irrelevant. In summary, although the majority of students provided justifications for their claims within their collaborative oral arguments, they struggled with the issue of appropriateness in terms of support, accuracy, and relevancy.

Students' independent written arguments

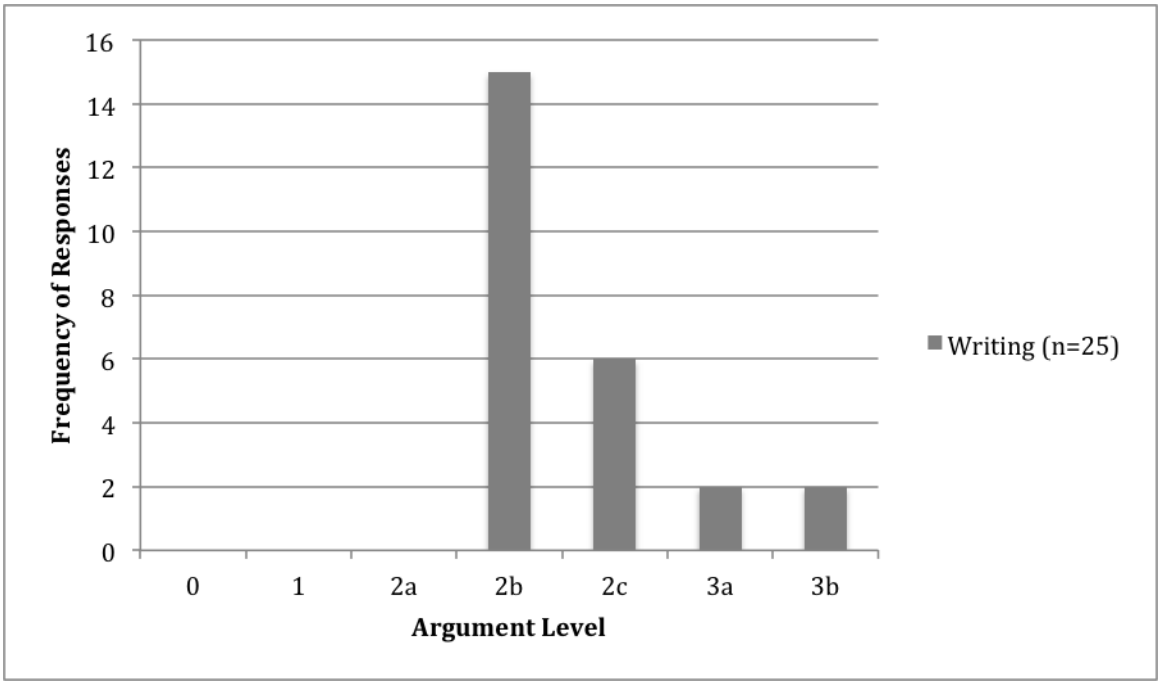

Figure 2. Frequency of individual written arguments at each level of the argument learning progression.

The frequency of the students' written arguments at each level of the learning progression is presented in Figure 2. Again we present the actual sample frequencies as opposed to percentages due to the small sample size. In looking at Figure 2, two themes emerge: 1) While the students regularly provided justifications in their written arguments, they tended to use inappropriate supports (e.g.

irrelevant, inaccurate, and/or non-supporting), and 2) Including a rebuttal was easier than including only appropriate justifications for written arguments. Evidence for each theme will next be discussed.

While the students regularly provided justifications in their individual written arguments, they tended to use inappropriate supports (e.g. irrelevant, inaccurate, and/or non-supporting). Similar to the students' oral arguments, the first theme is again immediately apparent: The students were providing justifications within their written arguments, however they tended to leverage irrelevant, inaccurate, and/or non-supporting supports (i.e. levels 2a, 2b, and 2c). While there are 15 responses at level 2b (i.e. justifications with mixed appropriateness), there are also 6 responses at level 2c (i.e. justifications with a mixture appropriateness as well as a rebuttal). In total, 84% of the students' written arguments included such inappropriate justifications. An example of these less sophisticated arguments is found within lesson 1. With the option to argue from his personal perspective about whether the Belo Monte dam should be built, Ben chose to take the pro-dam perspective. Specifically, he wrote:

> [T]here have been a lot of blackout in Brazil and those blackouts are caused by the powerline's not by the dams….In conclusion I think that the dam should be built because it brings electricity, prevents blackout.

This justification is relevant, but contradictory to his claim because if the power lines were the cause of the blackouts, then building a new dam would not prevent future blackouts. Rather, the justification appropriate for power line failures leading to blackouts should be on maintaining and/or preventing faults within the power line infrastructure. This piece of evidence is on-topic and, therefore, relevant; however it is contradictory to his claim that the dam should be built. As such, it does not function to support his argument.

Another example of students' use of inappropriate justifications is found within lesson 2 when the students were constructing written arguments around whether people should drink tap water or bottled water. One student, Jack, argued for tap water saying that the tap water "is tested 5-10 time a week, and it is the fifth (5^{th}) cleanest water in the nation". While Jack had data to support this, he continued to write, "this shows that bottle water companie's, do not taste bottle water as much as they taste tap water". This student's reasoning is based on a misconception about how the water quality was tested. More specifically, he understood testing to be in the form of taste tests as opposed to water quality tests. Therefore, he included a relevant, but inaccurate justification. As was also the case for the collaborative oral arguments, the students' individual written arguments tended to include justifications that were inappropriate in terms of accuracy, relevancy, and/or support.

Including a rebuttal was easier than including only appropriate justifications for individual written arguments. Unlike the students' collaborative oral arguments, a second theme is identified within Figure 2. Specifically, in terms of the students' sophistication in constructing individual written arguments, including a rebuttal was easier than including only appropriate justifications. In contrast, it appears that both relevancy and rebuttals could be difficult for students within their oral arguments because we found only one example at each of the levels. However, a larger sample would be necessary to further tease this apart. Regardless, rebuttals were present within levels 2c (i.e. justifications with mixed appropriateness as well as a rebuttal) and 3b (i.e. solely appropriate justifications with a rebuttal), and limiting justifications to being solely appropriate occurred at levels 3a (i.e. solely appropriate justifications) and 3b (i.e. solely appropriate justifications with a rebuttal). In Figure 2, we see that eight written arguments included a rebuttal (n=6 at level 2c; n=2 at level 3b), which represents 32% of the students' written arguments. This can be compared to the

written responses that included solely appropriate justifications, which were observed at levels 3 (n=2) and 4 (n=2) and represent 16% of the students' individual written arguments. Having a larger percentage of written arguments providing a rebuttal (32%) than solely appropriate justification (16%) suggests that including solely appropriate justifications was more difficult for the students than including a rebuttal within their written arguments. But, because frequency does not necessarily determine difficulty (i.e. many students could have mastered a difficult task), we next provide evidence to support this assertion.

The ordering of the levels of sophistication, in regards to the difficulty of providing rebuttals as compared to solely appropriate justifications, will be illustrated by comparing writing from two students—Bailey and Alfred. Bailey's argument was identified as being at level 3b (i.e. solely appropriate justifications as well as a rebuttal). More specifically, in addressing whether people should drink tap water or bottled water, Bailey made a strong argument that "people should drink tap water because it is clean and safe most of the time" by employing multiple relevant justifications as well as the following rebuttal to a bottled water company's advertisement:

> Nestle water company said Bottled water is the most environmentally responsible consumers product in the world. They are only saying that because they want people to think that tap water is bad so then people will buy a lot more bottle water.

We see that Bailey weakened the counter-claim by critiquing the appropriateness of the source and their underlying motivation, which served to further justify her claim that people should drink tap water. However, other students at the 2c level (justifications with mixed appropriateness as well as a rebuttal), including Alfred, provided rebuttals that were of equal quality. For instance, in arguing that the Belo Monte dam should not be built, Alfred rebutted:

> A person that disagrees with me would argue that the Dam is a good thing because it produces electricity for the country of Brazil. I would respond by saying the electricity doesn't even go to the people in Brazil that need it, it goes to the aluminum smolting factories that produce even more pollution to our world.

While Bailey critiqued the appropriateness of the source, we see that Alfred critiqued the appropriateness of the counter-reasoning. Alfred's rebuttal is no less sophisticated than Bailey's, although Bailey's argument as a whole was classified as being more sophisticated. What distinguished these two arguments, therefore, was the distinction that Bailey supported her claim with only appropriate justifications whereas Alfred also included inappropriate justifications. For instance, Bailey provided multiple lines of relevant support when arguing, *"people should drink tap water because it is clean and safe most of the time"*, such as her first justification centered on water quality tests:

> Scincetice [Scientists] studied Clevland's tap water and fiji's bottled water and found out that Clevland's tap water is cleaner than fiji's bottled water. When they were doing the studing they found out that fiji's water has Arsenic in it.

This justification aligns with and supports her claim, and, as such, is appropriate. She replicated this ability to limit her justifications based on appropriateness several times. For instance, she also provided the following marketing research results as further justification:

> They were also doing test were people came up and tried 6 different kinds of water and most of the people said that cup number 3 was the best but it was tap water. That means people like tap water better than bottled water.

Not only does this justification provide additional support for her claim that people should drink tap water, but it preempts a counter-argument concerning the taste. While Bailey displayed an ability to limit her justifications based on appropriateness, Alfred's justifications included some inaccuracies:

> [T]his dam shouldn't be built because of … the methane gas the reservoir produces. … The reservoir next to the dam will produce methane from decompose species. This makes our world over-heat. Most of Antarctica will melt if methane is produced. It can also make people severely sick and they can die.

While we see that Alfred did provide an accurate explanation of how the methane would be produced, his line of thinking is exaggerated. For instance, while the accumulation of greenhouse gasses, including methane, is associated with climate change and climate change is associated with the melting of icebergs in Antarctica, most of Antarctica is not going to melt from the methane produced from this reservoir. Similarly, while long-term exposure to high concentrations of methane gas can lead to death, people will not become severely sick or die from the methane produced from this reservoir. Bailey and Alfred, therefore, differed in their ability to limit their justifications based on appropriateness despite having rebuttals of similar quality. This supports our argument that appropriateness was more difficult than rebutting the opposing claim within the written arguments.

Comparison across students' collaborative oral and independent written arguments

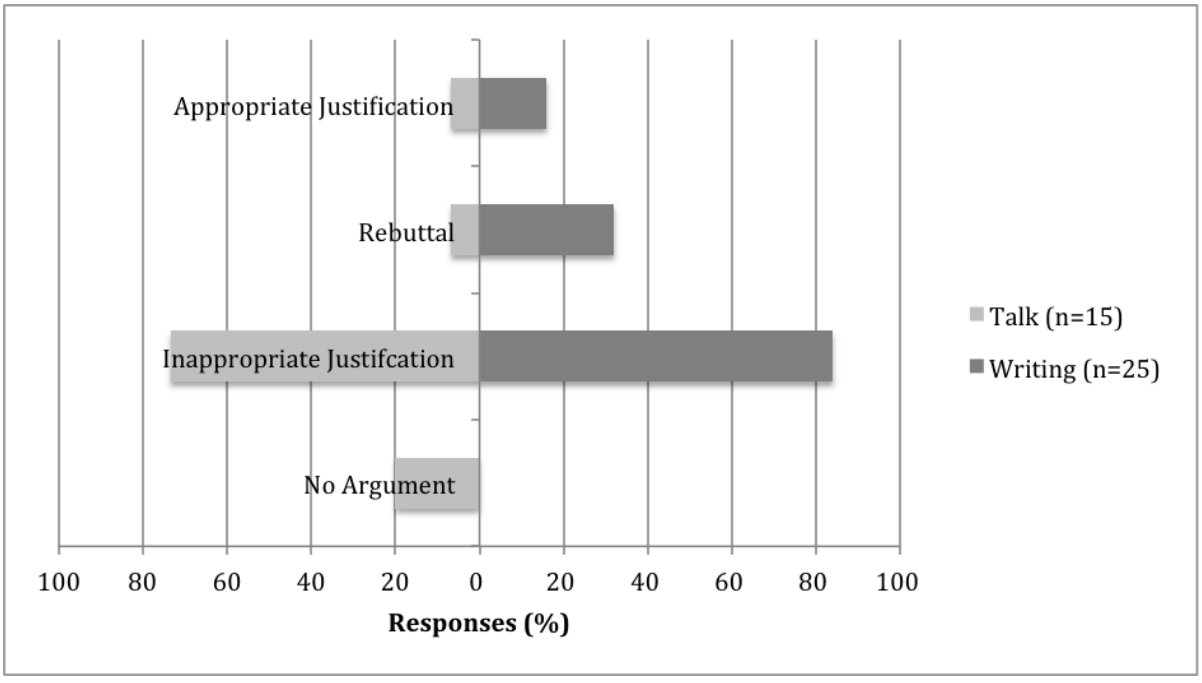

Figure 3. Percent of arguments at important markers within the learning progression organized by modality.

While we recognize that the oral arguments were at the group level whereas written arguments were at the individual level, we proceed with the comparison with the caveat that the modalities have different numbers of participants contributing to the final product (i.e. oral: group; writing: individual). Figure 3 presents both students' collaborative oral and individual written arguments at important categories within the learning progression. No argument represents arguments that were classified as level 0 (i.e. no claim) or level 1 (i.e. claim).

Inappropriate justifications occurred at levels 2a (i.e. only inappropriate justifications), 2b (i.e. justifications with mixed appropriateness), and 2c (i.e. justifications with mixed appropriateness as well as a rebuttal). Rebuttals occurred at levels 2c (i.e. justifications with mixed appropriateness as well as a rebuttal) and 3b (only appropriate justifications as well as a rebuttal). Appropriate justifications occurred at levels 3a (i.e. only appropriate justifications) and 3b (i.e. only appropriate justifications as well as a rebuttal). Because there was a difference in sample sizes, we present these responses as percentages. Additionally, it should be noted that some arguments fell into two categories. For instance, level 2c (i.e. justifications with mixed appropriateness as well as a rebuttal) was counted in the irrelevant category as well as the rebuttal category. Consequently, the summed category percentages within each modality can exceed 100%.

From Figure 3, we see that the majority of responses did include arguments in which the students included some justification for their claims. However, both the collaborative oral and individual written arguments tended to include irrelevant, inaccurate, and/or relevant-contradictory justifications (i.e. level 2a, 2b, or 2c). Therefore, the difference between the two modalities lies in the arguments that did not include inappropriate justifications. More specifically, according to our learning progression, arguments that include a rebuttal (i.e. levels 2c and 3b) or only appropriate justifications (i.e. levels 3a & 3b) are more sophisticated than those that include inappropriate justifications (i.e. levels 2a and 2b). In looking at Figure 3, we see that more students provided these higher-level arguments based on appropriate justifications and rebuttals in writing (appropriate: 16.7%; rebuttals: 32.7%) as compared to in talk (appropriate: 6.7%; rebuttals: 6.7%). Specifically, there were 2.6 times more arguments that included rebuttals and 4.9 times more arguments based on solely appropriate justifications in individual's writing as compared to the collaborative talk.

In comparison, the arguments that were less sophisticated than those that contained inappropriate justifications were those that did not make an argument because they did not provide a claim (i.e. level 0) or the claim was not justified (i.e. level 1). Whereas this did not occur in writing, the non-arguments occurred in 20% of the oral arguments. Additionally, we did not find any instances of level 1 non-arguments, therefore all of our students' non-arguments were a result of not providing a claim. An example of this is found within lesson 3 in which each group presented an argument on a different socioscientific issue for the community fair. The non-argument occurred within the group that addressed the following question: Why is it important to protect our oceans and rivers? Instead of presenting and supporting their claim, these students provided a summary of potential justifications. This is exemplified by the following quote:

> [T]his is a diagram [pointing to a table with 3 columns] showing micro-in, micro-invertebrates and their tolerance to pollution. So this picture pretty much shows the different kind of creatures, which are these, um, and which, and which creatures are sensitive, which is this column, moderate, which is this column, and tolerate, or tolerant, which is, what that is supposed to say.

While the students did discuss how micro-invertebrates are more and less tolerant to water quality, they never directly linked this back to why it is important to protect our oceans and river. As opposed to providing a persuasive argument that includes support for their claim, they described different diagrams and graphs that presented the information and data. This same phenomenon occurred again when they explained their global warming data: "And then, this is the global warming data, like, and it goes on for the years till 2010. First it was low, then high, then low, then high, then low, then high, then it went up dramatically." Again, they simply described the data as opposed to using it as a justification.

The previous examples from one group's collaborative oral descriptions (as opposed to arguments) can also be compared to the group member's individual written arguments. For instance, Bailey's argument addressing whether people should drink tap water or bottled water (i.e. lesson 2), which was presented in the previous section, employed multiple relevant justifications as well as a rebuttal (i.e. level 3b). Additionally, the other group members had individual written arguments at levels 2b and 2c for the first lesson. This suggests each group member knew to justify his or her claim in writing despite not doing so in talk. Moreover, the same trend occurred for the two other groups who made oral arguments at level 0. Therefore, this suggests that the most sophisticated arguments were in individual's writing whereas the least sophisticated arguments occurred in the collaborative talk.

DISCUSSION

Our findings suggest there are both similarities and differences between students' abilities to construct individual written as compared to collaborative spoken arguments. In terms of similarities, arguments in both modalities tended to include justifications for their claims; however, those justifications were often irrelevant, inaccurate, or non-supporting. In writing, the students' individual responses were more likely to include rebuttals than to limit their responses to solely appropriate justifications. Moreover, the students were more sophisticated in individually writing arguments as compared to collaboratively presenting oral arguments. We will next use the extant literature to discuss potential reasons for each of the themes.

Theme 1: While the students regularly provided justifications, they tended to use inappropriate supports (e.g. irrelevant, inaccurate, and/or non-supporting)

Although it is often the case that existing classroom norms tend to constrain students' engagement in scientific argumentation (Berland & Reiser, 2009; Driver et al., 2000; Jiménez-Aleixandre et al., 2000) and scientific argumentation is challenging for some middle school (McNeill et al., 2006) and high school students (Sandoval, 2003), overall we found that these middle school students were constructing both collaborative oral and individual written arguments with a range of abilities. While the middle school students mostly knew they were supposed to justify their arguments, it was not a practice they had mastered as evidenced by their routine use of irrelevant, inaccurate, and/or non-supporting justifications across both the individual written and collaborative oral arguments. This is not dissimilar to McNeill and Krajcik's (2007) finding that students struggle with selecting appropriate data to use as evidence. While McNeill and Krajcik (2007) were more specific as to the type of justification, we both conclude that students find the consideration of appropriateness as well as accuracy problematic when justifying an argument.

The findings of our study raise an important question as to why students were marshaling inappropriate justifications. Some researchers might attribute this to students' content knowledge (McNeill et al., 2006; McNeill & Krajcik, 2007; Osborne et al., 2004). However, others suggest that the socioscientific contexts, such as ours, reduce the content load by allowing students to reference their own experiences as well as ethical (Osborne et al, 2004; Patronis, Potari, & Spiliotopoulou, 1999), moral, and political influences (Cavagnetto, 2010). Our results seem to support this as students used such non-empirical forms of justifications when constructing their

socioscientific arguments; however these non-empirical influences could have also made it more difficult for students to limit their justifications to those that were solely appropriate. For example, in addressing the building of the Belo Monte dam in lesson one, the students had to consider moral issues, such as the rerouting of the Xingu River, which the Kayapo Tribe rely on fishing, drinking, and transportation. They also had to consider political issues including whether building the dam would solve the electricity shortage issues for the country as well as ethical decisions in regards to the dams negative impact on the ecosystem. Therefore, the students were required to evaluate the appropriateness of justifications in multiple dimensions (i.e. moral, ethical, political, scientific) as opposed to just within the scientific domain (Sadler & Zeidler, 2005), and applying evaluative criteria across these dimensions could have made it more difficult. Similarly, Kolstø (2001) found that students have difficulty evaluating information or scientific evidence within socioscientific issues, which he attributed to their acceptance with limited consideration of source, content, or stakeholders' interest in the issue, or inconsistencies in applying these evaluative criteria. Perhaps, then, differences in what count as evaluative criteria within these multiple dimensions (moral, ethical, political, and scientific) make it more difficult for students to evaluate between them. In comparison, a multivariate scientific issue would rely on only scientific epistemologies to determine what should be evaluated and how it should be evaluated (e.g. providing data as evidence). Clearly, the reasons behind how students evaluate the appropriateness of justifications in terms of accuracy, support, and relevancy within socioscientific contexts and between socioscientific and scientific contexts necessitates further exploration.

Theme 2. Including a rebuttal was easier than including only appropriate justifications for individual written arguments

It is plausible that the multiple influences (i.e. moral, ethical, political and scientific) involved in socioscientific arguments could have increased the difficulty of evaluating the appropriateness of justifications to a degree that made it more difficult than constructing rebuttals. But, it is also possible that the task of constructing rebuttals was made easier based on instructional moves within the learning environment. Specifically, the teacher framed his expectation for each of the lessons in terms of being persuasive or convincing. Berland and Hammer (2012) suggest this can support students' engagement in argumentation as well as influence the students' expectations for what should be constructed. Therefore, making the persuasive expectations clear in both lessons that included individual written arguments could have supported the students in constructing arguments that included rebuttals. But, because the majority of the written rebuttals came within the Belo Monte dam lesson (Lesson 1: n=7; Lesson 2: n=1), we suggest there was a difference in how the teacher framed his expectations for the two written argument lessons: The number of perspectives the students were required to consider. Specifically, in the Belo Monte dam lesson the students were originally separated into stakeholder groups—the power company, hydrologists, climate scientists, ecologists, and the Kayapo Tribe—and were required to represent the perspective of the group to which they were assigned as well as keep track of the other groups' claim and two pieces of evidence. Thus, the teacher framed the lesson to support the students understanding of multiple audiences despite there only being two claims (i.e. pro or con). Similarly, the tap water versus bottled water lesson only had two possible claims; however the students did not consider this issue from different audience perspectives. For instance, they could have considered how bottled water companies, public water treatment facilities, emergency response agencies, and the

public consumers viewed this issue. This suggests that designing the learning environment in such a way that the students recognize the perspective of different audiences supports the students in constructing rebuttals. Similarly, Kuhn and Udell (2007) found that students tended not to construct rebuttals unless explicitly prompted, which suggests that the students knew how to rebut, but did not recognize the expectation to do so. It is also possible that more rebuttals occurred during the Belo Monte Dam's written arguments (i.e. lesson 1) as compared to the Tap Water versus Bottled Water written arguments (i.e. lesson 2) because the Belo Monte Dam writing happened after the oral arguments had been completed. While this interaction is possible, we argue that the way the teacher developed the lesson and framed the instructions also contributed to the increased number of rebuttals.

Theme 3. The most sophisticated arguments were in individual writing while the least sophisticated arguments were in collaborative talk

While Berland and McNeill (2010) concluded that the students' verbal arguments were stronger than their written products, we found the students' individual written arguments to be stronger than their collaborative oral arguments. This discrepancy could very well be a result of samples with varying abilities or distinctions within the nature of the questions, but perhaps more pertinent to this discussion are differences in what was measured and how it was measured.

In terms of what was measured, both our study and Berland and McNeill (2010) regarded writing as an argument product; however, Berland and McNeill (2010) analyzed the process of constructing and debating oral arguments whereas we looked at the final product of the oral argumentation. Therefore, embedded in their comparison between the modalities are differences in both form (i.e. oral: process; writing: product) and number of students contributing to the final product (i.e. oral: group; writing: individual). Consequently, the three independent variables—modality, form, and number of students contributing—make it difficult to reliably attribute differences to only modality. In comparison, we attempted to control the form of argument variable by looking at products within both modalities, but were still left with a difference in the number of participants constructing the product (i.e. oral: group; writing: individual). As such, the discrepancy in findings between the studies could be related to how each study handled the form (i.e. process versus product) of the collaborative oral arguments.

Perhaps, the lower levels of sophistication in students' collaborative oral arguments within our study could also be related to affordances specific to each modality. For instance, when comparing writing and talk some have found that writing was more abstract, objective, explicit, and more rigorous (Goody, 1994), which could be because the "written language, stabilized on paper, invites kinds of reflection not so natural to oral exchanges" (Tishman & Perkins, 1997, p.371). This is also supported by research within science education. Specifically, Rivard and Straw (2000) found that "writing is an important discursive tool for organizing and consolidating rudimentary ideas into knowledge that is more coherent and well-structured" (p. 586). In our study, the complex network of socioscientific relationships involved in this lessons might have resulted in stronger written arguments because having their ideas recorded on paper provided the opportunity for the students to reflect and refine their ideas as they wrote their responses. This discrepancy, therefore, might be accounted for by the complexities involved in the socioscientific issues that required students to bring together the political, social, ethical, and scientific influences. Additionally, Rivard and Straw (2000) suggest that discussion helps peers to share, clarify, and distribute knowledge. Perhaps this is why we found instances of non-arguments within the oral modality, but not the

written modality. Regardless, this discussion provides grounds for additional research to tease apart the reasons why the oral arguments were less sophisticated than the written arguments.

CONCLUSION

We acknowledge that our exploratory study is limited by a small sample size and that there is a difference in the number of participants who constructed the argument products in individual writing as compared to collaborative oral presentations. We also acknowledge that our proposed learning progression is still in a theoretical phase and that a larger sample size is necessary to confirm or refute the ordering of the levels. Despite these limitations we believe our methodology and findings not only build upon the extant literature base, but also raise pertinent questions for future research. Specifically, the application of the argumentation learning progression to code the students' argument products was a productive and beneficial methodology, which provided the opportunity to make comparisons across the collaborative oral and individual written arguments. This is meaningful because, to date, little research has explored similarities and differences between these modalities in the scientific argumentation literature.

The results suggest that the sophistication of students' individual written and collaborative oral argument products are sometimes, but not always the same. For instance, regardless of modality, the consideration of appropriate justifications was challenging for students, however it is not clear why this is the case. Further research, therefore, could explore not only why students struggle with this construct, but also compare how students with high and low abilities in evaluating appropriate justifications actually go about the process. Additionally, our results suggest that, if students are provided with appropriate contexts, rebuttals in individual's writing may not be as difficult as previous research has suggested (Osborne et al., 2004). However, due to the sample size, we were not able to determine whether this is equally difficult in both modalities. Future research utilizing large samples could further explore these relationships as well as the reasons why students' abilities with collaborative oral arguments were less refined than with individual written arguments. With this knowledge, we could, in turn, better support students' understanding of argumentation by explicitly addressing aspects that are specific to writing and speaking.

REFERENCES

Aikenhead, G. S. (2005). Science-based occupations and the science curriculum: Concepts of evidence. *Science Education, 89*(2), 242-275. doi:10.1002/sce.20046

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the Web with KIE. *International Journal of Science Education*, *22*, 797-817. doi:10.1080/095006900412284

Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. *Journal of Research in Science Teaching*, *49*(1), 68-94. doi:10.1002/tea.20446

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765-793. doi:10.1002/sce.20402

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education, 93*(1), 26-55. doi:10.1002/sce.20286

Berland, L. K., & Reiser, B. J. (2011). Classroom communitites adaptations of the practic of scientific argumentation. *Science Education*, *95*(2), 191-216. doi:10.1002/sce.20420

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research*, *80*(3), 336-371. doi:10.3102/0034654310376953

Clark, D. B., & Sampson, B. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching*, *45*(3), 293-321. doi:10.1002/tea.20216

Common Core State Standards Initiative. (2010). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Retrieved from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Dawson, V. M., & Venville, G. (2009). High school students' informal reasoning and argumentation about biotechnology: An indicator of scientific literacy? International *Journal of Science Education*, *31*(11), 1421-1445. doi**:**10.1080/09500690801992870

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

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades k-8.* Washington D.C.: National Academy Press. http://www.nap.edu/catalog/11625.html

Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching, 50*(2), 209-237. doi:10.1002/tea.21076

Furtak, E. M., Thompson, J., Braaten, M., & Windschitl, M. (2012). Learning progressions to support ambitious teaching practices. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future* directions (pp. 405-433). Rotterdam, The Netherlands: Sense Publishers. doi:10.1007/978-94-6091-824-7_17

Goody, J. (1994). Entre l'oralite´ etl'e´criture. Paris: Presses universitaires de France.

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power*. Pittsburgh: University of Pittsburgh Press. doi:10.1234/12345678

Jiménez -Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). 'Doing the lesson' or 'doing science': Argument in high school genetics. *Science Education, 84*(3), 287-312. doi:10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F

Kelly, G. J., & Greene, J. (1998). The social nature of knowing: Toward a sociocultural perspective on conceptual change and knowledge construction. In B. Guzzetti & C. Hynd (Eds.), *Perspectives on conceptual change: Multiple ways to understand knowing and learning in a complex world* (pp. 145-181). Mahway, NJ: Lawrence Erlbaum.

King, P. M., & Kitchener, K. S. (2004). Reflective judgment: Theory and research on the development of epistemic assumptions through adulthood. *Educational Psychology*, *39*, 5-18. doi:10.1207/s15326985ep3901_2

Kolstø, S. D. (2001). To trust or not to trust, … —pupils' ways of judging information encountered in a socio-scientific issue. *International Journal of Science Education, 23*, 877-901. doi:10.1002/sce.1011

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

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child development*, *74*(5), 1245-1260. doi:10.1111/1467-8624.00605

Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning, 13*(2), 90-104. doi:10.1080/13546780600625447

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. doi:10.1111/j.1467-8624.2008.01190.x

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation and evidence and abilities to construct arguments over the school year. *Journal of Research in Science Teaching, 48*(7), 793-823. doi:10.1002/tea.20430

McNeill, K. L., Corrigan, S., Barber, J., Goss, M., & Knight, A. M. (2012, March). *Designing student assessments for understanding, constructing and critiquing arguments in science.* Poster presented at the annual meeting of the National Association for Research in Science Teaching, Indianapolis, IN.

McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. Lovett & P. Shah (Eds.), *Thinking with data: The proceedings of the 33rd Carnegie symposium on cognition.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McNeill, K. L., & Krajcik, J. (2012). *Supporting grade 5-8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing.* New York, NY: Pearson Allyn & Bacon.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences*, *15*(2), 153-191. doi:10.1207/s15327809jls1502_1

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203-229. doi:10.1002/sce.20364

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

Miles, M., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook (2nd edition).* Thousand Oaks, CA: Sage.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994-1020. doi:10.1002/tea.20035

Patronis, T., Potari, D., & Spiliotopoulou, V. (1999). Students' argumentation in decision-making on a socio-scientific issue: Implications for teaching. *International Journal of Science Education, 21*, 745-754. doi:10.1080/095006999290408

Rivard, L. P., & Straw, S. B. (2000). The effect of talk and writing on learning science. An exploratory study. *Science Education, 84*, 566-593. doi:10.1002/1098-237X(200009)84:5<566::AID-SCE2>3.0.CO;2-U

Rogers, S., Busch, K. C., & Berland, L. K. (2012, March) *Variation in how individuals argue about scientific and socioscientific questions.* Paper presented at the NARST 2012 Annual International Conference, Indianapolis, IN.

Ryu, S., & Sandoval, W. (2008). *Interpersonal influences on collaborative argument during scientific inquiry*. Paper presented at the Paper Presented at the American Educational Research Association (AERA), March 24-29. doi:10.1002/sce.21006

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*, 513-536. doi:10.1002/tea.20009

Sadler, T. D., & Donnelly, L. A. (2006). Socioscientific argumentation: The effects of content knowledge and morality. *International Journal of Science Education, 28*(12), 1463-1488. doi:10.1080/09500690600708717

Sadler, T. D., & Fowler, S. R. (2006). A threshold model for content knowledge transfer for socioscientific argumentation. *Science Education, 90*(6), 986-1004. doi:10.1002/sce.20165

Sadler, T. D., & Zeidler, D. L. (2005). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education, 89*(1), 71-93. doi:10.1002/sce.20023

Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education, 92*(3), 447-472. doi:10.1002/sce.20276

Sampson, V., & Clark, D. (2009). A comparison of the collaborative scientific argumentation practices of two high and two low performing groups. *Research in Science Education, 61*(1), 63-97. doi:10.1002/sce.20306

Sampson, V. Grooms, J., & Walker, J. P. (2010). Argument-driven inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education, 95*(2), 217-157. doi:10.1002/sce.20421

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences, 12*, 5-51. doi:10.1207/S15327809JLS1201_2

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction, 23*(1), 23-55. doi:10.1207/s1532690xci2301_2

Sandoval, W. A., & Reiser, B. J. (1997). *Evolving explanations in high school biology*. Paper presented at the annual meeting of the American Educational Research Association, Chicago. doi:10.1002/sce.10130

Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *Journal of the Learning Sciences*, *12*(2), 219-256. doi:10.1207/S15327809JLS1202_3

Schweingruber, H., Keller, T., & Quinn, H. (Eds.). (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.

Tishman, S., & Perkins, D. (1997). The language of thinking. *Phi Delta Kappan, 78*, 368-374.

Varelas, M., Pappas, C. C., Kane, J. M., & Arsenault, A. (2008). Urban primary-grade children think and talk science: Curricular and instructional practices that nurture participation and argumentation. *Science Education, 92*, 65-95. doi:10.1002/sce.20232

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*, 35-62. doi:10.1002/tea.10008