Elementary Students' Coordination of Claims and Evidence in Science and History

by

John-Carlos Marino

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Educational Studies)
In the University of Michigan
2020

Doctoral Committee:

Professor Elizabeth A. Davis, Chair
Associate Professor John Carson
Associate Professor Chauncey Monte-Sano
Professor Annemarie Sullivan Palincsar

John-Carlos Marino

jcmarino@umich.edu

ORCiD: 0000-0002-7313-5657

DEDICATION

To Alexander who taught me efficiency,

and to Juliette who forced it.

ACKNOWLEDGEMENTS

This dissertation was made possible by the help and support by a great many people who I would like to thank here. First, thank you to Matthews Elementary, Ms. Fischer, and the wonderful third-graders that took part in this study. They welcomed me into their school and classroom and gave me an opportunity to conduct my study. More importantly, the students did great work and demonstrated the possibility of ambitious instruction.

Next, I would like to thank my committee for their guidance, feedback, and encouragement through this process. To John Carson, who pushed me to think beyond my comfort zone and reminded me to challenge my assumptions about my own work. To Annemarie Palincsar, who gave me my first glimpse of what it meant to be a careful and considerate scholar during ELECTS meetings. To Chauncey Monte-Sano who graciously welcomed me to the world of social studies and took the extra time to make sure I had the preparation I needed. To Leah Bricker, who also served on my committee during my proposal process. Your enthusiasm for my work helped me remember to enjoy what I was doing and that it really was worth it. And to Betsy Davis, my committee chair and advisor. Thank you, Betsy, for your incredible support and listening from before I even decided to pursue my PhD. You have gone above and beyond what anyone could expect out of the best advisor, and have been a model researcher, teacher, and mentor.

I would also like to thank my friends and colleagues at the School of Education who gave me advice, commiserated with me, and listened as I talked out loud trying to figure things out. To Anna Arias, who guided me through many of the ins and outs of being a student at Michigan. To

Sylvie Kademian who was a great classmate and always had a way of grounding research in real life. To Adam Bennion whose help coding was indispensable. To Amber Bismack, who pushed me to crystalize my own understanding of argumentation. And to Jason Moore, for his support developing texts and, more importantly, making me take regular breaks where we didn't talk shop.

I also could not have come to Michigan, much less completed a dissertation, without the support of my friends and family. Thank you to Adam Streed for your thoughtful perspective when I needed to think about work and your friendship when I didn't. To my sister, Monica, who laughs at my jokes and makes me feel funny. Thank you to Abby and Steve, who graciously let me take their daughter to the hinterland and still did so much to help me finish. And I am so grateful to my parents, Carlos and Maureen, who have always supported me no matter what I do. Your confidence in me, not to mention cooking and babysitting, helped make this possible.

Finally, thank you to my amazing wife Tara. None of this would have been possible without you. You have been my best friend and biggest cheerleader. You believed in me even when I didn't and made sacrifices to help me pursue this degree. And, as I'm sure you'll never let me forget, you knew I wanted to go into education even before I did and helped me find my way there. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

Argumentation is a key practice in many disciplines and as such is emphasized as an important skill in K-12 education. Despite its relevance to all content areas, little is known about how the same students engage in argumentation in different contexts. Furthermore, there is little research examining how students who are novices in argumentation use evidence to develop and support a claim. This dissertation addresses this issue by studying how students new to the practice engaged in evidence evaluation and argumentation in science and history and looking at their understanding of evidence and argumentation across disciplines.

This case study examines a classroom of third-grade students who had no prior instruction in evidence analysis or argumentation. They engaged in two evidence-based investigations, one in science and one in history. Students were given an investigation question and worked together in small groups to analyze evidence and develop an argument supporting their answer to the question. Data was collected in the form of video recordings of the small group work and students' written responses on graphic organizers that supported evidence analysis and argumentation. Additionally, they completed pre- and post- surveys, in small groups, asking about their beliefs about evidence in science and history and the work of each discipline.

In the surveys, students demonstrated a generally limited sense of what evidence is and a stronger understanding of scientific practice in comparison with historical practice. When analyzing evidence in the investigations, students did more to identify details and summarize the texts than to make inferences about them, and several interpreted the sources in ways that confirmed their initial ideas. In the argumentation portion, students had difficulty articulating the

reasoning behind their evidence, and included evidence that was either inaccurate or did not support their claim. In both tasks, students did not engage in the social aspects of argumentation, namely discussion, and often just took turns stating their ideas.

Despite these challenges, the students demonstrated many strengths considering their age and unfamiliarity with the practices. They were able to find details that were relevant to their claim and made inferences about the texts that were often accurate and identified points not explicitly stated. In addition, they successfully incorporated multiple sources and used language markers, like "because," to indicate they were providing justification.

These findings suggest that while students of this age level may not have the background knowledge of the practices of evaluating evidence and constructing argument, they do have nascent abilities in both practices. With more support and stronger scaffolds, these students would have likely been very successful. Further research should continue to look at the reading strategies students use as they evaluate evidence and the reasoning students use to link evidence and claims, even if they do not explicitly state it. The methodological implications of the study are discussed, including the importance of using video of student work to analyze their argumentation practice, and the possibilities for future research. The findings also suggest instructional supports such as discussion protocols, graphic organizers, and exposure to accurate representations of historical and scientific practice.

CHAPTER 1

Introduction

In recent years, the practice of argumentation has received increased attention in the education and education research community. There is not only substantial body of research on how students learn and engage in argumentation, but argumentation is also prominent in recent standards documents from the four major K-12 academic content areas, mathematics, English language arts (ELA), and social studies[1] (National Council for the Social Studies (NCSS), 2013; National Governors Association Center for Best Practices (NGA Center) & Council of Chief State School Officers (CCSSO), 2010; NGSS Lead States, 2013). This is unsurprising as argumentation is the means through which practitioners in those disciplines (e.g., mathematicians, historians) propose and support their conclusions, and construct knowledge within their fields.[2] And, because argumentation requires the kind of careful thinking needed to develop ideas and justify beliefs that is essential to activities both in and out of academic settings, it is a worthwhile skill to teach to children.

[1] I use the C3 Framework's (National Council for the Social Studies (NCSS), 2013) definition of social studies which encompasses the content areas of history, civics, geography, and economics.

[2] I distinguish between discipline, field, and content area. Discipline is the most inclusive category and describes general areas of study as practiced by those with special training in that discipline (e.g., mathematicians, scientists). A content area is the study of a discipline within a school context and includes the knowledge and skills within that discipline that are typically taught to students. A field is an area of study within a discipline or content area. This includes broad fields such as biology within science, and more narrow sub-fields such as genetics within biology.

Despite the robust body of research on argumentation and the fact that it is so important to all of the academic disciplines, there is little research about how students learn argumentation across content areas and how their understanding of argumentation in one content area interacts with or affects their understanding of it in another. Although there are certainly important disciplinary differences in constructing arguments, there are still many aspects of argumentation shared across disciplinary lines. As such, it is important to understand if students see argumentation as a generic or discipline-specific activity, the different approaches they might take to arguing in different disciplines, and the differences in their ability to construct arguments in different disciplines.

Besides the connections between argumentation in different disciplines, there are other important reasons to look at students' learning across content areas. First, students typically have a very disjointed experience during the school day (Stevens, Wineburg, Herrenkohl, & Bell, 2005). They make several transitions between content areas and there is often little connection made between what they learn in their different classes. This is compounded by the fact that the lines between content areas can sometimes seem arbitrary and students may be confused by why one thing counts as "social studies" but not "ELA." Second, understanding student learning across content areas is especially important for elementary teachers who are typically teach all content areas and would be the teachers introducing students to the practice of argumentation. To be most effective, they need to have a strong understanding of how students learn across content areas, how their instruction in one area can affect another, and where they can leverage learning from one content area in another.

Although teachers' understanding of argumentation across disciplines is important, the focus needs to first be on students and how they understand the connections and engage in argumentation across content areas. While the content areas may typically be segregated for students, there are “natural” pairings that tend to occur when attempts are made to link them. Most often, mathematics and science are combined into what is known as STEM (science, technology, engineering, and math) education. This is evident from the large number of STEM-focused schools, grants, and programs, along with a body of research on STEM education. On the other hand, social studies and ELA are sometimes addressed at the same time because reading texts and producing writing are particularly important to history and civics, which usually fall under the social studies umbrella. Science and social studies, meanwhile, are rarely tied together. Despite their apparent differences, however, the two content areas have an important connection. This is especially true when comparing science with history, a content area contained within social studies. Knowledge in both disciplines is often tentative, meaning that theories and conclusions about scientific phenomena and historical events are often revised and adjusted based on new evidence. As a result, knowledge in these disciplines is constructed by building arguments and counter-arguments on the basis of evidence and explanations of why that evidence supports a conclusion. To engage in the disciplinary practices, then, students must understand how to construct and evaluate arguments, and recognize that these arguments can change over time as new evidence is uncovered. Furthermore, the two content areas also share a connection in the context of schooling, as they are often ignored in favor of teaching math and ELA which are heavily tested and thus “matter” more (Blank, 2013; Center on Education Policy, 2008; P. G. Fitchett, Heafner, & Lambert, 2014).

In the same way that there are few interdisciplinary connections made during schooling, there is relatively little research about learning across disciplinary lines. One exception might be STEM (science, technology, engineering, and mathematics) education, which is often framed as interdisciplinary. But even then, STEM is often shorthand for either only science or engineering, where mathematics often plays an auxiliary role as a tool used in science. In terms of argumentation, there is even less research across content areas, in particular between science and social studies. Aside from two studies (Herrenkohl & Cornelius, 2013; Iordanou, 2010), there are no major research articles comparing students' argumentation in science and history or social studies.

Furthermore, there is also relatively little research about argumentation among elementary-aged children. Although there are a few examples of studies involving elementary students (Fillpot, 2012; McNeill, 2011), research in this area more typically examines how middle- and high-school students engage in and learn argumentation. Consequently, there is little known about younger students' capability to engage in the practice or how to lay a foundation so that students are prepared for the more sophisticated argumentation they need to do in later grades.

Understanding how elementary students learn and engage in argumentation across science and social studies is an important concern, for both researchers and practitioners. The research in this area is scant and filling this gap will provide guidance to researchers developing interventions and models of student learning related to argumentation. It is also particularly useful to elementary teachers who have limited time to teach these two subjects. Knowing more about student learning across science and social studies may provide teachers, along with others such as curriculum developers and teacher educators, with ways of using instruction in science to support learning in social studies and vice versa. As such, a strong research base would help these constituents to find ways to design instruction in these areas as efficiently and effectively as possible. To these ends, this study aims to address two gaps in the literature. First, it targets younger elementary students than are typically studied to give insight into what it looks like when students begin to learn argumentation. Second, it directly compares students' argumentation in science and history by engaging them in parallel investigations in which they evaluate and use evidence to construct a claim.

Research Questions and Study Overview

In order to learn more about how students construct arguments and use evidence in science and history and how they understand the relationship between argumentation in the two fields, this research engaged students in tasks requiring them to use evidence to construct arguments in each subject area. The study describes the ways that students evaluated evidence and the reasoning they used to coordinate that evidence with the claims they made in each subject. Using a qualitative approach, I compare how students used evidence in each discipline to identify commonalities and differences in how they reasoned with evidence in each field. Students were also asked explicitly about what they perceived the relationship between argumentation in science and history to be. The research questions guiding this study are:

1. How do students interpret and make sense of evidence in science and history?
    A. What are the similarities and differences between these strategies in science and history?
2. How do students use evidence to construct arguments in science and history?
    A. What strategies for using evidence do students employ when constructing arguments?
    B. What are the similarities and differences in these strategies in science and history?
    C. How do students prioritize evidence when engaging in argumentation in science and history?

3. How do students coordinate evidence with claims when engaging in argumentation in science and history?
    A. What kind of reasoning do students use to coordinate evidence with claims?
    B. What are the similarities and differences between this activity in science and history?
4. How do students understand the use of evidence in arguments in science and history, and the relationship between the two?
    A. What connections, if any, do they see between their work engaging in argumentation in science and history?
    B. What connection, if any, do they see between their work engaging in argumentation and the work of scientists and historians?
    C. Is there a connection between how students understand the work of scientists and historians and how they coordinate claims with evidence in each subject area?

The analyses showed that students tended to have similar strategies when reading sources and constructing arguments. At the same time, however, they appeared to have a stronger understanding of scientific practice than of historical practice. This was evident in the pre- and post- surveys in which they were able to describe evidence and activities in science with more detail than in history. This difference in knowledge also appeared to have consequences for how they created their arguments. For instance, they were slightly more likely to make inferences based on the evidence in the science investigation compared with the history investigation.

Because the students were so new to argumentation, there were also findings that suggested what students new to the practice would most likely struggle with. One significant, and often unreported, challenge was indicated by how the students interacted. They rarely engaged each other in meaningful discussion about the investigation and rather than participate in dialogue that had the hallmarks of argumentative discourse, they more often simply took turns sharing their ideas.

Findings from this study support and extend prior research on argumentation in science and history and contribute to our understanding of how students new to the practice engage in argumentation. Reading and evaluating evidence in disciplinary-specific ways was a challenge for students (Barton & Avery, 2016; Britt & Aglinskas, 2002; Jeong, Songer, & Lee, 2007; Masnick & Klahr, 2003), and they often provided a list of details from sources rather than present a cohesive argument (Greene, 1994; Young & Leinhardt, 1998). As with students in other studies, these students also had significant difficulty with the reasoning portion of creating an argument (De La Paz & Felton, 2010; Ryu & Sandoval, 2012). While previous work suggests that students who engage in argumentation across science and history can develop ideas about argumentation that are specific to each discipline (Herrenkohl & Cornelius, 2013), the students in this study had not received prior instruction in argumentation and did not appear to have significantly different ideas about the practice across the subject areas.

Despite these evident challenges, the findings indicate that even without instruction in evaluating evidence and developing arguments, students at this age are capable of engaging in the sophisticated practice of argumentation. Students were not only able to make sense of challenging texts that incorporated novel content, but also drew conclusions from those texts that were relevant to their investigations and often served to support their claims. The students also successfully used graphic organizers to represent arguments that included multiple pieces of evidence and language indicating the students understood the importance of justifying arguments. Furthermore, the students' discussions included many indications that they were able to engage in argumentation even with little preparation and support.

In terms of methods, another contribution is that studying students' thinking at this age level poses many challenges which need to be acknowledged more consistently in the literature. Additionally, the findings suggest that combining analyses of written work with recordings of small group work provides much greater insight into student thinking than either on their own. Finally, it provides guidance for teachers on how to support students new to argumentation in both science and history, and suggests ways that younger students can develop skills with this practice. These findings support the conclusion that students of this age are not only ready to engage in argumentation in science and history, but that they also come with nascent abilities and understandings of the practice.

The following chapters present the foundation of the study, its findings, and the implications of the work. In Chapter 2, I present the theoretical framework that guides the study and our current understanding of argumentation across the disciplines of science and history. Chapter 3 provides the methods used to collect and analyze data for the study, including a description of the investigations students completed. Chapters 4, 5, 6, and 7 present the findings of my study. I proceed through the investigation in chronological order, beginning with the pre-survey students completed, continuing to the two-part investigation, and finishing with the results of the post-survey. I conclude in Chapter 8 with a discussion of the connections between this study's findings and the existing literature and a review of the implications of the work in terms of theory, methodology, and instruction.

CHAPTER 2

Conceptual Framework

This chapter presents the theoretical and empirical foundations guiding this study. I begin by reviewing disciplinary practices in education and propose argumentation as a practice worth investigating across disciplines. Next, I outline the conceptions of argumentation in the literature on science and history education and provide a definition for argumentation used in this study. Finally, I review the research on students and argumentation. This includes what is known about students' use of evidence and ability to construct arguments, the social aspects of argumentation in the classroom, and research examining argumentation across science and history.

Theoretical Framework

Involving students in the practices of the disciplines associated with school subject areas is seen by many as an essential aspect of education in those disciplines (Ford & Forman, 2006; NCSS, 2013; NRC, 2012]. There are numerous reasons for the focus on engaging students in disciplinary practices. First, a disciplinary practice-based approach can benefit student learning. In disciplinary practice-based instruction, students' learning should reflect the social practices of the discipline and the means through which each discipline achieves its aims. By making disciplinary practices central to the activity in the classroom, learning shifts from an individual mental process to something that occurs through social community-based activities (M. J. Ford & Forman, 2006). Engaging students in disciplinary practices also has the potential to make cognitive demands on students that strictly content-based approaches do not. With higher expectations of their performance and more demands on their thinking, students are more likely to learn the knowledge of the disciplines more deeply and effectively (Osborne, 2014).

Second, getting students involved in the practices helps them develop a deeper understanding of the discipline, and not just acquire a list of facts associated with the subject area. Taking a practice-based approach teaches students a more authentic image of the disciplines than other approaches (Osborne, 2014). In addition, a central goal of disciplinary activity is knowledge construction, which is a fundamentally social activity. To be educated in a discipline, students not only need to learn the knowledge constructed in that discipline, but also the ways in which that knowledge is constructed and the means that practitioners decided what does or does not count as knowledge. By getting students to “do the discipline” (Seixas, 1999, p. 332), educators go beyond simply delivering knowledge to students. Instead, students learn a subject by learning “how to know” (p. 332) in the subject and they not only learn the factual claims of a subject but also learn how those claims are generated and warranted.

Argumentation and the Standards

The importance of learning the disciplinary practices is reflected in recent standards documents. *A Framework for K-12 Science Education* (NRC Framework) (National Research Council (NRC), 2012), the framework from which the NGSS were built, includes the “scientific and engineering practices” as one of its three core dimensions of science and engineering education in grades K-12. According to the Framework, engaging in the scientific practices “helps students understand how scientific knowledge develops” (NRC, 2012, p. 42). Learning through scientific and engineering practices also helps students develop a better understanding of the content knowledge and concepts that are also integral to understanding science and engineering. In addition, the NRC Framework also argues that practice-based education improves student interest and curiosity in the content, thus promoting a deeper understanding of the subject matter.

Similarly, the C3 Framework includes “disciplinary concepts and tools” (NCSS, 2013, p. 29) as a key dimension in its inquiry arc for social studies education. According to the C3 Framework, learning and applying the concepts of the disciplines that make up social studies (civics, geography, history, and economics) helps students go beyond their everyday experience and “hunches” (p. 17) as they study social studies. The disciplinary concepts and tools allow students investigate questions more thoroughly, develop disciplinary “habits of mind” (p. 17), and generate sophisticated answers to inquiry questions in those fields. The C3 Framework also emphasizes that learning disciplinary concepts and tools is a means for students to develop an understanding of how practitioners organize and verify knowledge. This is another example of the importance of learning *how* practitioners in the discipline know what they know and not just learning *what* they know.

In mathematics, the Common Core State Standards (CCSS) (NGA Center, 2010) distinguish between “standards for mathematical practice” and “standards for mathematical content.” The standards for practice outline activities and habits of thought necessary for mathematical practice that not only demonstrates mathematical understanding, but also mirrors the work of mathematicians. This includes practices like “look[ing] for and express[ing] regularity in repeated reasoning” (p. 8) which reflect the ways that mathematicians develop generalizable rules through analyzing patterns. They also include how mathematicians find applications for their work with the practice of “model[ing] with mathematics” (p. 7) in which students connect abstract mathematical concepts to concrete problems.

Argumentation as a Common Disciplinary Practice

Despite the emphasis in the scholarship and standards documents on discipline-specific practices, there are practices that are common to multiple disciplines. Argumentation has received significant attention in research and is cited as a disciplinary practice in the CCSS as well as the NRC and C3 Frameworks. The CCSS names “Construct viable arguments and critique the reasoning of others” as one of the standards for mathematical practice (NGA Center, 2010, p. 6) and “Write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence” as an anchor standard in ELA. The NRC Framework includes “engaging in argument from evidence” (NRC, 2012, p. 3) as one of the scientific and engineering practices and the C3 Framework includes “causation and argumentation” (NCSS, 2013, p. 13) as one of the disciplinary tools and concepts in history.

Based on the characterization of argumentation in these standards documents, it is evident that making arguments in the four main school content areas share features with one another. The importance of evidence, for example, is cited in the standards for ELA, science, social studies, and mathematics. In addition to requiring arguments have “sufficient evidence,” the CCSS also want students to be able to “Draw evidence from literary or informational texts to support analysis, reflection, and research” (NGA Center 2010, p. 18). In science, the NRC Framework states that argument is rooted in evidence and that students must “formulate evidence” based on data and “examine their own understanding in light of the evidence” (NRC 2012, p. 52). Similarly, the C3 Framework says that students must be able to use evidence to make claims about the past and includes “developing claims and using evidence” (NCSS, 2013, p. 12) as one of the dimensions of the inquiry arc that describes ideal instruction in social studies in general. Although evidence appears to play a smaller role in the CCSS mathematics standards, it is nonetheless present in the expectation that students should “reason inductively about data, making plausible arguments that take into account the context from which the data arose” (NGA Center, 2010, p. 7).

Comparing Science and History

Although argumentation plays an important role in the standards documents discussed above and is an essential part of practice in multiple disciplines, research examining how students argue across disciplinary lines is scant. One of the few examples of this kind of work is a recent review of the theoretical underpinnings of argumentation in the NGSS and CCSS ELA standards (O. Lee, 2017) – that is, across science and language arts. In identifying the key differences in how argumentation is conceptualized in these two content areas, Lee demonstrated the importance of understanding argumentation across disciplinary lines. While there were certainly similarities between the two, the review reinforced the idea that there are disciplinary differences between argumentation in practice and in how it is taught in science and ELA, and likely in other combinations of disciplines as well.

While Lee’s comparison of the NGSS and CCSS is certainly useful to educators responsible for teaching using these standards documents, it does not tell us anything about how students practice argumentation across these disciplines. In fact, she notes that there were no available studies comparing argumentation across science and ELA. Similarly, searches for other combinations of school subjects yield almost no results. If research is to further our understanding of this important practice and support educators to teach it to their students, more research in this area is necessary.

To begin to address this gap, this study compares how students engage in argumentation in science and history. Comparing students’ argumentation between any content areas is important for both students and teachers. Students often have a disjointed experience of the school subjects throughout the day and have few opportunities to synthesize what they learn across content areas (Stevens et al., 2005). It also especially important for elementary teachers who are typically responsible for teaching all four major content areas. I have selected science and history in particular, however, for two main reasons. First, because of high stakes testing in mathematics and ELA, science and social studies (which encompasses history) are receiving increasingly less instructional time in the school day (Banilower et al., 2018; Blank, 2013; Center on Education Policy, 2008; T. L. Heafner & Fitchett, 2012). As a result, it is important to maximize the quality of instructional time in science and social studies (including history). By learning more about students' practice in argumentation across these two content areas, teachers may be able to leverage their instruction in argumentation in one content area to support argumentation in the other.

Second, the nature of knowledge and role of argumentation in science and history are more similar to one another than to other content areas. Namely, both disciplines rely on extant evidence to draw conclusions and develop explanations. As a result, knowledge in both disciplines can be tentative and can change based on new evidence. Teaching argumentation across science and history, then, has the potential to make this connection between the disciplines visible to students.

Argumentation

Although there are features of argumentation that appear across contexts, there are certainly aspects of argumentation that are specific to different disciplines. Before comparing students' argumentation in science and history, then, one needs to understand how the fields of science education and history education conceptualize argument. In addition to describing these conceptual frameworks, the following section also defines argumentation as it will be used in this study.

Argumentation in science education

Theories of scientific argumentation have been developed through argumentation theory, science studies, the philosophy of science, the history, anthropology, and sociology of science, and the learning sciences (Bricker & Bell, 2008). However, the extent to which that broad range of conceptualizations of argumentation has informed the research on scientific argumentation in education is limited. Arguably, the most widely adopted conception of scientific argumentation comes from argumentation theory and the work of Toulmin (Bricker & Bell, 2008; V. Sampson & Clark, 2008). Toulmin's structural model of argumentation, or the Toulmin argumentation pattern (TAP), identifies five main components of argument: *claim*, the statement being argued for; *data*, facts in support of the claim; *warrants*, the reasoning which ties data to the claim; and *backing*, the assumptions and foundations behind the warrants; and *rebuttals*, challenges to any of the other elements in an argument (Toulmin, 1958). In science education studies, TAP has been used to support students to construct scientific arguments and explanations (Driver, Newton, & Osborne, 2000; McNeill, Lizotte, Krajcik, & Marx, 2006), analyze discourse in the classroom (Erduran, Simon, & Osborne, 2004; Garcia-Mila, Gilabert, Erduran, & Felton, 2013), and evaluate students' understanding of scientific practice (Ryu & Sandoval, 2012; Sandoval, 2003).

In addition to their structure, arguments in science education are often defined by their purpose (Jiménez-Aleixandre & Erduran, 2007). Kuhn (1992) proposed that there are two types of argument: *rhetorical* and *dialogic*. In a rhetorical argument, the goal is to demonstrate the truth or falsehood of a claim through a reasoned course of thinking. In a dialogic argument, which occurs between two parties, each party presents their own assertion and then attempts to rebut the opposing party's assertion through counterarguments. Driver, Newton and Osborne (2000) made a similar distinction between these two types of arguments, but they assigned consensus building to be the main goal of dialogical arguments in which participants try to reach an agreement about what is the best claim or arrive at a shared understanding by deliberating over multiple claims and contrasting alternative arguments. Later, Garcia-Mila and colleagues (Garcia-Mila et al., 2013) proposed that dialogical arguments could have a consensus building purpose in addition to a persuasive purpose. In consensus building arguments, interlocutors present and respond to one another's arguments, but the purpose of rebuttals is not to convince the other party, but to compare the strength of claims in the hope of reaching a mutually acceptable conclusion. Much of the science education research has focused on argumentation as a consensus building activity to be engaged in as a means of "sensemaking" (Berland & Reiser, 2009). In this sense of the word, students engage in a consensus building argument to make sense of and understand a scientific phenomenon, rather than to persuade each other that their interpretation is correct.

Another key issue regarding scientific argumentation in the science education literature is the relationship between explanation and argumentation. There is some debate as to whether or not evidence plays a role in explanations and what, exactly, the difference between an argument and an explanation is (Berland & McNeill, 2012; Osborne & Patterson, 2011). I will not be addressing this question here. For one, part of the debate involves how explicitly the nature of scientific explanation and the features that make it different from everyday explanation should be explained to students (Osborne & Patterson, 2011). This study is focused on how students use evidence, not on their conceptions of scientific versus everyday explanation. Second, in the cases where researchers include evidence as a part of explanation (McNeill & Krajcik, 2008; McNeill et al., 2006) evidence plays the same role as it does in an argument; it is data that is warranted with theories or logic in support of a claim. I am interested in how students link data to their claims, and whether or not that activity can be called both argumentation and explanation is immaterial to my investigation.

Argumentation in history education

Unlike the research in science education, which frequently explicitly foregrounds argumentation, history education research often addresses argumentation in terms of students' historical thinking or reasoning (Fillpot, 2012; Seixas, 2015; van Drie & van Boxtel, 2008; Wineburg, 1999) and historical writing (Monte-Sano & De La Paz, 2012; Wiley & Voss, 1999; Young & Leinhardt, 1998). Several models of historical thinking have been proposed, but in essence historical thinking entails the concepts and ways of thinking that historians employ as they interpret and make sense of the past (Seixas, 2015; van Drie & van Boxtel, 2008; Wineburg, 1999). In these models the role of argumentation is often unstated or implicit. For example, Seixas' (2015) model of historical thinking only presents concepts students need to understand in order to think "historically" and produce historical arguments and does not explicitly tie any of these ways of thinking to argumentation. van Drie and van Boxtel (2008) do include argumentation as part of their framework for analyzing students' historical reasoning, but it is tied closely with history-specific means of warranting and critiquing arguments, such as evaluating the trustworthiness of documents. The empirical research in history education follows a similar pattern. Research on students' writing in history, for instance, often focuses on how historical thinking is evident in that writing and not whether or not students' written arguments follow any particular argument structure (De La Paz, 2005; Monte-Sano, 2010). Even though argumentation is not always explicitly named in these models, using evidence to construct claims remains the driving purpose behind these models and how they characterize historians' work.

The relative emphasis on historical thinking and reasoning in contrast with the elements of argumentation reflects the primacy of evaluating evidence in historical practice and the ways in which historians develop claims from that evidence. Unlike most scientific disciplines, historians examine events which cannot be replicated and for which further evidence cannot be produced (Hexter, 1971). Furthermore, the evidence historians use is a product of the past. Consequently, they must construct the context of a source by considering, among other things, who wrote it, when it was written, and why it was written (Hexter, 1971; Wineburg, 1991). In this analysis, the historian seeks to evaluate what the evidence means and not just if it is "false" or "true" (Collingwood, 1946). The historian must also work to determine which facts are significant and which are not, a consideration which is central to the analysis of sources (Carr, 1961). By examining multiple sources and pieces of evidence in this way, historians develop theories from the available evidence, compare sources with one another, test those theories against evidence, and then construct historical accounts and explanations (Mink, 1987; Wineburg, 1991). As opposed to developing a claim and finding facts to support it, this historical reasoning, rooted in the analysis of evidence, is the process though which historians develop their explanations and accounts. They then communicate their findings through arguments that present such accounts and the evidence that supports them.

Although the history education research does not often focus on the structure of arguments, when it is discussed, TAP, or something like it, seems to be the predominant model (De La Paz & Felton, 2010; Monte-Sano, 2010; Pontecorvo & Girardet, 1993). Pontecorvo and Girardet (1993) used TAP to break down students' discourse into argumentative operations in order to track the chains of reasoning that developed during discussions about a historical document. De La Paz (2005) worked with 8th-grade students to develop their historical reasoning and persuasive writing skills. She does not cite Toulmin or his argumentation pattern, but the scaffold provided to students closely resembled TAP, such as when it instructs students to “write facts that support [their] claim and explain how they do” (De La Paz, 2005, p. 146). A claim is being supported by data, or “facts,” and then warranted as evidence when students have to explain how they support the claim. Another study focused on students’ writing uses the TAP structure to analyze growth in students’ writing and the teaching that coincided with this growth (Monte-Sano, 2008).

Another possible reason that the history education literature emphasizes historical thinking as opposed to the structure of arguments is that TAP does not capture the disciplinary-specific traits of historical argumentation (Monte-Sano, 2010). Although historical writing may share the general elements of TAP with other disciplines, the kinds of data used in historical arguments and the means in which this data is presented, warranted, and interpreted involve a particular way of thinking unique to history. In science education, similar critiques have been made that propose that TAP does not adequately reflect the norms of argumentation and means of reasoning specific to science (Bricker & Bell, 2008; V. Sampson & Clark, 2008). I discuss the limitations of TAP as it applies to the disciplines in more depth below.

Argumentation in this study

Although the fields of science education and history education take different approaches to studying argumentation, there remain areas of overlap in the conception of argument in the two fields. In order to make comparisons of students’ argumentation in each field, argumentation needs to be defined broadly enough to cover the discipline-specific contexts in which it occurs while still allowing for the discipline-specific traits that are the hallmarks of argumentation in each field. Below, I provide a definition of argument that will be used in this study.

First, I distinguish argument from its everyday usage, which often means something very different from the kind of argumentation studied in science and history education and described in the NRC and C3 Frameworks. Hitchcock (2007) articulates that difference by proposing that arguments can be "disputational" or can be "reason-giving." In the disputational sense, an argument entails two parties who disagree about something and often have a heated and emotional exchange in which they try to convince each other to accept their point of view. In the reason-giving sense, an argument is when an individual expresses an idea or answer to question and provides one or more reasons supporting that conclusion. The argumentation that occurs in science and history, and by extension in science education and history education, is primarily of the reason-giving variety. Although there are disputes in science and history, these disagreements are settled by arguments that are rooted in reason-giving rather than other means, such as appeals to emotion, that could be a part of disputational arguments.

Using reason-giving as the primary goal of argument, Hitchcock (Hitchcock, 2007) provides a definition of argument as a "claim-reason complex" consisting of three acts on the part of the arguer: first, making a conclusion; second, premising that conclusion through propositions which support that conclusion; and third, inferences, either explicitly or implicitly stated, that show the premises lead to a conclusion. Working with this general definition of argumentation, I use TAP as a means for parsing an argument into its constituent parts. Toulmin's claim is equivalent to Hitchcock's conclusion and these claims can be premised using data. Linking premises to the conclusion, the final act in Hitchcock's definition, is roughly the same as providing warrants and backing under TAP. As I will discuss below, using Hitchcock's general definition of argument and Toulmin's framework for the structure of arguments provides a clear definition of argument while being flexible enough to accommodate the different contexts, purposes, and means of argumentation.

Although I favor the reason-giving meaning of argumentation, it is important to acknowledge the situations in which argumentation takes place and to understand that argumentation is a fundamentally social process (Bricker & Bell, 2008; Toulmin, Rieke, & Janik, 1979). Even when there is no interlocutor, arguments consist of some form of dialogue (Hitchcock, 2007). In other words, strictly rhetorical arguments, whose goal is reason-giving and can be presented as monologues, contain dialogical arguments, which presume a second party to which the argument is presented (Kuhn, 1992). Hitchcock's definition of argumentation is not tied to social context, and thus, there is room for arguments to take place in many types of dialogs and for many different purposes. For instance, claim-reason complexes can be put forward to convince or refute critics (van Eemeren & Grootendorst, 2004), persuade, negotiate, deliberate, seek information (Walton & Krabbe, 1995), or to reach a consensus (Garcia-Mila et al., 2013).

Another reason that Hitchcock's definition of argumentation fits well in this context is that it underscores the idea that argumentation is inherently tied to reasoning and thinking and also allows for a variety of modes of reasoning in arguments. By defining argument in terms of reason giving, argumentation becomes a rational act, and by extension, an act of thinking. This connection between argument and thinking has been made by researchers who propose that examining an individual's argumentation is a means of understanding how they think (Kuhn, 1992; Voss & Means, 1991). Toulmin (1979) also linked reasoning to argumentation and proposed that "reasoning is thus not a way of *arriving at ideas* but rather a way of *testing ideas critically*" (p. 9, emphasis his).

Regarding the modes of reasoning present in arguments, argumentation has historical roots in formal logic, where the validity of arguments is determined by how well they employ and follow systems of logic such as syllogisms (Bricker & Bell, 2008; Groarke, 2017). This approach, however, does not capture many of the features of everyday argumentation or the argumentation that is practiced in various fields of inquiry and thought (Bricker & Bell, 2008; Groarke, 2017; Toulmin, 1958). A practical approach to argumentation reveals that arguments are not typically translated into an artificial language, such as the ones provided by formal logic, and that argumentative reasoning can go beyond formal deductive and inductive reasoning (Groarke, 2017). Although formal and informal logic present very different means of reasoning, there is no reason why either should be excluded from playing a role in argument. Hitchcock's idea of arguments as complexes of claims and reasons is open to both types of reasoning because it does not prescribe the means by which the premise of an argument is established and linked to the conclusion.

In addition to allowing reasoning in the form of formal and informal logic, Hitchcock's definition also provides room for the disciplinary ways of thinking that are tied to argumentation in science and history. Argumentation and reasoning are context dependent, and take different shapes depending on when and why they take place. Toulmin (Toulmin, 1958; Toulmin et al., 1979) makes this point and proposes that argument has field dependent features and context determines what counts as appropriate reasoning. While his definition of "field" is unclear, he does attempt to describe some of the features of "practical reasoning" (i.e., argumentation) in fields that included law, science, fine arts, management, and ethics. Based on these categories of fields that Toulmin provides, I consider the concept of field to be roughly equivalent to discipline. Consequently, science and history will have field-dependent features of reasoning and argument unique to their disciplines. Despite these field-dependent features of argumentation, however, TAP and Hitchcock's definition of argument can still be used to describe argumentation in a variety of fields because neither framework dictates the mode of reasoning that should be used, only that it needs to be present.

The ability to accommodate argumentation in several fields is a double-edged sword, however, and leaves definition-based frameworks of argumentation such as Hitchcock's and Toulmin's open to the criticism that they are not sufficiently sensitive to the field-dependent features of argumentation and thus inadequate for an in-depth analysis of argumentation in any given field. The use of TAP as framework in education has been criticized for not being sensitive to disciplinary practices of argumentation in science (V. Sampson & Clark, 2008) and history (Monte-Sano & De La Paz, 2012) education. While TAP is useful for identifying the elements present in students' arguments or their relative complexity, it provides little guidance for evaluating how well the argument is warranted, or if the data provided sufficiently support the claim according to the norms of the discipline. This is why I do not use TAP to assess the quality of students' arguments or determine how well they align with disciplinary norms. Instead, I only use TAP to identify what elements are present in students' arguments, then analyze how students link those elements together to identify patterns of thinking. I then use the literature on argumentation practices in the disciplines to determine the extent to which students' argumentation aligns with disciplinary practices, such as contextualizing evidence in history (Wineburg, 1991) or relating an argument to a theory or model and tying it to trends in data (Kelly & Takao, 2002).

Because I am using TAP as the model of argumentation for this study, I also need to clarify how "data" and "evidence" are used here. While Toulmin labels information or facts provided in support of a claim as "data," much of the literature in science and history education uses the term interchangeably with "evidence." In order to be consistent with the language used in the literature, I also use the terms as equivalent to one another. Unlike some of the science education research, however, I do not be use "data" to mean only numerical data or recordings of measurements. The term data applies as equally to a table of temperatures as it does to a quotation from a historical document. In the cases where I refer to data as a set of recorded measurements I explicitly call this "numerical data."

Students and Argumentation

Many aspects of formal argumentation in the disciplines are difficult for students. The process of making sense of texts, especially multiple texts, is complex and because of the specific demands of argumentation and disciplinary practice may pose challenges to students. In particular they can have difficulty interpreting and using evidence in appropriate ways, and providing the reasoning tying their arguments together. Students, however, are capable of using evidence to create arguments, and there are supports that can help students do this. We know less about their understanding and engagement in argumentation across disciplinary lines, however.

Making Sense of Texts

Before they can begin to develop claims and supporting arguments, students must first read and make sense of texts (i.e., sources). Successfully reading sources for the purpose of developing a claim and constructing an argument requires students to construct models of text across sources within the context of discipline-specific text features, purposes for reading, and comprehension strategies. In the construction-integration model (Kintsch, 1998) when encountering a single text, comprehension begins with the process of building a *textbase* which is the network of propositions the reader builds from the words and phrases that captures the explicit meaning of the text. Next, the reader relates their background knowledge, both passively and through active inferences, to the textbase to create a *situation model* representing the reader's interpretation of the text.

While the construction-integration model addresses the basics of reading comprehension, it does not specifically address the disciplinary differences between texts. These differences across disciplines affect not just the way that a text is constructed and but also how readers should interpret the text. This is demonstrated by comparing, for one, the epistemological orientations of literary, scientific, and historical texts (Goldman et al., 2016). In literary reading, authorial intent plays an important role, which contrasts with the iterative and tentative nature of knowledge central to science, or the acknowledgement in historical texts that our understanding of the past is incomplete.

To further complicate matters, in the context of argumentation students must read and analyze multiple texts in order to develop a claim. This requires students to create new levels of representation that account for multiple texts and the relationships between them (Goldman et al., 2016). One, the *intertext model*, represents the information related to the source of the documents and details within, connections relating what information came from which source, and the level of agreement or disagreement between sources (Perfetti, Rouet, & Britt, 1999). The intertext model shares many features with the sourcing heuristic that Wineburg (1991) described historians using. The *integrated model* is the second representation developed in reading multiple texts and consists of the reader's holistic understanding of the situation and information presented across texts and not just within a single text. To build this model, the reader may make inferences to fill in the gaps of one text with details from another, build connections between information presented in different texts, or evaluate conflicting information. In this way, it is also possible that reading multiple texts can strengthen a reader's situation model of individual texts (Wiley & Voss, 1999)

Finally, the reader builds these representations based on their understanding of the task, or *task model* (Goldman et al., 2016). When encountering a text, readers have goals for reading in mind and ideas about strategies for achieving those goals. Thus, the task model plays a critical role in how a reader develops the other representations of text (Rouet & Britt, 2011). Additionally, the task model is influenced by the discipline in which one is reading. The reading goals in scientific, historical, and literary contexts are often quite different, and as a result require specialized reading strategies, such as attending to prose and structure in literary settings (Lee, Goldman, Levine, & Magliano, 2016).

Reading and Evaluating Evidence. Given the complexity of reading multiple sources in discipline-specific contexts, it is not surprising that students often have difficulty reading sources for the purposes of argumentation. Because of the important role that interpreting evidence plays in historical argumentation and practice, research on argumentation in the history education literature often examines how students interact with and interpret evidence. Given the complexity of making sense of texts in the context of explanation and argumentation described above, it is unsurprising that students often encounter difficulties in this area. In general, this research has found that, without preparation, students often have difficulty engaging in the critical reading practices that would enable to them to evaluate and interpret sources for the purpose of developing an argument in support of a historical interpretation (Barton & Avery, 2016). In contrast with the experts in his study, Wineburg (1991) found that high school students did not use the heuristics of sourcing, contextualization, or corroboration when evaluating the trustworthiness of documents being used to answer a historical question. Instead, the students ignored the attribution of the sources and did not qualify their analysis of the documents by referencing other documents. Wineburg attributed the difference in the high school students' treatment of evidence to a lack of understanding of the ways of knowing in history and how claims are warranted in the discipline. Similarly, Britt and Aglinskas (2002) found that students did not spontaneously use the sourcing heuristic. They did not record source information in their notes, were unable to answer many questions about source-related details of the documents they had access to, and included details from a novel (i.e., an inappropriate source) in support of their argument.

Students also have difficulty evaluating evidence in the sense that they often do not distinguish between evidence in support of an argument and details about a topic. Chambliss and Murphy (2002) found that, when reading, younger students tend to interpret arguments and evidence as a topic with details, rather than as a claim with supporting data and reasoning. Students were presented with a written argument that included a number of pieces of evidence in support of a main claim. When asked to represent that argument, students often represented that argument as a topical net of related details and their representations did not match the argument structure presented in the text. Similarly, high school and college students also have difficulty using details in a way that supports an argument and goes beyond elaborating on a topic (Greene, 1994; Young & Leinhardt, 1998). Young and Leinhardt (1998), for instance, found that high school students in an Advanced Placement U.S. history class often used details they pulled from evidence at face value, and did not transform those facts into evidence supporting a claim until they had more experiences with writing arguments. Although there were times when students integrated their interpretation of the documents into their argument as evidence, they were much more likely to present a list of details related to the topic of their essay.

In science, students have similar difficulty in evaluating the quality of evidence and its relationship to their argument. When collecting or given data, students can fail to attend to issues of reliability and error in the data they encounter (Jeong et al., 2007; Kanari & Millar, 2004; Masnick & Klahr, 2003). There do, however, seem to be differences in how students evaluate first- and second-hand data during investigations and when constructing arguments. When using first-hand data, students may be more likely to use more data in their argument and discuss the source of the data possible errors in the data when compared to using second-hand data (Delen & Krajcik, 2015; Hug & McNeill, 2008). On the other hand, using second-hand data that is presented as a primary source (e.g., a scientists' journal) may positively influence how students collect or record their own data during an investigation (Palincsar & Magnusson, 2001).

Constructing Arguments

Even though it can be difficult for students to use evidence and construct arguments, it is a skill that can be learned through guidance and practice. In science, proper scaffolding can help students evaluate and use evidence more effectively, improve the quality of their written arguments, and articulate the reasoning they use to connect evidence to claims (McNeill, 2011; McNeill et al., 2006; Osborne, Erduran, & Simon, 2004; Ryu & Sandoval, 2012). McNeill and colleagues (2006), for instance, provided students with scaffolds carefully targeted on getting students to articulate their arguments using claims, evidence, and reasoning. After fading out the scaffolds, they found that students wrote more robust and complete arguments after the intervention. Providing explicit instruction on the structure of arguments is not, however, the only way that students can learn to improve their argumentation skills. Ryu and Sandoval (2012) studied the elementary students in a classroom where students were given opportunities to engage in discussions about scientific phenomena and encouraged by the teacher to answer questions like “How do you know?” and “How do you convince others?” during the discussions. Although these supports did not address argument by name, through sustained practice students’ understanding about the need for evidence and explicit justification improved.

Research in history education has also shown that in the right instructional context students can learn to evaluate evidence and use that evidence in arguments in ways that reflect the disciplinary practices of history. This includes learning to read sources while employing the heuristics used by historians (De La Paz, 2005; De La Paz & Felton, 2010; Fillpot, 2012; Reisman, 2012), support arguments with evidence (Monte-Sano, 2008), and evaluate the evidence they use to support their written arguments (Monte-Sano, 2010, 2011). Learning to apply disciplinary strategies as they read not only improves students’ understanding of texts but can also improve the quality of their arguments. De La Paz and Felton (2010) tested an intervention in which they taught students a historical reasoning strategy for working with historical documents that supported students to evaluate the sources of the documents and to look across multiple documents. They found that students made better claims, included more rebuttals, and made more appropriate use of evidence in their arguments than students who did not receive the same instruction.

In addition to strategy instruction, manipulating the evidence given to students can result in improved argumentation and use of evidence. Providing multiple sources of evidence can encourage students to use it in more sophisticated ways (Paxton, 2002; Rouet, Britt, Mason, & Perfetti, 1996; Wiley & Voss, 1999). Making the author of texts “visible” is another way to improve students’ historical writing. Paxton (2002) provided students with textbook passages written so that the author was visible, such as by using the first person or hedging language like “perhaps”, as opposed to anonymous. Compared to students with texts by anonymous authors, those given texts with a visible author encouraged students to engage with the text in more meaningful ways. Those students wrote essays that were longer, referenced the source of documents more frequently, and included more personal perspectives on the topic.

Working in Groups

While the cognitive aspects of argumentation, such as the ability to evaluate evidence appropriately or link evidence to a claim, are important, the social aspects should not be ignored. Small group work is a common participation structure in elementary classrooms, and research on students' participation in this kind of work is relevant to understanding how they engage in argumentation.

In the same way that students typically need support to evaluate evidence and develop well defined arguments, they often need support to engage with one another in productive ways. Students trained in cooperative learnings skills or discussion protocols, for instance, can have more productive academic discussions than students without such training (Ashman & Gillies, 1997; Gillies & Ashman, 1998). Similarly, when teachers used and modeled questioning strategies with students, the students provided more elaborative and detailed responses when working in small groups (Gillies & Khan, 2009).

The different roles that students play during small group work can also shape the academic quality of discussions. Using research on business management, Maloney (2007) was able to identify various roles that students play and their effects on group work. One group of students, for instance, consisted largely of students who took on roles that contributed to or managed the discussion. As a result, their discussions were the most successful, particularly along criteria related to argumentation. Similarly, Hogan (1999) categorized students' behavior during small group work into "sociocognitive roles," some of which promoted co-construction of ideas and some of which detracted from it.

The goal of a task is also relevant to how students interact with one another and engage in argumentation. Students' perceived purpose of a discussion, for instance, can affect the ways in which they engage in a task. When students see a discussion as "something to get through" rather than a meaningful exchange of ideas, they may not engage in high-quality discussions (Prado-Olmos, 1994). Instead of responding to one another's ideas, they can simply take turns expressing their ideas in ways that resemble the teacher and student roles typical in a whole class context. The stated purpose of a discussion can also affect how students interact with one another and engage in argumentation. Garcia-Mila and colleagues (2013) found that the framing of a discussion was related to the kinds of arguments students produced.

Research Across Disciplines

There is currently very little scholarship about students' argumentation in science and history or in different contexts. One exception is the work of Herrenkohl and Cornelius (2013) that investigated the norms of discourse around argumentation in fifth and sixth graders' science and history classes. They found that after engaging in argumentation in both subject areas, students developed some similar norms around argumentation in science and history. There were some key differences, however, and students more often included analogous and imaginative thinking in history and were more likely to define arguments as theory driven in science. This indicates that students may have some ideas about the generalizable features of argument while still understanding that there some aspects of argumentation that are specific to the disciplines. It is important to note, however, that this study took place in a classroom where argumentation was heavily supported with materials and instructional scaffolding. The norms that developed occurred after student sustained practice in argumentation in both science and history. It is unclear, then, what students' initial ideas about argumentation in each field were and whether they began with a generic conception of argumentation or one that saw argument as being different according to the discipline.

Another study that examines students' argumentation in different contexts comes from Iordanou (2010). In this study, students took part in discussions about a social topic and a science topic via instant messaging software. Iordanou found that students who were prepared in advance of the discussion with a list of possibly relevant facts included more counterarguments and requests for clarification than those who did not receive the preparation. Furthermore, students who were only given facts related to the science topic demonstrated more sophisticated argumentation on both topics than students who were only prepared with facts about the social topic. This indicates that transfer of argumentation skills across domains is possible, and that there may be a relationship between the content of instruction in argumentation and the kind of transfer that occurs.

Together, these studies suggest that when students engage in argumentation across domains, the interaction between their ideas about argumentation across those domains is complex. As students develop skills in argumentation, their learning in one subject area may have an effect on their skills in another subject. And, as students gain practice in argumentation, they may develop discipline-specific ideas about argument even if those ideas are tacit and not explicitly stated or taught. While this research is promising and indicates that studying students' understanding of argumentation across contexts is worthwhile, it is not sufficient for making many claims about how students understand argument and evidence in different content areas. A number of questions about students' use of evidence and approaches to argumentation across disciplines remain. It is clear, for instance, that students can struggle to evaluate evidence effectively in both science and history, but it is unclear whether those struggles are related to their understanding of evidence in general or if they are tied to students' skills or beliefs about the discipline. The study proposed here hopes to contribute to our limited understanding of this topic by investigating how students evaluate and use evidence in both science and history.

School Subjects vs. Disciplines

Another likely challenge when students attempt to engage in disciplinary practices such as argumentation is the disparity between the school subjects and disciplines as they are practiced in the field. What students do in "school" history and "school" science is very different than what scientists and historians do as they engage in their work. Many of these differences are tied to the difference in epistemologies and assumptions on which science and history in the two settings are based.

One notable difference between the disciplines in the field and in school is the way in which knowledge is presented and the ideas about the nature of knowledge implied by those representations. Science textbooks and trade books, for instance, often represent scientific knowledge as something that is concrete and known with great certainty (Abd-El-Khalick, Waters, & Le, 2008; D. J. Ford, 2006). This is reinforced by the frequent use of *boosting* language, such as "certainly" or "definitely," in these texts, as opposed to the *hedging* language, such as "may" or "possibly," more frequently found in authentic scientific texts (Oliveira, Akerson, Colak, Pongsanon, & Genel, 2012; Parkinson & Adendorff, 2005). As a result, the nature of scientific understanding as something that is open to debate and the result of many layers of analysis is hidden, and science becomes predominantly a list of facts in books (Rice, 2002).

Similarly, history instruction often focuses on “facts” and obscures the analysis and reasoning necessary to construct knowledge in the field. History instruction can lack “authentic intellectual work” in which students are mostly asked to reproduce knowledge (Au, 2009). This is evidenced in the literacy practices in history classes that are often centered around basic comprehension and students’ ability to locate and recall information (Kiuhara, Graham, & Hawken, 2009; Nokes, 2010), and which result in students developing strategies for finding the answer to the questions at the end of the chapter rather than attending to the substance of the text (Greenleaf, Schoenbach, Cziko, & Mueller, 2001). High stakes testing, which typically does more to assess students’ ability to recount existing narratives or details rather than their ability to engage in historical reasoning (J. Lee & Weiss, 2007), often promotes this kind of instruction (Au, 2009).

In both cases, students’ experiences with the disciplines of science and history in the classroom do not reflect the practices or epistemological stances of scientists and historians. As a result, when asked to engage in activities involving the disciplinary practices promoted by standards documents, students are likely to base their understanding of what is expected of them on their prior non-disciplinary practice-based work. Instead of analyzing the perspective of a source’s author, for instance, they may simply produce a basic summary because that is what has been asked of them in the past. The difficulty of engaging students in sophisticated disciplinary practices, then, is compounded by their previous experiences with a focus on pre-fabricated narratives, basic comprehension, and recall of the "right" answer.

Conclusion

This chapter reviewed the empirical and theoretical literature on argumentation supporting this study. I first highlighted the importance of disciplinary practices in education, and suggested argumentation as an area of interdisciplinary study. I also reviewed the conceptualizations of argumentation in the science and history education literature and provided a definition of argumentation guiding this study. Finally, I reviewed the empirical literature on how students engage in argument in science and history. In the next chapter I describe how I used this research base to develop an argumentation-focused activity and to guide my study design and analysis.

CHAPTER 3

Methods

This qualitative case study examines the ways that a class of third-graders made sense of sources and used them to develop arguments in support of a claim. I engaged the students in two investigations, one in science and one in social studies, in which they were given an investigation question, read and evaluated sources, and then constructed an argument supporting their answer to the question. To characterize the students' thinking I collected and analyzed the written work they produced during the investigation along with video recordings of the students completing the tasks. All of the students also completed pre- and post-surveys which allowed me to evaluate their thinking about the disciplines, the nature of evidence, and thoughts on their own work. Interviews with a focus-group of students helped triangulate the data collected from the other parts of the study and add depth to my understanding of the students' beliefs.

Study Setting

This study took place in a third-grade classroom at Matthews Elementary,[3] a K-5 elementary school in a small city in Michigan. The school serves a linguistically and socio-economically diverse population, see Table 1. Matthews Elementary is also an International Baccalaureate (IB) school. The IB curriculum is centered around what it calls "Transdisciplinary Themes" (International Baccalaureate, 2009, p. 8) which are "themes that identify areas of shared human

[3] All names are pseudonyms.

experience and have meaning for individuals from different cultures and ethnicities" (p. 8) and include ideas such as "Who we are," "Where we are in place and time," and "How the world works." IB uses these themes to develop units of inquiry that integrate content from a variety of content areas to address these themes.

Table 1: Matthews Elementary Demographics 2015-2016.

Students	317
Economic status	
Economically Disadvantaged	62%
Non-Economically Disadvantaged	38%
Ethnicity	
White	28%
Hispanic/Latino	27%
African-American	24%
Other	20%
Language	
English Language Learners	24%

* From the state's school demographics website

There were three third-grade classrooms at Matthews Elementary, taught by Ms. Norris, Ms. Adams, and Ms. Fisher. My study was conducted in Ms. Fisher's classroom. The school's principal worked with the third-grade team to decide which classroom I would work in. In their decision, they considered the experience each teacher had in third-grade, other obligations the teachers had, and the temperament of the students in the class.

Role of the Researcher

I selected Matthews Elementary as the site for this study because my previous relationship with the school made it easier to access the site and obtain permission to conduct the study. Prior to this study, I served as a field instructor to student teachers completing their fieldwork requirements at Matthews Elementary. I was also involved with the school outside of my position as a field instructor. I attended the IB training seminars with the teachers and administrators as the school was transitioning into the IB program. In addition, I volunteered as an instructional assistant in their summer school program. As part of my activities at Matthews Elementary, I worked with two of the third-grade teachers in some capacity. I was the field instructor to a student teacher in Ms. Norris' classroom, and an instructor in Ms. Adams' teacher education program. I did not, however, have previous experience working with Ms. Fischer.

During the study, I took the role of lead teacher in the classroom. Ms. Fischer was present in the classroom during the study but took a secondary role. She allowed me to lead whole class instruction, such as introducing the graphic organizers, and manage the classroom during small group work. While students were working independently, Ms. Fischer worked with one group, Group 2, while I circulated through the classroom among the other groups.

Study Methods

This study consisted of four main elements: a pre-survey, two investigations (one in science and one in history), a post-survey, and three focus-group interviews. Data collection occurred over the course of nine days. Table 2 outlines the data collection timeline. In the sections below, I describe the in-class discussions and activities in greater detail, followed by a description of the interviews conducted with students.

Except where noted, the students completed the activities in small groups. There was a total of 4 groups consisting of four students and one group with five students. Previous studies have collected data from small groups of students to capture their thinking as they evaluate sources and construct arguments (Maloney, 2007; Pontecorvo & Girardet, 1993). My goal was to examine students' thinking in as naturalistic setting as possible. In addition, students worked in small groups to give more students the opportunity to share their thinking and ideas; in a whole class discussion I would have likely heard from far fewer students.

Table 2: Timeline and overview of data collection.

Day	Activity	Description	Data Collected
1	Beliefs Pre-Survey, day 1	Small group discussions and questionnaire about argumentation and evidence in science, history, science class, and social studies class.	Questionnaires Audio/video recordings (~45 minutes each)
2	Beliefs Pre-Survey, day 1	Small groups complete discussions and questionnaires about argumentation evidence in the disciplines.	Questionnaires Audio/video recordings (~10 min. each)
2	History investigation introduction and source analysis day 1	History content knowledge pre-test. "Who stole the cookies" introduction to investigation materials. Whole class introduction to history investigation. Small groups begin to analyze sources for history investigation.	Content pre-tests Evidence analysis worksheets Audio/video recordings (~10 min. each)
3	History investigation source analysis day 2	Small groups complete analysis of sources for history investigation.	Evidence analysis worksheets Audio/video recordings (~45 min. each)
4	History investigation argument construction	Small groups create graphic organizers representing their argument supporting their answer to the investigation question.	Graphic organizers Audio/video recordings (~35 min. each)
5	History investigation interview	Focus group interview with 5 students about the evidence and arguments in the history investigation.	Evidence rating sheets Audio/video recordings (~30 min.)
6	Science investigation introduction and source analysis day 1	Science content knowledge pre-test. Review using materials. Whole class introduction to science investigation. Small groups begin to analyze sources for science investigation.	Content pre-tests Evidence analysis worksheets Audio/video recordings (~10 min. each)

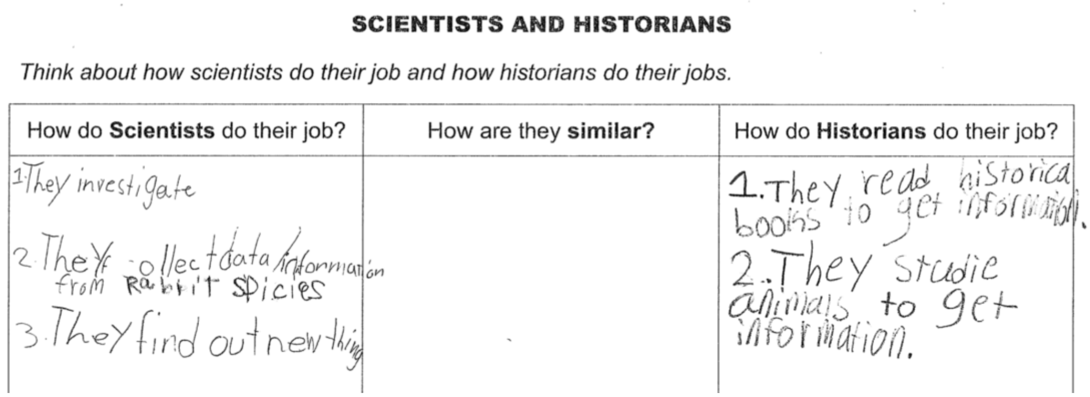
7	Science investigation source analysis day 2	Small groups complete analysis of sources for science investigation.	Evidence analysis worksheets Audio/video recordings (~45 min. each)
8	Science investigation argument construction	Small groups create graphic organizers representing their argument supporting their answer to the investigation question.	Graphic organizers Audio/video recordings (~35 min. each)
9	Beliefs Post-Test	Small group discussions and questionnaire comparing the investigations in science and history.	Questionnaires Audio/video recordings (~45 min. each)
10	Science Investigation Discussion	Whole group discussion about the solution to the investigation question.	Audio/video recording (~30 min.)
10	Science investigation interview	Focus group interview with 4 students about the evidence and arguments in the science investigation.	Evidence rating sheets Audio/video recordings (~30 min.)
11	Investigation comparison interview	Focus group interview with 4 students comparing the two investigations.	Audio/video recordings (~30 min.)

Ms. Fischer composed the groups and took two main factors into account when selecting students for each group. One consideration was the students' academic progress. Ms. Fischer tried to create groups of mixed ability, primarily using her own assessment of their reading levels, with at least one of her "stronger" students in each group. She also took into account the students' ability to work together and put students who she knew did not work well together in separate groups. The one exception to this was Group 2. Ms. Fischer grouped those students together because she anticipated that they would need extra help during the investigations. Two of the students were English language learners and the other two were considered by Ms. Fischer to need the most support reading of all the students in her class. She put these students in a group together with the intent of working with them and being able to provide them support during the activities. As a result, I was not able to use Group 2's data for this study. The group received so much teacher support that I did not feel data from the group would have been comparable to the work the other groups did on their own.

Pre-Survey

To learn about students' initial ideas about evidence and argumentation in science and history I administered a pre-survey. McNeill (2011) studied students' changing ideas of explanation, argument, and evidence over the school year as they engaged in scientific argumentation. Although I did not study the change in students' beliefs over time, McNeill does provide a good template for learning about what students believe about evidence and argumentation. In this study, I am similarly interested in how students' beliefs about evidence come into play when they engage in argumentation using evidence, and thus use McNeill's questionnaire as a guide.

The pre-survey consisted of five sets of open-ended questions in which students were asked about scientific and historical practice and evidence and its use in the domains of science, history, science class and history class. The first question asked students to compare how scientists and historians do their jobs. Figure 1 presents the question and one group's responses. Because this study examined possible differences in students' argumentation across disciplinary lines, it was important to understand students' existing beliefs about the disciplines of science and history and how work is done in each field.



SCIENTISTS AND HISTORIANS

Think about how scientists do their job and how historians do their jobs.

How do **Scientists** do their job?	How are they **similar?**	How do **Historians** do their job?
1. They investigate 2. They collect data information from Rabbit Spicies 3. They find out new thing		1. They read historical books to get information. 2. They studie animals to get information.

Figure 1: Pre-Survey: "How do scientists and historians do their jobs?"

The next set of prompts focused on the kinds of evidence used in different contexts and the purpose of using evidence. In this set of eight questions, the same two questions were repeated for the domains of science, science class, history, and social studies class[4]. First, students had to create an "idea web" of everything they thought could be used as evidence in each domain. Figure 2 presents one group's answer to this question for scientists. Second, students were asked

[4] The survey asked about "social studies class" and not "history class" because the students had not yet been introduced to history as its own subject.

why evidence was used in each domain. Figure 3 presents the question as phrased for science class along with one group's answers.

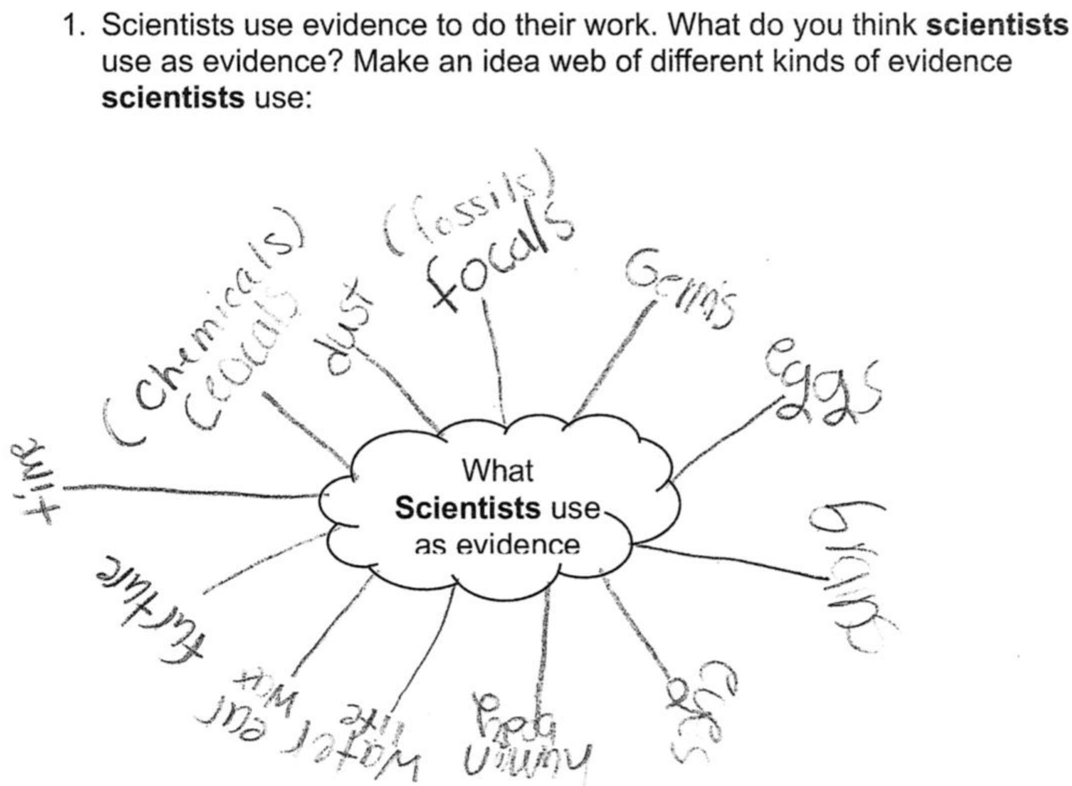


Figure 2: Pre-Survey: "What do scientists use as evidence?"

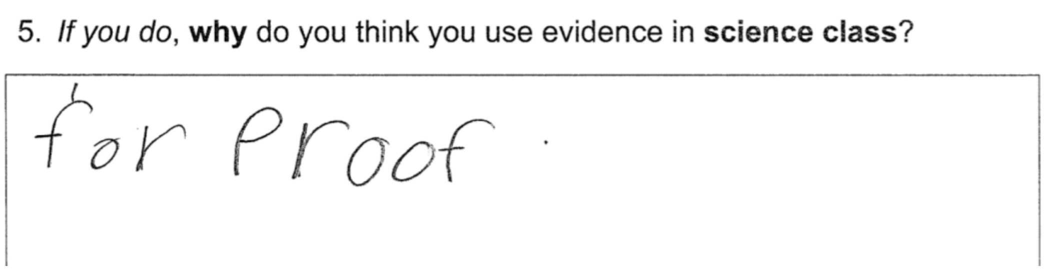


Figure 3: Pre-Survey: "Why do you use evidence in science class?".

The students worked in their small groups to answer this question across the first two days of the study. This was done prior to any introduction to the investigations. Because of time constraints they were unable to complete the last two questions on the survey and completed the survey on the second day of the study. See Table 2 for more details and Appendix 1 for the complete set of survey questions.

Investigation Overview

The majority of this study comprised of two source-based investigations, one in science and one in history. Each investigation consisted of four phases: *content pre-assessment, introduction, evaluating sources*, and *argument construction.* The *content pre-assessment* was a brief pre-test designed to assess students' prior knowledge of the investigations' content. In the *introduction* I gave the students the investigation question and provided background knowledge and context necessary to complete the investigation. During the *source evaluation* phase students read 6 sources related to the investigation question and evaluated how each source helped answer the investigation questions. Finally, students decided on a claim answering the investigation questions and developed a supporting argument using a graphic organizer during the *argument construction* phase. The following sections describe each phase in detail.

Building on existing literature, I developed two source-based investigations, one in science and one in history. These investigations were designed to present students with an investigation question, texts to analyze and base their answers on, and a graphic organizer to supports students to construct an argument in support of their investigation question.

Task content. The content used in the tasks was based on the third-grade science and social studies standards used by Matthews Elementary. This was done to ensure that the content would fit into the work that students were already doing and make the investigations a more natural part of the students' experience. In addition, I chose content that students had not yet covered in their class and both investigations introduced new topics to the students. I also had to ensure that I would be able to craft authentic investigation questions around the topics. In order to build an investigation around the content there needed to be some uncertainty in the scholarship or room for debate in the question.

The history task was developed around the following standards: “Describe the causal relationships between three events in Michigan’s past (e.g., Erie Canal, more people came, statehood),” and “Describe how Michigan attained statehood.” After consulting with the classroom teacher and surveying teaching materials on the topic of Michigan’s statehood, I chose the conflict between Michigan and Ohio over Toledo as the topic. Rather than focus on the timeline of events, which is well-established in the scholarship and not generally up for debate, I chose to focus on the causes of Michigan’s final decision to give up their claim on Toledo in exchange for statehood. Prior to becoming a state, there was a conflict between Ohio (already a state at this point) and the Michigan territory. Both claimed Toledo as their own. When applying to become a state, the federal government offered Michigan statehood on the condition that it conceded Toledo to Ohio. A convention of delegates rejected these terms on September 28, 1836. Three months later, on December 14, a second convention of delegates re-voted on the issue, unanimously accepted the terms of the deal, and gave up claim to Toledo to become a state. During the debate on this issue there was considerable disagreement between the Whigs who wanted to fight for Toledo and the Democrats who prioritized achieving statehood.

Based on this, the investigation question was, “Did the second vote for Michigan’s statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?” The students were given three claims to select from: “The second vote happened mostly because of popular opinion,” “The second vote happened mostly because of the work of the Democratic party,” and “The second vote happened both because of popular opinion and also because of the work of the Democratic party." The sources in the investigation slightly favored the claim that the second vote was the result of the work of the Democratic party because more sources mentioned their role in the event. Table 3 provides a summary of the six sources used in the investigation. For the complete set of sources and evaluation worksheets see Appendix 2.

Table 3: Overview of History Sources.

Title	Description
"Results of the First Convention"	A table presenting the attendance and voting tallies of the first convention for Michigan statehood. Demonstrates that the first vote was fairly close.
"Biography of Michigan's Governor"	An excerpt from a biography of Stevens T. Mason. Describes the variety of meetings that were held in Michigan communities after the first vote and that five men were sent to find delegates for a second convention.
"Results of the Second Convention"	A table presenting the attendance and voting tallies of the second convention for Michigan statehood. Includes the unanimous approval and an explanation that many Whigs did not attend.
"Report of the Democratic Meeting"	Newspaper article from 1836 reporting on a Democratic meeting and its recommendation that citizens call for a second convention on statehood.
"Book on Michigan History"	Excerpt from a book on the history of Michigan. Explains that Democrats put out a notice about a second election and that many Whigs did not join in.
"Letter from the Governor"	Excerpt of a letter from Stevens T. Mason to a group of Democrats. The letter suggests that the people of Michigan should select new delegates so the decision is "from the people of Michigan."

The science task was developed around the following standard, "Make observations and/or measurements of an object's motion to provide evidence that a pattern can be used to predict future motion." After surveying materials used to teach this standard in Michigan schools and in Matthews Elementary school district, I developed an investigation in which students needed to make predictions about the motion of objects based on force and mass. The investigation was based around a hypothetical push-cart race. In the race, participants built carts similar to soapbox derby cars, but that were raced on a flat track with one person pushing the cart and a second person in the cart pushing. Figure 4 presents an image given to students as an example of a push-cart.

Figure 4: Example of a Push-cart.

In the investigation, there were two teams. Karl and Edgar built the “Blue Monster” and Ebony and Asia made the “Green Bolt.” The Blue Monster was heavier but was pushed with greater force than the Green Bolt. The Green Bolt was pushed with less force but was lighter than the Blue Monster. Figure 5 provides the information provided to students about the weight and force of each push-cart.

Racing Teams

	Blue Monster (Karl & Edgar)	Green Bolt (Ebony & Asia)
Weight	100 kilograms	75 kilograms
Push Force	25 newtons	20 newtons

Figure 5: Information Given to Students about the Push-Carts.

Based on this information, the investigation question was, “Which car will win the race? The Blue Monster or the Green Bolt?” As with the history investigation, the students were given three claims to choose from: “The Blue Monster will probably win,” “The Green Bolt will probably win,” or “It will probably be a tie.” The sources in this investigation slightly favor the claim that the Green Bolt will win because it has an aerodynamic advantage. Table 4 provides a summary of the six sources used in the investigation. For the complete set of sources and evaluation worksheets see Appendix 3

Content pre-assessment. Because prior knowledge can affect the quality of arguments, it was important to assess the students’ existing knowledge of the investigations’ content. Prior studies have used content pre-assessments as a way to eliminate students’ prior knowledge as an explanation for differences in writing and argumentation (Monte-Sano & De La Paz, 2012). In this study students took a content pre-assessment immediately before each investigation. Unlike the rest of the tasks, this was completed individually. The pre-assessments were multiple choice

Table 4: Overview of Science Sources.

Title	Description
"Cart Weight Test"	Table of results from an experiment testing different weights of push-cart. Demonstrates that lower weights yield faster times.
"Engineer's Notes"	Notes from an engineer who makes push-carts as a hobby. In the results, a lighter but weaker cart went faster than a heavier but stronger cart.
"All About Air Resistance"	An article explaining air resistance and the effects of smooth shapes vs flat surfaces on aerodynamics.
"Edgar's Notes"	Notes from one of the team members in the investigation scenario. Describes the Blue Monster as having a square front and the Green Bolt as having smooth curves.
"Push Force Test"	Table of results from an experiment testing different forces on a push-cart. Demonstrates that higher push strengths yield faster times.
"Force, Weight, and Speed"	Article about force and weight. Explains the relationship between force weight and speed and that heavy things can go faster than light things if pushed with enough force and that if something is light enough it can go faster than something pushed with greater force.

tests and asked five questions about facts and concepts important to understanding each investigation question and completing the tasks. See Appendix 4 for the tests. Although I did not do a statistical analysis comparing different experimental conditions, I used descriptive statistics to characterize the level of students' prior knowledge on the investigations' topics as a way to flag students whose arguments may have been influenced by their prior knowledge, or the lack thereof.

Introducing the materials. Prior to beginning the history investigation (the first of the two investigations) I needed to introduce students to the materials and tools they would be using to complete the investigations. To explain to students how to use the various materials, I created a sample investigation asking the question, "Who stole the cookies from the cookie jar?" The investigation presented a scenario in which cookies were taken from a teacher's desk and the goal was to figure out what student had taken them. Figure 6 presents the scenario given to students.

School Mystery!

On October 1, 2017 in a school called Brookline Elementary, someone stole cookies from a teacher's classroom. Mrs. Plano the music teacher baked a batch of cookies at home and brought them to school to share. When she got to school in the morning she put the cookies in a cookie jar on her desk. By the end of the day, they were gone. The principal wants your help to figure out who is to blame for this awful crime.

Figure 6: Sample Investigation Scenario.

After setting up the investigation, I then modeled how to use each of the materials that students would be using in the history and science investigations. This included explaining each part of the source evaluation worksheets, reading the sources and making notes on the source evaluation worksheets, deciding on a claim based on my findings, explaining the elements of the SenseMaker boards (a graphic organizer for creating an argument) and constructing an argument using the SenseMaker board, as I explain below.

This introduction happened on day two of the study (see Table 2) immediately after the history content pre-assessment and prior to providing the topic of the investigation question. I also reviewed the “Who stole the cookies from the cookie jar?” investigation and materials on day six of the study, so that students would be reminded how to use the source evaluation worksheets and SenseMaker boards.

Investigation introduction. Once the students understood the elements and materials of the investigations, I introduced the investigation with a brief PowerPoint presentation that provided students with background knowledge necessary to understand the investigation question and help ensure that all groups were entering the task with a common understanding of important terms and concepts. While front-loading in this way runs the risk of discouraging students' sensemaking, there was limited time in which students could conduct the investigations. In order for students to make sense of the texts in the investigations it was necessary to explicitly introduce them to a few key concepts and events for context.

In the introduction to the history investigation, I began by defining key terms from the investigation: delegate, political party, convention, Whig, and Democrat. I then reviewed for students how a territory became a state so that students would understand the political process in which these events took place. Finally, I reviewed the timeline of Michigan applying to become a state, the disagreement between the Whigs and Democrats, and the outcome of the revote on Michigan's statehood.

For the science investigation, I introduced the concept of force and provided students with a definition of "newton" and "kilogram" along with providing examples to give students a sense of the scale of each unit (e.g., an average third-grader weighs 25 kilograms, it takes 1 newton to pick up an apple). I then explained to students what a push-cart was by showing them example photographs (e.g., Figure 4) and a video of a push-cart race.

After providing students with the necessary background information, I presented the investigation question and took questions from the students about the background information and investigation question.

Evaluating sources. Immediately after the presentation introducing the investigation question, the students began working in their small groups to evaluate the sources[5] they would

[5] In this study I differentiate between "sources" and "evidence." I consider sources to be the texts or raw data that are analyzed in order to answer a question or develop an explanation. Sources, or more specifically the data or

use to answer the investigation questions and construct an argument. Previous studies in history education have separated source evaluation into a separate task that occurs prior to argument construction (Britt & Aglinskas, 2002; De La Paz, 2005; De La Paz & Felton, 2010; Monte-Sano & De La Paz, 2012). In this study, source evaluation occurred on a separate day for two reasons. First, this provided me the opportunity to collect data on how students make sense of sources and data. Designing this as a separate task helped students to focus their efforts on interpreting the sources rather than immediately trying to make it fit into an argument. Second, having a variety of kinds of documents (e.g., both primary and secondary) and sources that supported multiple claims necessitated a relatively large set of sources.

The source evaluation process was the same for each source and across both investigations. I first read the source aloud to the class and instructed them to follow along while the source was projected on the board. I read all parts of the source and named each section of the source as I did so (e.g., "Title, 'Push Force Test'"). I then reread the prompt, "What does this piece of evidence tell you about the investigation question?" and told students to work together to record their ideas. I gave students approximately 7 minutes per source to work. In both the science and history investigations the source evaluation portion of the investigation took two days of approximately 40 minutes each. During their work time I circulated through the room keeping the students on task and prompting them for their thinking.

Source evaluation worksheets. Each source evaluation worksheet was comprised of five parts. Figure 7 provides an example of a source from the history task. The worksheets were

details from the sources, do not become evidence until they are used to support an argument. In other words, when the students in the study read the texts they were looking at "sources", but when they used information from the source to support their claim it became "evidence."

divided into two columns. The left-hand column contained all the elements of the source material. First, each source had a source number and title. Second, there was a "headnote" included in each source. The headnote provided background information about the source itself that would help students make sense of the source. This included information about the context of the source, details about the author, and/or descriptions of the source's contents. Below the headnote was the text of the source itself. This was to be the primary focus of students' reading and basis of their arguments. Next, each source was given an attribution, which provided the author, the publication (if applicable), and date of publication. Prior studies and curriculum materials have used this format to provide students with the information necessary to engage in historical reading practices (De La Paz et al., 2017; Martin, Wineburg, Rosenzweig, & Leon, 2008).

The righthand side of the worksheets were for students to record their ideas. This part of the worksheet was the same for every source. At the top the investigation question was repeated and followed by the prompt, "What does this piece of evidence tell you about the investigation question?"

Designing the sources. The sources used in the investigations were designed with a few considerations in mind. First, the sources had to support multiple claims since there was more than one way to answer the question. The sources were written so that there were a roughly equivalent number of sources supporting each "side" of the investigation question. In addition, some sources included details that supported both sides of the investigation question. This was done to provide insight into how carefully the students were reading the sources and how they handled evidence that countered their own claims. In cases where a source could support multiple claims, it was important to see whether students acknowledged the conflicting details or

Source 4 – Report of the Democratic Meeting	**Investigation Question:** Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?
Headnote: This report comes from a newspaper article printed between the first and second conventions to vote on statehood. It describes a meeting the Democrats had. It tells what the Democrats decided to do after the first convention voted not to join the United States. It comes from a newspaper called "The Democratic Free Press." The newspaper was mostly read and written by Democrats.	What does this piece of evidence tell you about the investigation question?
We have decided: The members of the Democratic party in Wayne county are worried about the dangers of voting against statehood for Michigan. We recommend meetings of our fellow citizens in every county. At the meetings people can show they want to call another convention and vote for Michigan to join the United States.	
Source: *Democratic Free Press* Nov. 16, 1836. "Report of the meeting of the Democratic Convention of Wayne County"	

Figure 7: History Investigation — Source 4.

if they only focused on details that supported their existing ideas about the answer to the investigation question.

Second, the sources needed to reflect the kinds of materials that would typically be encountered in each field. With this consideration in mind, I considered two sets of traits for the texts: primary/secondary and text/numerical. As discussed previously, primary sources are essential to historical practice and so needed to be a part of the investigation. They are also important to science, however, as direct observations and measurements serve as the basis of explanations and arguments in science. While primary sources in science typically consist of data collected by scientists first-hand, I decided to provide students with second-hand data in this investigation. This was primarily to keep the investigations limited in time and complexity. Additionally, involving the students in data collection would have introduced a host of complicating factors when analyzing their use of that evidence in their arguments. Secondary sources are also important to both fields. Historians, for instance, may build upon previous scholarship or contrast their own interpretations of events with what others have argued. In science, secondary sources can provide the scientific principles that tie an argument together.

The second set of traits, text/numerical, describes the format of the sources, both of which are important in both fields. Historians, for instance, make use of texts such as essays, letters, and newspaper articles, as well as numerical data like population records or ledgers. In science, while quantitative data plays a significant role, scientists still make use of texts in the form of journal articles, reference texts, or qualitative observations.

Using these two sets of traits, I developed three kinds of sources: *primary textual* sources such as Source 4 in the history investigation which was an excerpt from a newspaper article from 1836; *secondary textual* sources such as Source 3 in the science investigation which was an encyclopedia article; and *primary numerical*, such as Source 5 in the science investigation which was a table presenting the results of an experiment. From these three kinds of sources, I created six total sources for each investigation so that there were two of each type of source, with each claim being supported roughly equally between each pair of sources.

Although these traits encompass a broad range of sources encountered in each discipline, there are still a wide range of formats that are excluded. Photographs, artifacts, models, and specimens are just a few kinds of sources left out. I limited the sources to the three types I described for two reasons. First are the practical limitations of using a large number of sources. I was limited in the time available and the students' attention spans. Second, the students' experience evaluating sources such as images was limited. Introducing a format or genre they were very unfamiliar with would have likely introduced further difficulty into an already challenging task.

Once I decided on the format and number of sources I created the sources themselves. I did this by either adapting existing texts or creating new ones. In the history investigation, all of the sources were adapted from existing texts. The sources in the science investigation were made by a combination of adapting existing texts and creating new ones. Because the investigation was novel (i.e., not adapted from existing curriculum materials) I had to generate both the numerical data and the primary textual sources, namely Source 4 (Edgar's notes to himself about the cart design) and Source 2 (an excerpt from an engineer's journal).

The sources that came from existing documents such as newspaper articles or historical monographs needed to be adapted so they would be readable to the students. I used guidelines proposed by Wineburg and Martin (2009) to help simplify the texts while maintaining the original meaning and key details necessary for students to construct arguments. To ensure the texts were accessible to third-graders I consulted with an elementary reading specialist and revised the text with their input. See Appendices 2 and 3 for the complete set of sources.

Previous work on students' use of evidence has included works of fiction or sources of questionable credibility (e.g., Britt & Aglinskas, 2002). The work presented here follows the lead of other studies, however, and only presented students with sources that would generally be considered reasonable. This is partly because the children in this study were so new to the practice of developing arguments and explanations from evidence. Introducing misleading or inaccurate sources would have likely made the process much more difficult. Furthermore, it would have required adding to the already large number of sources.

Constructing Arguments. After students evaluated all six of the sources, they moved on to the argument construction phase of the investigations. In this phase, students decided on an answer to the investigation question and used a graphic organizer to create an argument supporting their claim. The students in Ms. Fischer's class had not engaged in argumentative practice such as this prior to my study. Consequently, it was important to support their work creating arguments.

Scaffolds for student thinking have been used in previous research on argumentation. Bell and Linn, (2000) developed an interactive computer-based environment called SenseMaker that helped students construct arguments using evidence, promoted the sharing of ideas, and made students' thinking visible. Herrenkohl and Cornelius (2013) adapted the SenseMaker environment into a paper-based tool that students used in constructing arguments in both science and history. I adapted one of the SenseMaker boards from Herrenkohl and Cornelius for use in this study.

Figure 8 presents an example of one of the groups' completed SenseMaker boards. Each board consisted of five elements. First, the investigation question was repeated at the top of each board. Second, there was a space labeled "Claim" where students put their answer to the investigation question. For each investigation, there were three possible claims. Each of these claims was printed on a piece of paper, and the groups selected their claim and glued it directly to the SenseMaker board. Next, there were two columns labeled "Why I Think That" and "Evidence." In the "Why I Think That" column, students provided reasons that supported their claim. In the "Evidence" column, they cited the sources that their reasons came from. Each group got slips of paper with the names of each source written on them. When the students identified which source their supporting reason came from, they glued that slip onto the SenseMaker board and then drew an arrow connecting the source to their supporting reason. The students were also given slips of paper with “AND” printed on them to use to connect multiple supporting reasons or pieces of evidence. Finally, there was a space where students glued the slips of the sources they found irrelevant to their argument.

The design of the SenseMaker board in this study differs from Herrenkohl and Cornelius in two ways. First, Herrenkohl and Cornelius also provided students with “BUT” slips of paper, which students could use to include counter arguments or rebuttals. I excluded the “BUT” slips because this was the students’ first encounter with formal argumentation. I thought it was unlikely that they would readily be able to consider counter arguments in their work and did not want to confuse them. Second, I asked the students to include their reasoning on the SenseMaker board. The intention was for students to provide their source on one side, their supporting evidence on the other, and the reasoning linking the two along the arrow connecting the two. As Figure 8 shows, however, the students struggled with this and did not include it in their final arguments.

In both the history and science investigations, the argument construction phase took one day lasting approximately 45 minutes. As with the pre-survey and source evaluation phases, the students worked in their small groups. During their work time I circulated through the room keeping the students on task and prompting them for their thinking.

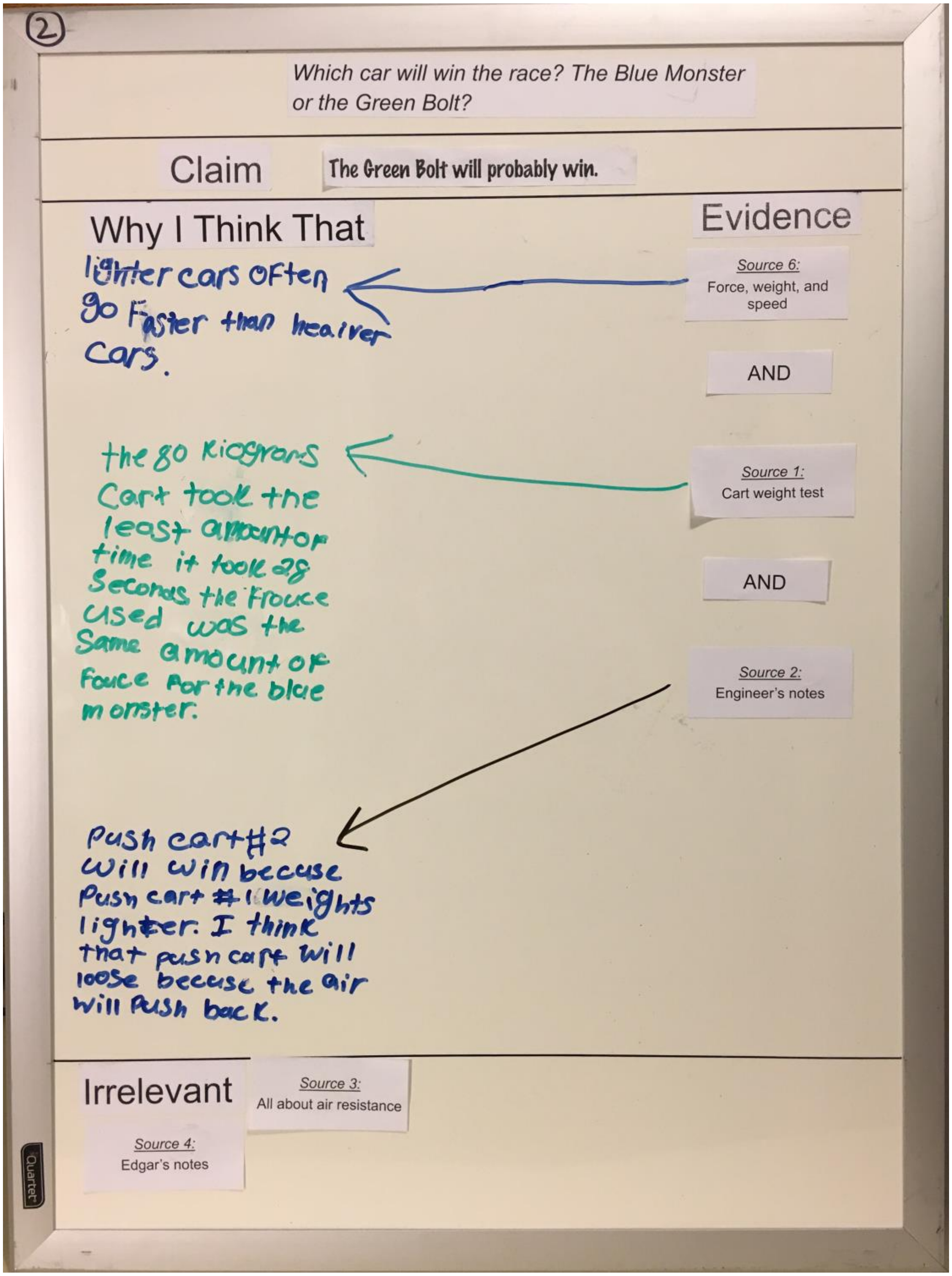
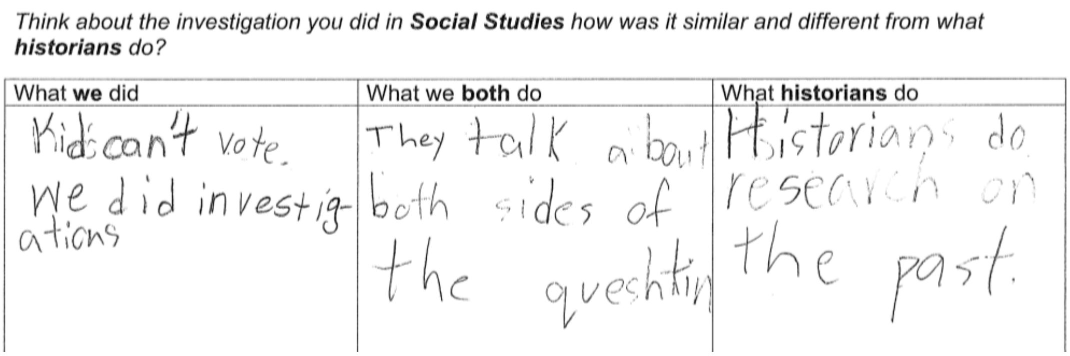


Figure 8: Group 2 — Science SenseMaker Board.

Post-Survey

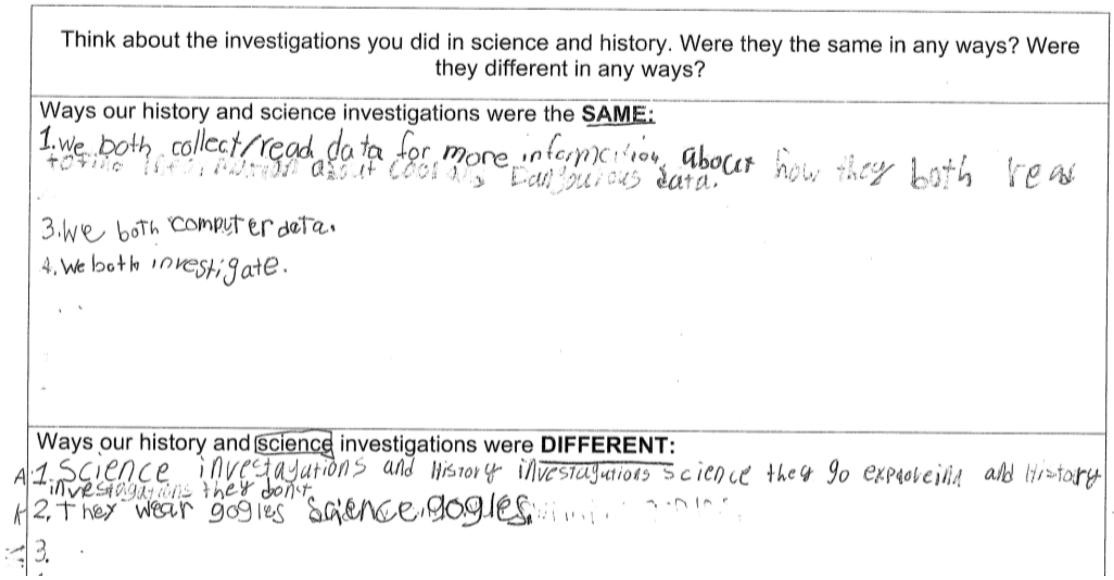
After the students completed both the history and science investigations, they completed a survey that asked them to compare what they did in each investigation with the work of scientists and historians and to compare the two investigations with one another. In the first two questions, students were asked to compare the work they did with what historians do and what scientists do. Figure 9 presents Group 3's responses to the question about historians. The question comparing scientists and the science investigation was presented in the same format.


*Think about the investigation you did in **Social Studies** how was it similar and different from what **historians** do?*

What **we** did	What we **both** do	What **historians** do
Kids can't vote. We did investigations	They talk about both sides of the queshtin	Historians do research on the past.

Figure 9: Post-Survey — Comparing the Investigation with Historians.

The final question on the post-survey asked students to consider their experiences across the investigations. They were asked to list how the investigations were the same and how they were different. Figure 10 presents Group 5's responses to the questions. The complete post-survey is included in Appendix 1. As with the pre-survey and work during the investigations, the students worked in small groups to complete the post-survey and their discussions were video- and audio-recorded.



Think about the investigations you did in science and history. Were they the same in any ways? Were they different in any ways?

Ways our history and science investigations were the **SAME:**

1. We both collect/read data for more information about how they both reas
to find information about cool and Dangerous data.

3. We both computer data.

4. We both investigate.

Ways our history and science investigations were **DIFFERENT:**

1. Science investagations and History investagations science they go exploreing and History investagations they don't.

2. They wear goggles science gogles

3.

Figure 10: Post-Survey — Comparing the Investigations with Each Other.

Interviews

In addition to collecting data from the whole class during the history and science investigations, I also conducted three focus group interviews to further probe students' thinking about argumentation and their use of evidence. Focus group interviews are interviews with a group of participants in which interaction among the members is encouraged and attended to (Barbour, 2007). One benefit to using focus group interviews is that in certain cases they can be a more efficient means of collecting data from individuals than one on one interviews (Onwuegbuzie, Dickinson, Leech, & Zoran, 2009). Given the constraints of fitting into the students' school day schedule this was an important consideration.

There are also a number of benefits to focus group interviews beyond logistical concerns. For instance, because participants are primarily interacting with one another and not the interviewer, they may be less likely to try to answer in a way that meets what they believe the interviewer's expectations are. The interaction between participants can also yield surprising results. Focus group interviews have the potential to illuminate what is important to the participants, which may not align with the researcher's expectations (Barbour, 2007). The discussion involved in focus group interviews also encourages students to respond to one another's ideas and defend their beliefs. Introducing students to the possibly unexpected ideas of their peers had the potential of making their thinking about their own ideas more visible.

I selected students for the focus groups based on a number of factors. First, the students in the focus groups needed to be volunteers to ensure that the students in the focus groups felt comfortable sharing. Second, I wanted each group to be represented in the focus group. Having one student from each group in the interviews helped to make sure that I got a range of ideas from the students' work, and not just the experiences of a single group. Third, I consulted with Ms. Fischer to select students who represented a range of skills. Because the activities were so reading intensive, I used reading level as a proxy for academic achievement and relied on Ms. Fischer's assessment of their reading level. Finally, given the first criteria, I attempted to reflect the demographics of the school as much as possible in the students selected for the focus group.

There was a total of three interviews during the investigation. The first interview took place on day 5 of the study, on the day following the completion of the history investigation (see Table 2). In this interview I asked students about their thoughts on the history investigation. They began by reviewing their arguments on the SenseMaker boards. I then went through each source and asked the students to explain how helpful the source was to their argument and how they used it in their argument to support their claim. I also asked them to select the most and least helpful sources. The second interview occurred on day 10 of the study, after they had completed the science investigation. This interview followed the same protocol as the first interview. The final interview took place on the last day of the study and was intended to complement their work on the post-survey. In this interview I went through each question of the post-survey and asked the students to elaborate on their answers and respond to one another's thinking. See Appendix 5 for the complete interview protocols.

Data Collection

Data for this study came from the students' written work and video-recordings of the small group work as they completed the tasks and interviews. The written artifacts consisted of: content tests, pre-surveys, source evaluation worksheets (6 each for the history and science investigations), SenseMaker boards (one for each investigation), and post-surveys. I also video-recorded each small group as they worked. Video-recordings were essential as discussions often involve non-verbal cues or gestures to texts or worksheets that would otherwise be invisible in an audio only recording (Driver et al., 2000). There were many instances when students silently looked at sources or referred to a source as "this one." There were also times when it was valuable to see which student was recording an answer on the worksheet or SenseMaker board. The video-recordings were transcribed and annotated with important details about the activity going on in the video, such as "Angela reading independently."

As noted previously, the data from Group 2 was excluded from this study. They received a good deal of support from Ms. Fischer, and most of their discussions centered around making sense of the literal meaning of the sources rather than utilizing the sources to construct an argument. In addition, I did not have permission from all of the students in the class to video-record their work. Consequently, I did not collect video data from Group 4. I did, however, collect their written work and that data is included in my analysis of the class's written responses. In addition, one student from the group was present during the focus group interview.

Data Coding and Analysis

This study employed qualitative case study methods, viewing the class as a case. Data analysis began by using existing literature to identify a set of analytical questions that supported each research question and data analysis. I used constant comparative analysis (Strauss, 1987) and began by open coding the data according to my analytical questions. I then categorized and refined these open codes into sets of codes describing the students' activities and traits of their work. The following sections describe the coding and analysis used to answer my four main research questions:

1. How do students interpret and make sense of different kinds of evidence in science and history?
2. How do students use evidence to construct arguments in science and history?
3. How do students coordinate evidence with claims when engaging in argumentation in science and history?
4. How do students understand the use of evidence in arguments in science and history, and the relationship between the two?

Question 1 Coding: Evaluating Sources

The data used to answer my first main research question, "How do students interpret and make sense of different kinds of evidence in science and history?" came primarily from the work the students completed during the source evaluation portion of each investigation. This included the source evaluation worksheets and four days' worth of video recordings (two for history and two for science).

In my analysis, I looked for a number of traits in their responses and discussion, and used the following analytical questions to guide my analysis:

- What was the quality of their evaluation of the sources?
- How did they interact with the text?
- What did they focus on in their reading of the text?
- How did they interact with each other?

I began by looking at their written responses on the source evaluation worksheets. These codes were a combination of *a priori* codes and emergent themes. Namely, the codes related to the accuracy, clarity, and source of the response were developed prior to looking at the data. The codes characterizing the utility of the response (i.e., “detail” and “interpretation”) emerged after my initial analysis of the students’ discussion. These codes are discussed in detail below. Table 5 provides the codes applied to the source evaluation worksheets.

The codes used for the video of students’ work were developed through open-coding the transcripts and video to identify the students’ activity as they read and evaluated the sources. These are organized around the categories of student interaction, reading strategy, and interaction with text and investigation, as described below. These codes are presented in Table 6. Because of the “messy” nature of students’ discussions, these codes were used to identify exchanges worth examining as opposed to using them quantitatively

Table 5: Source Evaluation Worksheet Codebook.

Code	Description	Example
Accuracy & Clarity		
Accurate	Response is an accurate report or interpretation of the source. Includes responses that make reasonable conclusions based on the text. Does not include accurate responses primarily about details found outside of the text.	"Democrats get five men to vote for becoming michigan." (History Source 2)
Inaccurate	Response is an inaccurate report or interpretation of the source. Includes responses that make inaccurate conclusions based on the text. Response is also considered inaccurate if it is primarily made up of details not in the source	"Michigan did not want Toledo." (History Source 1, results of vote indicate the opposite)
Unclear	Response includes an unclear referent (e.g., "they" without an antecedent)	"They were going to vote for Michigan" (History Source 4)
Copying text	The response includes text copied verbatim from the source	"many people really wanted to have a new vote about becoming a state." (copied from History Source 2)
Misinterprets Tables	Response indicates a misreading of a table.	"71 people wanted to hear and be where they are at." (History Source 3)
Relevant	Response is relevant to the investigation question. It could be used as evidence in an argument, includes a possible answer to the investigation question, or sheds light on the investigation question in another way (Can be double coded with other *Utility* codes).	"That it is Democrats wanted make Michigan a state" (History Source 4)

Utility		
Detail	Response includes a detail quoted or paraphrased from the source.	“It also tells us where it happened at” (History Source 5)
Summary	Response provides a summary of the source. Includes vague responses.	"It tells you that something can be aerodynamic." (Science Source 3)
Inference	The response indicates some interpretation of the source and goes beyond repeating or reporting a detail from the source. This includes drawing a conclusion about the investigation question, making a generalization based on the text, or stating an inference.	“It maters about the shape” (Science Source 3, source did not directly say the shape mattered.)
Origin		
Title	Response includes an interpretation or detail that comes from the source's title	“It's the results of the second convention” (History Source 3)
Headnote	Response includes an interpretation or detail that comes from the source's headnote	“(By new articals) leader's let citizen's decide to become Michigan.” (History Source 4, headnote explains the text is from an article)
Main Text	Response includes an interpretation or detail that comes from the source's main text	“The congress says that the convention that decides about statehood should be a convention of delegat elected by the people of the state.” (History Source 6)
Attribution	Response includes an interpretation or detail that comes from the source's attribution	“Mr. Marino made it” (History Source 3)
Outside Information	Response includes information or details not included in the source text (e.g., from other sources, introductory lesson to investigation)	“There opinions changed after 3 months to become a state” (History Source 1)

Table 6: Source Evaluation Discussion Codebook.

Code	Description	Example
Student Interaction		
Discussion	Students engage in an authentic back and forth discussion about the source or investigation. Excludes on-task talk in which students do not engage with each other's ideas. (e.g., "put down Source 1," or simply calling out a detail to record).	Aisha: "What do you think is evidence?" Yasmin: "I think evidence is people." Taylor: "What do you think is evidence from what it tells you here?"
Dictating	One student dictates to another what to write without any input from other students.	
Independent Writing	One student records a response independently without any input from the group.	
Role Management	Student addresses the students' roles in the discussion or manages turn-taking.	"No, recorder. She's the recorder"
Reading Strategy		
Literal	Discussion focuses on basic comprehension of the source. Examples include paraphrasing, summarizing, or directly quoting the text.	"It tells you that some things are aerodynamic"
Inferential / Evaluative	Discussion of the source includes interpretation that goes beyond the text itself. Examples include making an inference based on the text, using the text as a justification for a claim, connecting outside information to the source.	Marino: "The newspaper was mostly read and written by Democrats." Anna: "So Democrats were the ones that did wanted to make it a state, so…"
Interaction with Text and Investigation		
Read Source	Student reads directly from the source.	
Read Investigation Question/Prompt	Student reads the investigation question or worksheet prompt ("What does this piece of evidence tell you about the investigation question?")	"Who thinks the Blue Monster will win?"
Answer Investigation Question	Student suggests an answer to the investigation question.	"That it was because - they [Democrats] worked as hard - to make the democratic party…That they was working so hard - that they want to— they was working so hard to be a state."

The codes for the videos of students' discussions were strongly guided by the analytical questions provided above. Because social interaction is one key aspect of argumentation, I looked for ways the students engaged, or did not engage, with one another as they worked. The codes *dictating, independent writing, discussion,* and *role management* allowed me to describe how students interacted during the small group work. *Dictating* and *independent writing* captured examples where the students were on-task and may have made statements relevant to the investigation but were not interacting in meaningful ways. There were many examples, for instance, where students appeared to collaborate, but in fact only one student was doing the "thinking" as they dictated a response to another student who was a passive recorder of the information (i.e., *dictating*).

In contrast, the *discussion* code captured instances where the students were engaged in an actual back and forth discussion about the investigation. In exchanges coded as *discussion*, the students were not only on-task and saying things relevant to the investigation, but they also responded to or built on one another's ideas. Daniel, for example told his group that "This is the proof right here" and Zahra responded, "How?" (G5SD1:246-247). After Daniel suggested where the text supported a claim Zahra directly responded to his claim by asking him to justify it further. Later, Jason provided an interpretation of one source saying, "The lighter cart is going to be slowed down by the air pressure." Amira then said, "You just go" (G5SD1:551-552). Amira did respond to Jason's suggestion, but not in a way that challenged, built upon, or meaningfully engaged in his ideas. Instead, she focused on whose turn it was to write the response. So even though Jason provided an on-task contribution to the group's work it was not coded as *discussion* because Amira's response was not meaningful.

Because the students' arguments would rest on how they read the sources, I also looked for indications of the reading strategies that students used as they read. Based on open-coding, memoing, and prior research on reading comprehension, I developed codes *literal* and *inferential/evaluative* to characterize how the students read and interpreted the sources. Literal, inferential, and evaluative comprehension have been proposed categories of reading comprehension (Basaraba, Yovanoff, Alonzo, & Tindal, 2013; Clymer, 1968). Literal comprehension consists of the meaning that readers make of what is stated explicitly in the text and is bounded within the confines of the text itself. Inferential comprehension consists of making inferences about what the text might imply to go beyond what a text directly states and fill in the gaps left by the author. Evaluative comprehension involves bringing one's background knowledge and understanding of context to bear on a text and making meaning that extends beyond the scope of the text. In addition to being used as a framework for comprehension in a number of studies (Basaraba et al., 2013), similar distinctions between explicit and implicit meaning in the text have been made in other strands of research, such as the difference between recall questions (i.e., focused on literal comprehension) and critical thinking questions (i.e., focused on inferential and evaluative comprehension) (King, 1990).

The codes *literal* and *inferential/evaluative,* then, were defined by whether or not students were going beyond what was explicitly said in the source. In *literal* readings, the students focused on details from the source and did not provide any interpretation. This included doing things like quoting directly from the source, paraphrasing details from the source, or providing a summary of the source. In contrast, *inferential/evaluative* readings, involved students making an inference about something that was not said explicitly in the source or explicitly connecting the source to the context of the investigation, including to answer the investigation question or link two sources together.

A single source almost never provides exactly the detail necessary to support a claim or answer a question and arguments are typically stronger when supported by multiple pieces of evidence. As a result, to successfully develop a claim and construct an argument based on evidence, one needs to engage in inferential and evaluative readings of texts. It was important, then, to be able to identify the kind of reading that students did during their deliberations in order to describe their process for developing claims.

The distinction I make between literal and inferential and evaluative reading does not necessarily mean that in exchanges coded as *literal* the students were engaged in less sophisticated reading than in exchanges coded as *inferential/evaluative.* The reading context can affect what one attends to, even when recalling explicit details from a text without commentary. What is included in a summary, for instance, is often dependent on the reasons for reading a text, and a detail that is superficial in one context may be especially important in another. For example, when Taylor paraphrased a history source by saying, "It gives us information that the Whigs hated the decision" (G3HD2:751), she likely focused on the Whigs' opinion on the decision because it was relevant to their goal of deciding whether or not all of public opinion supported that decision. Even though Taylor appeared to be taking the investigation question into account here, this was still coded as a literal reading of the text. This is because the students' discussions did not typically give the information necessary to determine how much they took into account the context of reading the source and the relationship between the investigation question and source text.

Questions 2 and 3 Coding: Using Evidence to Construct Arguments

The data used to answer my second and third main research questions, "How do students use evidence to construct arguments in science and history?" and "How do students coordinate evidence with claims when engaging in argumentation in science and history?" came primarily from the work the students completed during the argument construction portion of each investigation. This included the Sense Maker boards and two days' worth of video recordings (one for history and one for science).

In my analysis, I looked for a number of traits in their responses and discussion, and used the following analytical questions to guide my analysis:

- What was the quality of their arguments?
- What kinds of justification did they provide?
- How did they interact with and use the materials (i.e., SenseMaker board, source evaluation sheets, sources)?
- What did they focus on in their use of the materials?
- How did they interact with each other?

I began by looking at their written responses on the SenseMaker boards. These codes were a combination of *a priori* codes and emergent themes. First, I determined the kinds of justifications the students provided for their arguments. These codes are presented in Table 7. Partly based on how the students read the evidence in the previously described analysis, I identified two main types of justifications, *facts* and *interpretations. Facts* consisted of responses that were either a quote from a source, a paraphrase from a source, or a detail that included information from outside the cited source. *Interpretations* were responses that consisted of the kinds of interpretive statements described in the previous section. These responses were more than just repeating or restating details from the sources. Instead, they included some kind of inference that went beyond what was explicitly stated in the source.

Table 7: Type of SenseMaker Responses Codebook.

Code	Description
Logical Connector	Response includes a logical connector, including: because, since, so.
Supports Claim	Response can reasonably be interpreted to support the claim students are making.
Reasoning Present	Response includes reasoning that ties the cited evidence to the claim.
Unclear	Response is written in a way that makes it difficult to evaluate according to the previous codes. Includes responses with an unclear referent, such as “they” or “it.”

Next, I assessed the accuracy of the justifications the students provided. In the case of *facts*, I evaluated whether or not each statement accurately repeated or rephrased what the cited source said. The cited source was especially important in this determination. There were instances where students made a statement that accurately presented a detail found somewhere in the investigation but not in the source they cited. Responses such as these were coded as *inaccurate.* For interpretive responses, I assessed the reasonableness of the inference the students made. Table 8 provides the codes used for this portion of analysis.

Table 8: Accuracy and Reasonableness of SenseMaker Responses Codebook.

Code	Description
Accurate	A response that provides a "fact" (e.g., paraphrase of source, quote from source) accurately represents what the source says
Inaccurate	A response that provides a "fact" (e.g., paraphrase of source, quote from source) inaccurately represents what the source says
Both	A response that provides a "fact" (e.g., paraphrase of source, quote from source) has both inaccurate and accurate elements
Not Applicable	The response provides a "fact" (i.e., a detail) that does not come from the named source. This includes responses of this kind that are both accurate and inaccurate.
Reasonable	A response that provides an inference from the text is a reasonable conclusion based on the text and investigation.
Unreasonable	A response that provides an inference from the text is an unreasonable conclusion based on the text and investigation.
Unclear	A response of either kind (i.e., "fact" or inference) is written so that judging accuracy or reasonableness is difficult, such as including a "they" with an unclear referent.

Finally, I applied the codes presented in Table 9 to evaluate how and how well students' justifications supported their claims. Identifying where the students' responses reasonably supported their claim and when they included the reasoning element allowed me to evaluate the quality of their overall argument. Looking for logical connectors such as "because," "since," and "so" helped make it clear when students were actively trying to link their evidence to their claim even if they didn't articulate the reasoning which connected the two.

Table 9: SenseMaker Response Justifications Codebook.

Code	Description
Logical Connector	Response includes a logical connector, including: because, since, so.
Supports Claim	Response can reasonably be interpreted to support the claim students are making.
Reasoning Present	Response includes reasoning that ties the cited evidence to the claim.
Unclear	Response is written in a way that makes it difficult to evaluate according to the previous codes. Includes responses with an unclear referent, such as "they" or "it."

When analyzing students' discussions as they developed their arguments and completed the SenseMaker boards, I used many of the same codes as when I examined their discussions during the source evaluation portion of the investigation. By applying this same set of codes to the argument construction discussions I was not only able to characterize the students' interactions, but also compare their behavior during this portion of the investigation with what they did during the source evaluation tasks. Because the goal of the argument construction task was different, I replaced the *comprehension* and *interpretation* codes with *justification.* The former two codes focused on the students' reading of the sources whereas the latter attended to when students were applying the sources to their argument. These codes are presented in Table 10.

Table 10: SenseMaker Discussion Codebook.

Code	Description	Example
Discussion	Students engage in an authentic back and forth discussion about the source or investigation. Excludes on-task talk in which students do not engage with each other's ideas. (e.g., "put down Source 1," or simply calling out a detail to record).	Angela: "Hey, you guys all have to agree with it." Solomon: "Guys, let's do both, both because—" Angela: "You have to give a reason why." Solomon: "Because more people voted, they—"
Dictating	One student dictates to another what to write without any input from other students.	
Independent Writing	One student records a response independently without any input from the group.	
Justification	Student provides justification for their claim or makes a bid for justification from the other students. Includes citing evidence, providing a logic-based rationale, and referencing a scientific principle.	"What is this? We need to get our evidence out." "The one that has more light, more light can win."
Source	Student interacts with source. Includes reading out loud from a source, naming a source, and reading a source independently.	"This source tells us that the cart that—" "What source is this from?"
Read Investigation Question	Student reads or rephrases the investigation question.	"Okay. Got to do this first. Do you think the second one had the most because the work of the democratic party? Or do you think the second vote had the most because of the population, or opinion?"
Answer Investigation question	Student suggests an answer to the investigation question.	"The second part both, because of the popular opinion and because of the democratic party." "Green Bolt! Green Bolt!"
Role Management	Student addresses the students' roles in the discussion or manages turn-taking.	"We're waiting for you to pick something." "Now it's my turn to write, what else?"

Question 4 Coding: Disciplinary Beliefs about Argumentation

The data used to answer my fourth main research question, "How do students understand the use of evidence in arguments in science and history, and the relationship between the two?" came primarily from the pre- and post-surveys the students completed at the beginning and end of the study and third focus group interview. This included their written responses and video recordings.

In my analysis, I looked for a number of characteristics in their responses and discussions and used the following analytical questions to guide me:

- What categories of things do students consider as evidence?
- Are their ideas appropriate to the discipline in question?
- How accurate are their ideas about evidence and the disciplines?
- What purpose do they think evidence serves?

I began by looking at the students' responses to the pre-survey and evaluating how appropriate each response was to the discipline in question. These codes were intended to capture the extent and strength of students' understanding of disciplinary practice in the disciplines of science and history, and the school subjects of science and history. These codes are presented in Table 11.

Table 11: Disciplinary Appropriateness of Pre-Survey Responses.

Code	Description	Examples	Non-Example
Appropriate to both science and history	Response for questions about both science and history is appropriate to both disciplines. Includes responses that may not be appropriate to the question, but are appropriate to the discipline (e.g., “to prove things” for “what is used as evidence in both?”	Bones, for proof, books	What people did in the past, record books, microscope, bones, google
Appropriate to science	Response for questions about science/science class is appropriate to the discipline. For questions about science class, the response may be appropriate to the discipline of science even if it is not appropriate to the classroom setting (e.g., fossils). Includes responses appropriate to both disciplines (e.g., bones) Includes responses that may not be appropriate to the question, but are relevant to the discipline (e.g., “tweezers” for “what is used as evidence in science?”)	DNA, microscope, bones, google	What people did in the past, record books
Appropriate to history	Response for questions about history/history class is appropriate to the discipline. For questions about history class, the response may be appropriate to the discipline of history even if it is not appropriate to the classroom setting (e.g., ruins) Includes responses appropriate to both disciplines (e.g., bones). Includes responses that may not be appropriate to the question, but are relevant to the discipline	What people did in the past, record books, bones, google	Tweezers, microscope,
Science instead of history	Response for a question about history, history class, or both science and history describes an activity, item, or tool, that is predominantly related to science. Does not include responses that are appropriate for both disciplines.	"skin" for evidence that historians use, "borax" for evidence used in both science and history class	"digging" for evidence used in both science and history, “Google”
History instead of science	Response for a question about science, science class, or both science and history describes an activity, item, or tool, that is predominantly related to history. Does not include responses that are appropriate for both disciplines.	"record book" for evidence that scientists use.	"digging" for evidence used in both science and history, Google

In addition to evaluating whether or not the students' responses were appropriate to the discipline in question, I also recorded the scientific and historical topics evident in their responses. In combination with the codes on disciplinary appropriateness, these codes helped to shed light on the depth of students' knowledge in each discipline and what aspects of the discipline they were most familiar with. Naming a number of things related to chemistry but not including responses related to a historical sub-discipline, for instance, might suggest that students' conception of science was both stronger than their ideas about history and closely related to the sub-discipline of chemistry. These codes are presented in Table 12.

Table 12: Topics of Students' Responses to the Pre-Survey.

Code	Description	Examples	Non-Example
Crime	Response describes an activity, item, or tool related to crime or crime scene investigation.	Blood, white chalk, fingerprint dust	microscope
Historical Science	Response clearly indicates a historical science such as paleontology. This includes responses which also imply archeology. (Can be co-coded with "chemistry" and "biology")	by digging, bones	Chemicals
Chemistry	Response is clearly related to chemistry. This includes responses describing materials or activities related to classroom chemistry. (Can be co-coded with "historical science" or "biology")	chemicals, baking soda, making potions	blood cells
Biology	Response is clearly related to the biological sciences (e.g., animals, human body, blood) this does not include responses clearly specifically to historical sciences (e.g., fossils) but does include borderline cases such as bones. (Can be co-coded with "historical science" or "chemistry")	animals, DNA, microscope, bones	fossils
Historical sub-discipline	Response is related to a sub-discipline or topic within history (e.g., American history, George Washington)	George Washington, the constitution	history book

Next, I looked specifically at the examples of evidence that students provided on the pre-test. Through open-coding and memoing, I developed the coding scheme presented in Table 13. These codes describe the traits of the examples of evidence students gave. This set of codes was intended to capture how well students understood the concept of evidence in general. Many responses, for instance, were appropriate to science but did not constitute an example of what could be used as evidence and were instead a tool used to collect evidence or a material used in the practice of science. This indicated that students may have had some understanding of the practice of science but were less clear on what evidence is.

Table 13: Traits of Evidence in Pre-Survey Codebook.

Code	Description	Examples	Non-Example
Tool	Response to the question "What is used as evidence?" is a tool used during inquiry, rather than a piece of evidence	Google, tweezer	bones, DNA, books
Material	Response to the question "What is used as evidence?" is a material used during inquiry, rather than a piece of evidence	potion, borax	bones, DNA, books
Topic of Study	Response is more of a topic or object to be studied rather than an example of evidence. Although response may conceivably be used as evidence, it requires a significant justification to do so. Excludes responses which could be used as evidence in a crime scene setting	Humans, homes/habitat	people, blood, DNA
Canonical Example of Evidence	Response is clearly an example of evidence. Rather than a subject of study, material, or tool, the response indicates a measurement, record (e.g., observation, drawing), or artifact (e.g., book, bones) that is typically used as evidence. Must be appropriate to the discipline asked about (e.g., "DNA" is acceptable if given as a response to a question about history, but “fossils” is not)	DNA, bones, record book	humans, tweezers, borax, buildings

The final set of codes captured what students believed the purpose of evidence in each field to be. Working from my open coding of the students' responses, I developed three main purposes, presented in Table 14. Although the codes *research/investigate* and *discovery* appear to be similar, they do capture different ideas in the students' responses. Responses coded as *research/investigate* focused on an individual learning knowledge that was already established. Here, I mean "research" in the sense of writing a research paper compiling information from existing sources of "established" knowledge as opposed to conducting original research, such as the kind of research historians might do in collecting, comparing, and corroborating sources to draw a conclusion. This is research that involves the kind of *knowledge telling* that students often do in school compared to the *knowledge transformation* that practitioners engage in (Scardamalia & Bereiter, 1987). In contrast, *discovery* is for responses that reference discovering or developing previously unknown knowledge, such as "They get new stuff like fossils" (G1Pre) or "They find out new things" (G5Pre).

Table 14: Students' Purposes for Evidence in the Pre-Survey.

Code	Description	Examples	Non-Example
Research / Investigate	Response describes a process or activity related to researching or investigating a topic.	Research, study, books, Google	for proof, drawing
Discovery	Response describes a process or activity related to discovery or building knowledge.	Find, know about, learn	for proof
Proof	Response describes proof, finding the correct answer, or justifying beliefs to others as the purpose of an activity, tool, or item.	So people can believe me, for proof	Research, learn

I analyzed the post-survey using a combination of emergent codes and codes carried over from the pre-survey. Some of the new codes, however, were connected to the themes identified in the coding schemes for the pre-survey. *Experimentation/data collection*, for instance, is only a code for the post-survey but is closely related to the *tool* and *material* codes from the pre-survey. Responses coded as *experimentation/data collection* often incorporated or implied tools or materials used in science. The code was different from the pre-survey because the questions were slightly different. In the pre-survey, students were asked to provide examples of evidence, thus it was important to distinguish the different types of things students provided as an example. In the post-survey, on the other hand, students were mostly describing the activities related to the investigations and disciplines, so the code needed to capture the processes they were describing rather than objects. Table 15 presents the codes used to analyze the post-surveys.

In my analysis of the video recordings for both the pre- and post-survey, I used the codes for analyzing their written work to identify exchanges that would shed light on their written responses. When possible, I also connected the transcript to specific written answers in order to clarify their responses or characterize their thinking about it. For instance, I was able to link Group 5's written response of "record books" to an exchange in the transcript, which revealed they meant books such as the *Guinness Book of World Records* and not books of historical records.

Interrater Reliability

Two researchers coded all of the students' written responses from the pre- and post-surveys, source evaluation task, and argument construction. The two raters agreed at least 90% of the time for all coding schemes related to written responses. After performing an inter-rater reliability check, the two researchers reviewed the coded data together and discussed instances of disagreement about the codes. The two researchers reached 100% agreement through these discussions.

Table 15: Characteristics of Students' Responses to the Post-Survey.

Code	Description	Examples	Non-Example
Research / Learning	Response describes a process or activity related to researching or investigating a topic.	Research, study	for proof
Discovery	Response describes a process or activity related to discovery or building knowledge.	Find, know about, learn	for proof
Proof	Response describes proof, finding the correct answer, or justifying beliefs to others as the purpose of an activity, tool, or item.	So people can believe me, so they know it's true	to know it
Scientific Sub-Discipline	Response names a process, tool, or item related to a scientific sub-discipline.	Fossils, potions	Book, scientist
Historical Sub-Discipline	Response names a process, tool, or item related to a historical sub-discipline.	George Washington, the Constitution	history book
Experimentation / Data Collection	Response names or suggests scientific experimentation or a means of collecting data in science.	Experiment, chemical reaction, go exploring	Read, write
Evidence	Response names or suggests using evidence. Includes references to "sources."	We used sources, collect evidence	Put fossils together
Content	Responses names or suggests the topic of the investigations, or topics within the fields of science or history.	Different topics, presidents from long ago	They study history
Collaboration	Response names or suggests collaboration, working together, or teamwork. Does not include response that just say "we" or "both" without explicit reference to collaboration.	We work together, we get help from people	We both study
Reading / Writing	Response names reading or writing as an activity.	Write books	Google
Other	Response does not fall into any of the other categories.		

Overarching Analysis

To analyze the data related to students' reading of sources, I began by running descriptive statistics on the codes applied to their written artifacts. This allowed me to identify trends in their responses, such as their general accuracy and the extent to which their responses captured literal (i.e., summaries and details) versus inferential/evaluative reading strategies. I also identified, when possible, exchanges in which students discussed specific responses. In addition to linking these exchanges to responses, I also used the video data to flag which students were recording the answer and what sources, if any, they were using as they worked. I used this information to clarify students' written responses, which were often vague, and revise the coding of the written artifacts when necessary.

After characterizing the trends in students' written responses, I went on to examine their discussions. After coding the video data, I first entered the codes into data matrices to identify trends in how the groups read the evidence and constructed arguments. I then created descriptive summaries of each group's work. Using these trends and summaries as a guide, I developed theories about the students' thinking and work and returned to the video data to test those theories. My final analysis was the result of an iterative process of developing and testing theories across groups.

Conclusion

This chapter described the setting, participants, methods, and analysis used in this study. I used qualitative case study methods, treating the class as a case, to better understand how students without prior experience in formal argumentation engage in the practice in science and history. I developed an activity in which students were posed with investigation questions in science and history, analyzed sources to develop an answer to those questions, and then created an argument using a graphic organizer. In addition, I administered surveys asking about students' ideas about argumentation and evidence in the disciplines. Through a process of open-coding, memoing, and categorization, I developed coding schemes to analyze students' ideas about evidence and argumentation and how they engaged in the practice. Focus group interviews were used to triangulate the data when necessary.

The next chapters provide the results of my analysis. First, Chapter 4 analyzes the students' responses to the pre-survey and their initial ideas about evidence and how scientists and historians do their work. Chapter 5 presents my analysis of how students read and interpreted evidence in the investigations. In Chapter 6, I describe how students worked together to generate their arguments and characterize the various aspects of those arguments. Next, in Chapter 7, I review their responses to the post-survey and their reflection on the activities and their connection to the disciplines. Finally, Chapter 8 contextualizes these findings in the relevant literature and discusses the implications of this study for future research and classroom pedagogy.

CHAPTER 4

Students' Initial Ideas about Science, History, and Evidence

In this chapter I present findings related to the students' ideas about the disciplines of science and history, the school subjects of science and social studies, and evidence in these four domains. Based on my analysis of the pre-surveys and the recordings of the small groups as they completed the survey, I make five major assertions:

A) The students had a vague initial understanding of disciplinary activity in science and history.

B) The students had a limited initial understanding of what constitutes evidence in the disciplines of science and history.

C) The students had reasonable and appropriate beliefs about the reasons for using evidence.

D) The students had generally stronger conceptions of science and science class when compared to history and history class.

E) Students' ideas about science and history as school subjects differed in comparison to their ideas about the disciplines practiced by professionals.

How Scientists and Historians do their Jobs

I begin by discussing students' beliefs about how scientists and historians engage in their work and develop knowledge in their fields. This is based primarily on students' responses to the first question in the pre-survey, Figure 11.

Think about how scientists do their job and how historians do their jobs.

How do **Scientists** do their job?	How are they **similar?**	How do **Historians** do their job?
• they find new discvrys • they get new stuff like focals • they make cemicall reacthons • they observe.	• By traveling • and Discovaring new things. they Put thing together ~~they set up~~ • they work together to find new discuverys	• they find new species of animals from the past. • By diging By comparing other animals

Figure 11: Beliefs pre-survey – “How do scientists and historians do their jobs?”.

Three main trends were apparent in the students’ responses to this question. First, the students often gave vague responses that indicated they had a general sense of scientists’ and historians’ work but lacked clarity on more specific aspects of that work. Second, they generally had a stronger understanding of scientific practice than of historical practice. For science, they understood that experimentation and data collection were part of scientists’ work and they demonstrated an awareness of some of the fields of study within science. In history, on the other hand, their definitions of disciplinary practices were vaguer and more inaccurate, and they had more explicit questions about what historians were and how they do their jobs. Additionally, they often confused science for history and listed activities relevant to the historical sciences (e.g., paleontology) or other sciences as part of historical practice. Finally, although their general definitions of these practices were somewhat weak, they did believe that research and discovery were important elements of the work in both disciplines.

Vague Conceptions of Disciplinary Activity

One indication of the students' vague conceptions of the disciplines is from the appropriateness of their responses to the question "How do (scientists/historians) do their jobs?" Table 16 presents the results of coding their responses to this question based on the appropriateness of their response to each discipline. In this table, "Strictly Science" and "Strictly History" mean that the response was appropriate to only that discipline and could not be applied to another area of study, and "Both" refers to practices that could apply to both disciplines. For instance, "They make chemical reactions" is a response that strictly applies to science, whereas "They look on Google" and "They work together and talk together" are appropriate to both science and history. The table shows a generally weak understanding of the disciplines because a minority of their answers for what scientists and historians do were strictly appropriate to that discipline— only 23% for history and 36% for science. Instead, the majority of their responses for science were in fact applicable to both disciplines, such as "There [they're] very brainy," and "They do research." For how historians do their jobs, the majority of their responses actually described activities that *scientists* engage in. Responses from three different groups, for example, included "animals" in their description of historical activity: "They [find] other animals," "They study animals to get information," and "They find new species of animals from the past."

Table 16: Appropriateness of Responses for Disciplinary Activity.

Discipline in Question	Strictly Science	Strictly History	Both	Other	Total
Science	5 (36%)	0 (0%)	9 (64%)	0 (0%)	14 (100%)
History	7 (54%)	3 (23%)	3 (23%)	0 (0%)	13 (100%)
Both	2 (18%)	0 (0%)	7 (64%)	2 (18%)	11 (100%)
TOTAL	14 (37%)	3 (8%)	19 (50%)	2 (5%)	38 (100%)

Second, the students' responses suggest that students' understanding of the disciplines is vague. This is evident in the number or responses coded as applying to both science and history. A majority of their responses (64%) for "How do scientists do their jobs?" described activities that were appropriate for both science and history. These responses described dispositions and attributes common to most academics (e.g., "They are serious"), the learning involved in careful study of any topic (e.g., "They read books"), and general activities applicable to most disciplines (e.g., "They investigate"). Table 17 presents the students' responses that were coded as applying to both science and history. Responses such as "They find new discoveries" and "They are super smart" show that the students have a sense that both disciplines are related to generating knowledge, but not of how practitioners achieve their goals.

Table 17: Responses Coded as "Both" for "How do (scientists/historians) do their jobs?".

Responses for Science	Responses for History	Responses for Both Science and History
they read books	they look on google	there both help each other
they study everything	by digging	they are serious
they investigate	they make books	they are super smart
they find new discoveries		by traveling and discovering new things
they find out new things		they work together to find new discoveries
there very brainy		
they learn new things		
they work together and talk together		

Students' Understanding of Science

Although the students had a difficult time articulating what they thought scientists did, they did have some beliefs about scientific practice that were accurate. First, some of the students appeared to know that experimentation and data collection were an important part of scientific practice. Second, they demonstrated a knowledge of fields within science, which included chemistry, biology, engineering, and paleontology.

Experimentation and data collection. Some of the students' responses indicated that they believed that experimentation and data collection were part of scientific practice. Two of their written responses, "they make chemical reactions" and "they do their job by making chemical reactions" suggest that the students knew something about experiments and chemistry. There were also more explicit mentions of experiments in their discussions. Students in Group 5 said that scientists did their jobs by "testing" (G5PreD1:6), proposed that scientists "test their inventions" (G5PreD1:27), and then later suggested that "They do their job by—they do their job by doing experiments on people—" (G5PreD1:80).

Relatedly, Jason, a student in Group 5, made numerous mentions that scientists collect data and information as part of their job. This first came up in the following exchange in which he was prompting his group mates for more input.

> Jason: But how do they do their job? How do they—like what do they use—
>
> Daniel: Test? Oh, what do they use?
>
> Jason: What kinda data do they collect?
>
> Daniel: They collect data about their machines and stuff, or about stuff they use (G5PreD1:28-31)

The idea of scientists collecting data seemed to be important to him as he repeated this later on, suggesting "They do investigate by weather data," "They collect data from— they collect data information," and "They collect data, information from rabbit species—" (G5PreD1:41; 100; 129). It is worth noting that Jason came up with this language on the first question of the survey, and prior to being asked about the kinds of evidence that scientists and historians use. This suggests that he had existing ideas about data that were not initially influenced by the questions on the survey.

Knowledge of fields within science. Several students also showed that they had knowledge of some of the fields within science and understood that science included a range of topics of study. Their responses demonstrated a knowledge of chemistry, biology, and engineering. In addition to two of the groups' written responses about chemical reactions, Zahra in Group 5 suggested that scientists do their jobs by, "figuring out their potions" (G5PreD1:1). The students in Group 5 also demonstrated their knowledge of biology as a field of study within science. Jason proposed that scientists collect data about the "human body" (G5PreD1:32) and Daniel suggested that, "they collect information from a— species" (G5PreD1:76). Daniel's suggestion then sparked a discussion among the students in Group 5 that showed the students' background knowledge in science that may have been supporting their beliefs about scientific practice. As a follow up, Jason asked, "From any species? … Wait, do you think we should give a pacific [specific example]?" (G5PreD1:101-105). In response, the students generated a list of possibilities that included: "mammals, spiders, bacteria, rabbits, birds," and "any type of fish" (G5PreD1:106-117). Finally, Daniel indicated a nascent understanding of engineering by proposing that scientists "test their inventions" and "collect data about their machines and stuff" (G5PreD1:27; 31).

The students also seemed to have particularly strong knowledge of paleontology. Surprisingly, this was mostly evident in their discussions, addressed below, about how historians do their jobs. This is somewhat problematic, however, as it indicates a knowledge in the field but confusion about what discipline it falls under. This is discussed more below in discussing their beliefs about historical practice.

Students' Understanding of History

In comparison to their understanding of science, the students had a more complicated and comparatively limited understanding of history. Many of their responses were vague or did not strictly apply to the study of history, whereas more of their responses about science were accurate or at least appropriate. In addition, there was a good deal of confusion about historical practice in their discussions. Some students questioned what historians were, while others had accurate, but vague, ideas of historical practice at the same time other students disagreed with those suggestions. In addition, students often confused science with history. There were many instances where they believed the historical sciences, paleontology in particular, were part of historical practice. And, in a few other cases, they described biology as being part of history.

Vague understanding of historical practice. The first indication of the students' weaker understanding of historical practice comes from their written responses to the survey. Only three of their 13 responses to "How do historians do their jobs?" were strictly related to history, and another 7 were strictly related to science (Table 16). In contrast, the students never listed historical practices as part of the work of scientists, and all of those responses were either directly or at least somewhat related to science (i.e., coded as "both"). Some of the students' lack of clarity about historians' work was also evident in their discussions, with two students explicitly expressing their confusion. Eric, in Group 1 asked, "What are historians?" (G1PreD1:74) when working on the pre-assessment with his group and Amira in Group 5 asked "What does 'historians' mean?" (G5PreD1:164) after reading the prompt on the survey.

Although not all of the students were as explicitly confused as Eric and Amira, some of the other students' responses indicated understanding of history, but one that was still vague or unclear. Some students provided tautological answers to the question "How do historians do their jobs?" Group 3 wrote that, "They study history," and Amira said that historians were "People that investigate history" (G5PreD1:167). These responses establish that the students had some idea about the goal of history but not necessarily how that is accomplished. Jason's definition was even more vague when he proposed that they got "information on data" (G5PreD1:177).

Even when students gave an appropriate description of historians or historical practice, there was typically disagreement from other students. In Group 1, Angela suggested to the group that, "Historians are people that try finding things in the past, like history" (G1PreD1:76). Solomon, however, immediately interjected with, "Like dinosaurs" and Eric followed up with, "Like discover new species they haven't found yet in the—that were from the past" and Caitlin later agreed that they "Dig to find new animals" (G1PreD1:77-82; 103). In Group 5, Zahra was on the right track when she responded that historians do their jobs by, "Taking out historic books… By reading historical books" (G5PreD1:170-175). Jason picked up this idea and said, "—Define history, like old records, longest or something shortest" (G5PreD1:171-173). At first, it seemed like Jason's response provided an excellent example of how historians do their job. His comment about "longest or something shortest," however, revealed a limited understanding of historical practice because, as he explained later, he was talking about the kinds of records that would be in the "Genius [Guinness] world record books" (G5PreD1:563).

Another of Group 5's exchanges demonstrates that while the students sometimes had ideas about historical practice that were close to being accurate, there was still a good amount of uncertainty or confusion.

Jason: Okay, so get information. What else do they do?

Amira: They find out about dead people and how they die.

Jason: No, that's not what they do. No. They study—

Daniel: They maybe watch—

Jason: What do they study?

Daniel: They make history.

Jason: So they study about—

Zahra: They study history.

Jason: —say animals, or—

Amira: History,

Jason: They study about history, like—What do you mean by history, so I mean like animals—

Daniel: They make history.

Amira: People that are not famous that passed—

Jason: No, you put different people, because even though if they might be about history. Well, we could do animals you guys. Who wants to do dogs? (G5PreD1:179-192)

Ultimately, the group recorded "They read historical books to get information" on their survey, but that response belies the confusion evident in this discussion. The students clearly had the idea that historians deal with history but did not seem to understand what "history" entailed. Both Daniel and Zahra reference "history" but do not say anything about what that meant. Amira's idea that historians find out about "dead people and how they die," on the other hand, was an accurate, though simplistic, picture of what historians do. Jason, however, disagreed for reasons that were unclear and later offers that history has to do with studying animals. Taken together, these exchanges and written response suggest the students had a nascent understanding of the topic of history but less clarity on how historians do their jobs.

Confusing history and historical sciences. Besides having an unclear definition of history and historical practice, students showed confusion about what historians actually do, and often confused evidence or activities in science for what historians do. Seven responses out of the 13 responses to "How do historians do their jobs?" named something that was related to science instead of history, and three more responses were applicable to both science and history (see Table 16). Four of the responses that were appropriate to science instead of history ("by digging," "they find new species of animals from the past," "a fossil digger," and "they clean fossil") were related to the historical sciences (e.g., paleontology, geology). In addition, one group said that historians and scientists "both study dinosaurs."

Although "digging" and "fossils" are relevant to the field of archeology and not just paleontology, it is unlikely that the students had this in mind when recording their responses. Group 1 recorded "by digging" on their survey worksheet, and it was clear from their discussion that they were referring to paleontology. In the discussion, Eric suggested that historians, "like discover new species they haven't found yet in the—that were from the past" (G1PreD1:82). Caitlin picked up on this idea, followed up with, "Like digging" and later said, "Yes, they dig to find new animals. … They'll find something, like what are they called? … Fossils, dinosaurs" (G1PreD1:83; 103-108). These students had a clear idea that history deals with the past but seemed to be confused about the difference between natural history and human history.

Group 1's discussion also seems to have been influenced by a book about Benjamin Waterhouse Hawkins that the students all seemed to be familiar with. Waterhouse Hawkins was an artist who produced sculptures for a variety of museums and academic institutions in the midto late-1800's. In addition to creating reconstructions of dinosaurs, he is credited with constructing the first mounted dinosaur skeleton. He came up after the exchange above, when Caitlin said, "Like that story that we read … Remember the Waterhouse Hopkins or whatever it was called? What about that?" (G1PreD1:112-115). At this, the other students acknowledged remembering the book, and Eric went on to recount his experience of when he, "was at the Natural History Museum" and "saw, like, the littler sculptures that he—that–of the dinosaurs he made, and other animals" (G1PreD1:121). In addition to remembering the book read in class, the visit to the "Natural *History* Museum" is very likely influencing Eric's beliefs about historical practice. This confusion between natural and human history continued as the groups discussed other questions on the survey.

Confusing other sciences and history. Biological science was another set of activities that students attributed to historians instead of scientists. On the survey, Groups 1, 2, and 5 responded that historians "[compare] other animals" (1), "[find?] other animals" (2), and "study animals to get information" (5). Additionally, Group 3 said that both scientists and historians "study ocean life." It is unclear as to why the students described historians as doing these things, but there may be some connection to paleontology and the historical sciences in some of these cases. Group 1, for instance said that historians both "find new species of animals from the past" and "[compare] other animals." The second response is not explicitly tied to the historical sciences, but it may be that the students were describing the comparison of animals that takes place when finding new species of animals from the past. Similarly, Group 2 responded that historians "other animals" (without a verb). My interpretation is that the students meant to say they "find other animals," and that this is related to their beliefs about paleontology. The responses prior to this one were that a historian is a "fossil digger" and that they "clean fossils." It is possible that students were thinking historians find “other animals” by digging and cleaning fossils. If that is the case, this reinforces the conclusion that some students were conflating history and the historical sciences.

Uncertainty about social studies. Although this part of the survey did not ask about the school subject of Social Studies, it is worth noting that later on several students had a difficult time articulating what social studies is, much the same way they were challenged by defining historical practice. There were indications that some students did not know what social studies was or believed that they had never done social studies in school. When they were asked to consider whether they used evidence in social studies one student remarked “Do we even do social studies?…We don’t do social studies” (G1PreD2:9,12). A discussion in Group 3 on the second day of the activity was sparked by one student saying, “I’ve never done social studies.” This started the following exchange about what social studies actually is:

> Brandon: I’ve never done social studies.
>
> Aisha: I have.
>
> Student: What have you—
>
> Taylor: What is even social studies?
>
> Yasmin: I’ve only done social studies.
>
> Brandon: Social studies is learning about history.
>
> Aisha: Is it?
>
> Taylor: No, social studies is math and read—social studies is math, reading—
>
> Aisha: It’s learning about our environment.
>
> Yasmin: No, it’s not.
>
> Taylor: Yeah, learning about our environment.
>
> Yasmin: That’s math.

Taylor: Reading, learning about our environment. It's a whole bunch of other stuff.

Brandon: No, no. Social studies is about history.

Aisha: Seriously? No, it isn't. No, it isn't. It's not even about our environment.

Taylor: Well, sometimes you might learn about your environment in social studies. (G3PreD2:19-36)

Here there was obvious disagreement and uncertainty about what the school subject of social studies covers and its role in the school day. Brandon had the accurate idea that it includes the study of history but asserted that they had not done that in school. Yasmin's response that she has *only* done social studies is unusual because she did not seem to agree with Taylor or Aisha's definitions that encompass a wide variety of activities. It is unclear, then, what exactly Yasmin thought social studies is. Taylor and Aisha were the most confused. Aisha suggested that "it's learning about our environment," but then disagreed with herself later on. Taylor, on the other hand, proposed that social studies covers nearly every other subject *except* social studies. Learning about the environment could conceivably fall under the umbrella of social studies, as that can easily be framed as a socio-scientific issue which might face students with both the science and politics or civics of the environment. There is little in what Taylor said, however, that indicate that is what she had in mind.

Research, Discovery, and Disposition

Although the students' knowledge of disciplinary practices in science and history was limited, they did seem to understand that research and discovery were important to both scientists and historians. Written answers such as "they find new things," "by traveling and discovering new things," and "they find new discoveries" show that the students held a strong belief that both scientists and historians were responsible for constructing new knowledge in their fields. Table 18 presents the rest of the students' written responses coded as "research" or "discovery" In addition to making discoveries, the students also indicated they believed that learning and studying were an important part of disciplinary practice. This is evident in the responses which reference "research," "study," and reference materials such as "books" or "google."

Table 18: Responses coded as "Research" or "Discovery" for "How do scientists/historians do their jobs?"

Discipline	Research	Discovery
Scientists	they read books they study everything they investigate	the get new stuff like fossils they find new discoveries they find out new things they learn new things
Historians	they look on google they study history they read historical books to get information they study animals to get information	they find poop and clean it they find new species of animals from the past they [find?] other animals
Both	they both study dinosaurs they do research they study they both do the same research they study ocean life	they work together to find new discoveries by traveling and discovering new things

There were also many students who felt there were certain dispositions or characteristics that were important to being a scientist, historian, or both. In the students' eyes, collaboration was a key aspect of both practices. They stated that both scientists and historians "work together," and "both help each other," and that scientists also "work together." Additionally, practitioners needed to be "super smart," "brainy," and "serious" to do their work. While the attention to intelligence is not particularly surprising, the mention of collaboration was unexpected. Typical narratives about scientists in school focus on individuals making discoveries independently, such as Newton "discovering" gravity, or Einstein developing his theory of relativity. Consequently, the fact that scientists work together is often obscured or hidden. Unfortunately, the students did not discuss these responses in their small group work, so it is unclear where they got these ideas. Notably, however, the students only made mention of collaboration for "scientists" and "both" but not for "historians" on their own. There are two possible explanations for this. One, the students genuinely saw collaboration as part of both practitioners' work, so they did not include it under the "historians" column because it applied to both. It may also suggest, however, that they may not have believed or known that historians work together. Many of the students' other responses in the "both" column, were actually only applicable to scientists, suggesting their view of science was dominating their interpretation of the "both" column. If this is the case, the fact that collaboration was a part of the "both" column but not the "history" column may mean that they still saw it primarily as a practice of science.

What is Evidence?

In this section, I describe students' initial understanding of, and beliefs about evidence in the disciplines of science and history and the school subjects of science and social studies. This is gathered from students' answers to the questions "Make an idea web of different kinds of evidence (scientists/historians) use" and "Make an idea web of as many things as you can that you use as evidence in (science class/social studies class)," see Figure 12 for an example.

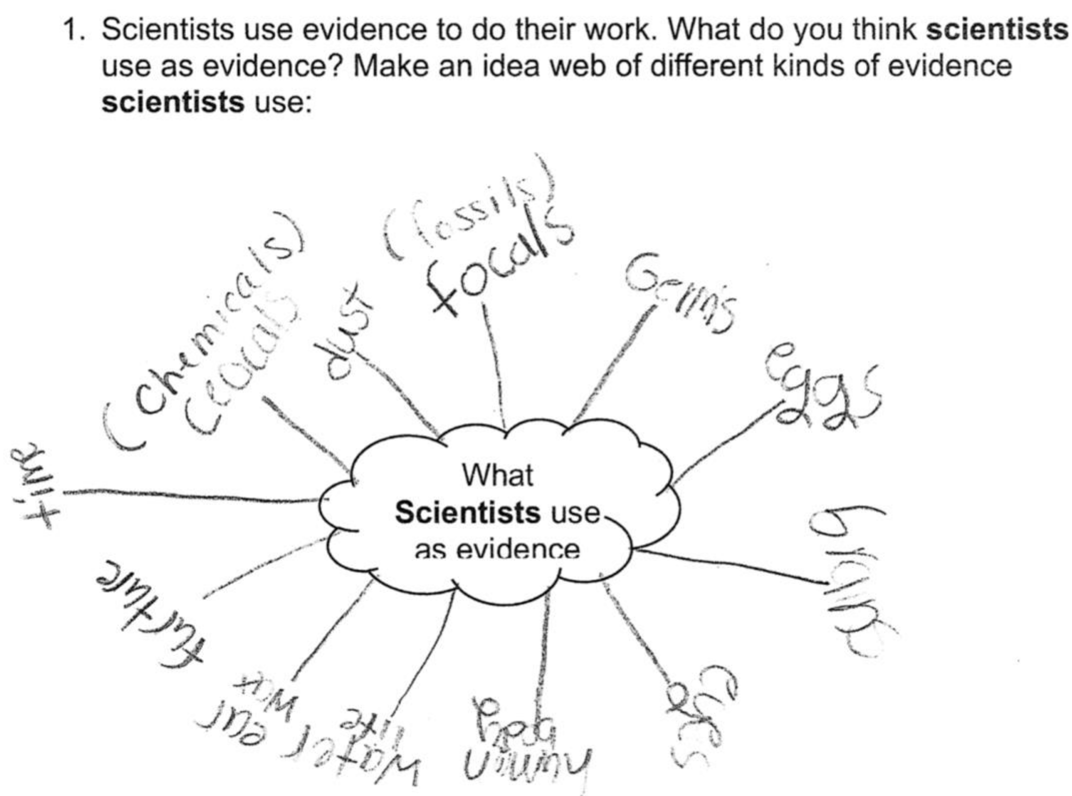


Figure 12: Beliefs Pre-Survey — What is used as evidence?

Based on the students' answers to these questions the students had a limited understanding of what can be used as evidence or even what evidence is in these domains (i.e., science, history, science class, social studies class). In addition to typically providing generic examples of evidence, the students often provided subjects of study, tools, and materials related to the domains. Their responses to these questions support the previous finding that they had a fuller understanding of science, but also reveal that their understanding of *evidence* in science was nonetheless weak.

Weak Understanding of Evidence

In general, most students had a limited understanding of what constituted evidence in the disciplines of science and history and the school subjects of science and social studies. As Table 19 shows, only 18% of students' responses to the questions about what can be used as evidence in were coded as "canonical" examples of evidence (e.g., blood, fossils, fingerprints). The relatively few examples of canonical evidence in the students' responses indicates a weak understanding of what evidence is.

Table 19: Responses to "What do X use as evidence?"

Discipline	Total Answers	Appropriate to discipline	Canonical Examples of Evidence	Tool*	Material*	Subject of Study*
Science	65	55 (85%)	13 (20%)	15 (23%)	4 (6%)	25 (38%)
Science Class	32	27 (84%)	4 (13%)	3 (9%)	6 (19%)	6 (19%)
History	32	14 (44%)	2 (6%)	6 (19%)	1 (3%)	14 (44%)
Social Studies Class	24	15 (63%)	8 (33%)	3 (13%)	4 (17%)	1 (4%)
Total	153	111 (73%)	28 (18%)	27 (18%)	15 (10%)	46 (30%)

* These responses may or may not have been appropriate to the discipline.

Table 20 presents all of the responses coded as "canonical." Although they were able to provide a number of canonical examples of evidence in science, they had a harder time listing canonical examples of evidence in the domains of science class, history, and social studies class. Furthermore, the responses they did provide for these areas were vague and fairly generic examples of evidence that could be applied in any discipline. If students had a strong understanding of what constituted evidence in these fields, they would likely have given more detailed answers. "Drawing" and "pictures," for instance, were given as examples of evidence used in science class and social studies class. It is uncertain, however, if the students had an understanding that there would likely be differences between the drawings and pictures used in the two areas. In particular, they do not indicate that those drawings and pictures should be of something that is important to those fields, and that the content of the picture is one thing that determines its usefulness as evidence. “Picture of pottery,” for example, is a response that would demonstrate that the students had a clearer understanding of the kinds of evidence relevant to social studies class (e.g., artifacts that provide evidence of the events of the past), and that it might be different than a picture used in science class (such as a picture of an organism, which might help one understand characteristics and adaptations).

Table 20: Students’ Canonical Examples of Evidence.

Discipline	Response
Science	Blood/blood stains (3 instances), bones (2 instances), books, DNA, finger prints (3 instances), fossils, hair/fur, teeth
Science Class	books, drawing, mass, pictures
History	birthdays, books, DNA
Social Studies Class	books, drawing (2 instances), pictures, videos, what we see, writing (2 instances)

Similarly, “books” is another notable response because it was given as evidence in all four domains. None of the groups’ written responses, however, specified the kinds of books or the content of the books that would make them useful as evidence in any specific domain. Historians, for instance, use a variety of kinds of books and multiple kinds of documents in their work. They use newspaper articles, diaries, government records, speeches, and books from the past just to name a few. The students’ answers suggest a general sense that practitioners in the field need to use facts and information, which can be found in books and images, but lack a clear idea of what those texts and sources actually are.

Although the written responses did not articulate the kinds of books used in history one of Zahra’s comments during her discussion with Group 5 did suggest a possible understanding of the kinds of evidence necessary in historical practice. She told the group that historians do their jobs by “reading historical books” (G5PreD1:175). On its face, this suggests that Zahra knew that historians used books from the past to do their work and learn about events from the past. Unfortunately, however, there is no other mention or discussion of "historical books" in this group. It is unclear, then, if she was talking about books from the past or books *about* history.

The fact that students saw books as being important to all disciplines may also be connected to their beliefs that research is part of scientists' and historians' work (discussed above) and that evidence is used for proof (discussed below). If research is important to how scientists and historians do their jobs, books are essential to do this work. In addition, books are a definitive source of knowledge in the classroom and when a student needs to demonstrate that something is true, they often turn to books for backing. Consequently, books can play an important role in all the disciplines.

Tables 19 and 20 also show that students were able to provide the greatest number of canonical examples of evidence for the discipline of science. Their responses for what is used as evidence in science were also very specific, especially compared to their responses in the other domains. Additionally, they were more likely to provide domain appropriate responses for science and science class than they were for history and social studies class. Their comparatively better understanding of evidence in science is discussed below.

Topics, Tools, and Materials as Evidence. Besides providing vague examples of evidence, the most common difficulty students had in providing examples of evidence was telling the difference between evidence and subjects of study. Although this confusion was apparent in all subject areas, it was especially true for science where 38% of their responses were coded as "topics of study." See Table 21 for example responses. For example, students listed "gems," "amoebas," and "lava" as items that scientists might use as evidence. While there are situations where these could arguably be considered evidence (e.g., the presence of amoebae as evidence that water is unsafe to drink) it seems unlikely that the students had these situations in mind. There were no indications in their discussions that suggested the students considered these things as evidence supporting a conclusion. Students gave similar responses in the other disciplines as well, including "people" and "what people did in the past" for history, "history" for social studies class, and "rocks" and "weather" for science class. While these are all appropriate topics of study in these areas, they are not good examples of evidence.

Table 21: Sample Topics of Study, Tools, and Materials Listed as Evidence

Discipline	Topic of Study	Tool	Material
Science	Amoebas, brains, homes/habitat, lava, skin, water life	Microscope, pliers, tweezers, white chalk	Chemicals, potions
Science Class	Rocks, weather, people	Computer, technology, timer	Baking soda, materials, chemicals/borax, mitoes [Mentos]
History	History, houses, jobs, what people did in the past	Cameras, computers, tools	No responses
Social Studies Class	History	Cameras, video recorders	No responses

Another confusion the students had was about the difference between evidence and the tools used to collect evidence, such as "microscope," or the materials involved in doing science or history, such as "chemicals." Groups gave tools and/or materials as examples of evidence for all four questions (scientists, historians, science class, social studies class). While some of these responses may have been appropriate tools for collecting evidence in a particular domain, such as timer in science class or computers for history, they were still not examples of something that could itself be used as evidence.

Most of the tools that students listed were provided as examples of evidence in science. This included, among other things, "computers," "gloves," "goggles," and "magnifying glass." In addition, although the group did not record this on their survey, Jason from Group 5 repeatedly suggested "biometric scanner" as an example of scientific evidence. At another point, Zahra suggested tweezers as an example of evidence and Daniel agreed, saying, "Oh, like to pick up evidence. [Laughing] Yeah, they use tweezers like to pick up the evidence, and like put it in a container or something like that" (G5PreD1:266). These responses indicate that students have reasonable ideas about the work that scientists do and the ways they might collect evidence. It also indicates, however, that students were unclear about what actually constitutes evidence in science.

Finally, several responses provided examples of materials used in a domain rather than evidence, in particular for science class. These answers reflected materials that were used in classroom activities, but not things that could conceivably be used as evidence. For example, "mentos" and "soda" likely came from students' knowledge of the well-known activity in which Mentos candy are put into a bottle of soda to cause a large eruption. Similarly, "baking soda" and "vinegar" probably come from their knowledge of the classic model volcano activity. Although these are appropriate to the domain of science class and could conceivably be used in an activity which involved data collection, the materials themselves would not be used as evidence in an argument.

Another possible explanation for their tendency to name tools and materials rather than evidence is the wording of the question and follow-up prompts. On the worksheet, question was worded as "What do you think scientists use as evidence? Make an idea web of different kinds of evidence scientists use." I used similar language in my follow up questions, asking things like

“What do they use as evidence?” and “What else do they use?” The word “use” may have thrown the students off and caused them to think about tools and materials rather than evidence. Amira, for instance, said that “They [scientists] use little stethoscope things” (G5PreD1:241). If she was attending to the first part of the prompt, “What do scientists use...” but not the second part “… as evidence” this is an entirely appropriate response and may account for the students’ focus on tools used to collect evidence rather than evidence itself.

Fuller Understanding of Science in General

While many of their responses about science and science class made the same mistakes discussed above (i.e., confusing subjects of study, tools and materials for evidence), these responses did still the reinforce the finding that the students generally have a more robust understanding of science compared to their understanding of history. First, there were considerably more responses for what scientists use as evidence than responses for the other areas. The sheer number of responses to the question about what is used as evidence in science demonstrate that the students had more robust beliefs about science, and evidence in science, than in the other areas. The students gave more than double the responses for what is used as evidence in science (65) compared to science class (32), history (32), and social studies class (24). Because they had more ideas about science than the other areas, it suggests that their conception of science was more developed than those other areas.

Second, their responses in science and science class were more likely to be appropriate to the discipline, than their responses for the other areas; 85% and 84% of their responses were appropriate to science and science class, respectively (Table 19). Responses like “germs,” “water life,” and “human body” for science and “feather,” “mass,” and “chemicals/borax” for science class were relevant to their domains and indicated the students understood what those disciplines entailed. On the other hand, only 44% of their responses for history and 63% of their responses for social studies class were appropriate to those domains. Many of these answers, such as “dead animals” and “tracker” for history and “vinegar” and “DNA” for social studies, were more appropriate to science. The greater number of examples of evidence, subjects of study, tools, and materials relevant to science and science class reveals the students had a stronger understanding of what these domains entailed when compared to history and social studies class.

Weak Understanding of Scientific Evidence. Although the students demonstrated they had a stronger understanding of science and science class in general, their knowledge of evidence, specifically, in science and science class was still weak. In other words, their ability to list items relevant to science and science class did not translate into more canonical examples of evidence. In fact as Table 19 shows, in terms of the proportion of their total responses, there were actually fewer canonical examples of evidence in science (20% of responses for science) than there were for social studies class (33% of responses for social studies class). Instead, as discussed above, the students were much more likely to provide a subject of study, tool, or material related to science and science class, rather than a piece of evidence.

Looking at the canonical examples of evidence in science (Table 20) also suggests the students had a very narrow understanding of evidence in science. Nearly all of these responses are connected to criminal forensics. This is unsurprising as it is likely where most students are accustomed to hearing the word “evidence” used. In fact, Daniel and Anna nearly said as much during the investigation and interview. While completing the pre-survey Daniel told his group mates, that “I seen some shows where they did that [use gloves]” (G5PreD1:274). During the interview Anna suggested that scientists might use a “pin board with string:”

> Anna: They might use a pin board with string, and they might put the evidence, and they might put when it was at, and then, and—or where did—where they were once that was written and things like that, to find evidence.
>
> Marino: Where did you get that idea about the pin board and the string?
>
> Anna: 'Cause I remember in a bunch of movies, there's pin boards or—and then there's strings that they put, and then I was like, "What if they can use that, or maybe they would." (IntPost:178-180)

It seems, then, that despite having a better understanding of scientific practice in general, the students still had limited and narrow knowledge about evidence, specifically, in science.

Good Examples of Evidence and its Use

Even with the clear difficulties the students had in naming examples of evidence, there were also indications during the small group work that some students had some knowledge of evidence, and how it might be used in the disciplines. While these discussions were not representative of the students' ideas about evidence, they do suggest that some students did have exposure to and knowledge of disciplinarily accurate representations of evidence and its use.

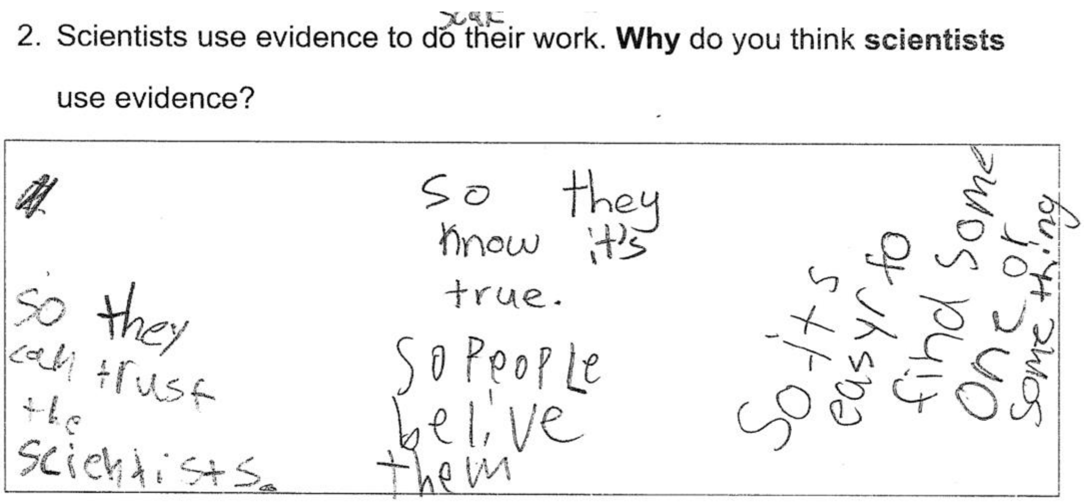
In science, there were two especially positive examples of students' beliefs about evidence. First, Jason from Group 5 asked, "What kinda data do they collect?" (G5PreD1:3) when discussing how scientists do their jobs. This reference to "data" was not prompted by me or the classroom teacher and shows his knowledge of both the vocabulary and its importance to science. The other students in the group picked up on this idea and followed up with statements like, "They collect data about their machines and stuff, or about stuff they use," and "They do their job by data" (G5PreD1:31; 79).

At another point, some students were not only able to identify specific examples of evidence in science, but also how it would be used. Group 1 recorded “bones” and “fossils” in the written pretest as something that scientists use as evidence. This was coded as an example of evidence that was appropriate to the discipline and canonical. The audio recording, however, revealed that not only did one of the students, Eric, know that bones were appropriate examples of scientific evidence, he had a specific purpose in mind. In it, he said: “I have an idea. How to get evidence of like big animals that eat other smaller animals—they like—their skeleton of the dead animals— in the mouth— and eat another skeleton” (G1PreD1:311-315). Here, Eric did not just demonstrate knowledge of what can be used as evidence, but how that evidence can be used to find out something, in particular the predator/prey relationship of animals from the past.

In history and social studies, there was one clear instance when students provided more specific examples of what could be used as evidence. During Group 3’s discussion of what is used as evidence in social studies class Brandon said, “—evidence of social studies is history. You can use history for evidence— Like the Constitution” (G3PreD2:120-122). By naming a specific document this response indicates a stronger sense of what constitutes historical evidence than simply “writing” or “books.” This answer suggests that Brandon is aware that historians use important documents in their work, and not just books.

Why is Evidence Used?

To characterize students’ beliefs about the reasons for using evidence in the domains of science, science class, history, and social studies class I analyzed their responses to the questions asking, “Why do scientists/historians use evidence?” and “Why do you think you use evidence in science/social studies class?” Figure 13 presents an example of one of the questions and one group’s responses.


2. Scientists use evidence to do their work. **Why** do you think **scientists** use evidence?

So they can trust the scientists.

So they know it's true. So People belive them

So it's easy to find some one or some thing

Figure 13: Beliefs Pre-Survey — Why is evidence used?

Although the students had a mostly weak understanding of *what* is used as evidence, they demonstrated a fairly good idea of *why* evidence is used in these domains. They provided three main reasons for using evidence: for discovery, for proof, and for research / learning. (There were also a few other responses that did not fit neatly into a category, such as “they team up” or “we use it when we solve word problems”.) Table 22 presents the number of responses in each of these categories. As the table shows, the greatest number of responses indicated that the students felt that evidence played a role in discovery, in particular for science where 46% of their answers were coded as “discovery.” This included answers like “to discover a new species,” “to figure out stuff,” “to know about other animals/fossils,” and “to explore new experiments.”

The students also seemed to understand that evidence played an important role in proving or supporting one’s ideas. Of their total responses, 28% were coded as “proof” and included answers such as “for proof,” “so they can prove stuff to nerds,” “so people can know it’s true,” and “so they can trust the scientists.” These responses demonstrate that many students understood two things about evidence. First, that evidence is necessary to support the truth of a

Table 22: Why is Evidence Used?

Discipline	No. of Responses	Discovery	Proof	Research/ Learning	Other
Science	13	6 (46%)	5 (38%)	1 (8%)	1 (8%)
Science Class	8	2 (25%)	2 (25%)	4 (50%)	0 (0%)
History	9	4 (44%)	1 (11%)	0 (0%)	4 (44%)
Social Studies Class	9	2 (22%)	3 (33%)	2 (22%)	2 (22%)
TOTAL	39	14 (36%)	11 (28%)	7 (18%)	7 (18%)

statement or claim. Second, that evidence, or “proof,” plays a social role and that arguments have an audience that needs convincing. A number of these responses included a second party at whom the evidence is aimed. For instance, in the response “so they can trust the scientists,” “they” refers to an audience whom the scientists are providing evidence to with the goal of making that audience believe the scientists. After suggesting this answer to his group, Solomon said, “No, some people don’t trust them [scientists]” (G1PreD1:506), explicitly acknowledging the scientists’ audience and why they might need to be given evidence. Responses such as these indicate that the students had some idea of the dialogic nature of arguments and the role of evidence in that exchange.

The final category of reasons for using evidence, “research/learning,” suggest that the students saw differences between the disciplines of science and history and the school subjects of science and social studies. The majority of the responses coded as “research/learning” for the question on why evidence is used were provided for the school subject of science class. These responses, “to find out about it,” “so if you can see if you got something right or wrong,” “to get smarter,” and “to know and learn about it,” suggest that students saw a difference between the discipline of science and science as an activity in school. As opposed to scientists, who according to the students primarily use evidence for proof and discovery, science students’ use of evidence seems to be tied to learning content, ostensibly the primary purpose of school. The idea that students mostly saw evidence as something to be learned, rather than used, in science class is reinforced by their tendency to provide subjects of study as examples of evidence in science. In school, one is more likely to learn *about* amoebae rather than *use* them as proof of a claim.

In addition to suggesting the students had different ideas about the use of evidence in the disciplines versus school, their responses also suggested (again) that the students had a stronger understanding of science than of the other domains. As with the other questions, they provided a greater number of answers for science than for science class, history, or social studies class. Second, a greater proportion of these responses regarding science as a discipline cited discovery and proof as reasons that evidence is used. Compared to “research/learning” and the responses coded as “other,” these two uses of evidence are more appropriate to the disciplines. Finally, some of their responses for history and social studies class were not appropriate to those disciplines, echoing the confusion students had when listing examples of evidence or describing how historians do their jobs. Responses such as “to know about other animals/fossils,” “to explore new experiments,” and “we use it when we solve word problems” fall outside of the discipline of history and the school subject of social studies.

Conclusion

This chapter described students’ initial beliefs about evidence, the disciplines of science and history, and the school subjects of science in history. Figure 14 presents the landscape of findings related to students’ beliefs. The figure presents “successes” and “challenges” on the y-axis to indicate to what degree students struggled with a practice or had a misconception about an idea. Things that were especially challenging are represented as being closer to the extremes of the axes, whereas less challenging elements are closer to the intersection of the x- and y-axes. The disciplines of science and history are presented along the x-axis. Labels on the extreme ends of the x-axis are more specific to one discipline or the other, and those at the center are equally relevant to the two investigations. Presenting the findings in this way makes the differences and similarities between how students engaged in the two investigations evident, as well as making the students' relative strengths and weaknesses visible. This figure will be used throughout the findings chapters to present the landscape of findings related to each element of the study.

Figure 14: Landscape of Findings Related to Students' Initial Ideas about Evidence.

As Figure 14 shows, there were a number of similarities in the students' ideas about evidence in the disciplines of science and history. In general, they had a sense of what general academic practice looked like, but a weak understanding of the disciplinary practices specific to science and history. They also seemed to have a fuller, although not always articulate, understanding of science, and generally provided a greater number of responses for questions about science that demonstrated more background knowledge of the discipline. The students had a relatively poor understanding of what constituted evidence in both fields, although they were able to name slightly more appropriate examples in science than in history. Despite this weak knowledge of evidence, the students seemed to have a good understanding of the *uses* of evidence in both the disciplines and the school subjects. Finally, there were indications that students had ideas about evidence and its uses that were specific to the school subjects and different than those about the disciplines of science and history. This is represented in the figure as a strength because the practice of the disciplines in the field is typically quite different from their practice in schools. In the next chapter, I build on these results by analyzing how students used evidence in practice and turn to the results related to the stage of the investigation in which they read and evaluated sources of evidence.

CHAPTER 5

Students' Reading and Evaluation of Evidence

In this chapter I present findings related to the ways that students read and evaluated evidence in preparation for answering an investigation question and constructing an argument. This analysis is based on the portion of the science and history investigations in which students worked in small groups to read sources, six per investigation, and record their ideas about what the sources told them about the investigation question. Figure 15 provides an example of one of the source evaluation worksheets from the history task.

Based on the students' written responses and discussions during small group work, I make 6 major assertions:

A) The students were more likely to focus on basic comprehension of the sources in the history task compared to the science task.

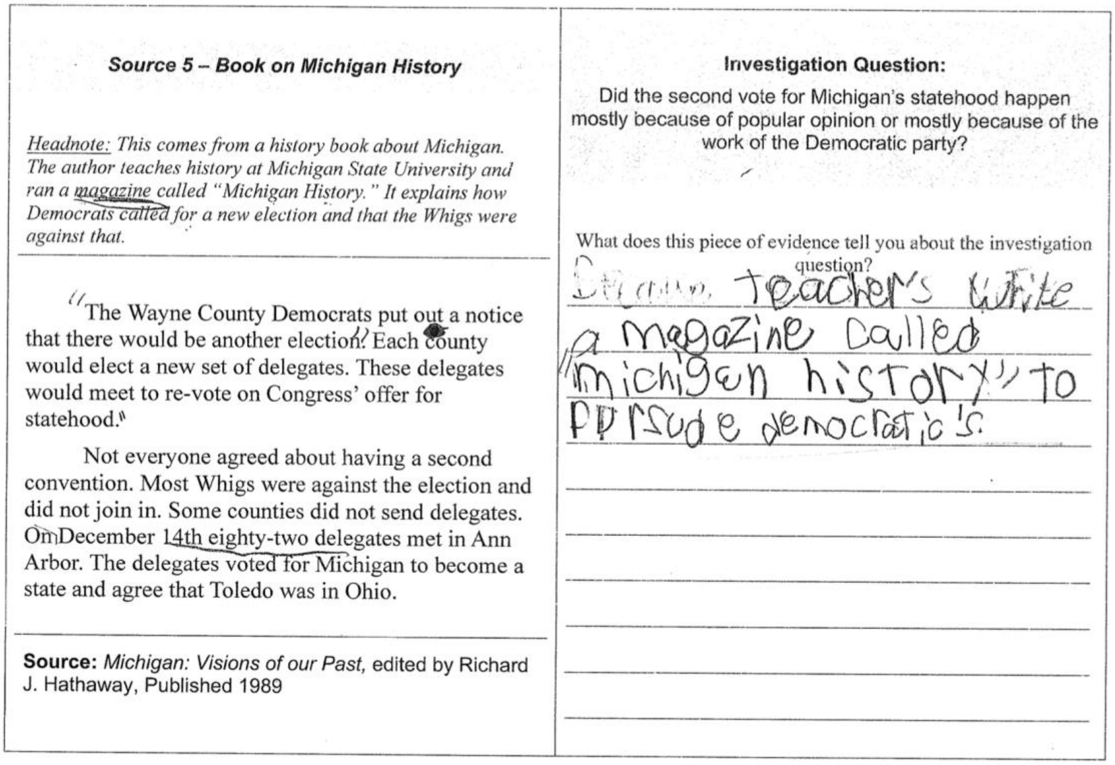
B) The students were more likely to make inferences or draw conclusions based on the sources in the science task compared to the history task.

C) The students had stronger initial ideas about the answer to the investigation question in the science task than the history task.

D) The students' existing beliefs about the answers to the investigation questions resulted in readings of the sources that supported those ideas.

E) There was relatively little genuine discussion during the small group work.

F) The students appeared to have a better understanding of the science task and content when compared to the history task.



Source 5 – Book on Michigan History

Headnote: *This comes from a history book about Michigan. The author teaches history at Michigan State University and ran a magazine called "Michigan History." It explains how Democrats called for a new election and that the Whigs were against that.*

"The Wayne County Democrats put out a notice that there would be another election. Each county would elect a new set of delegates. These delegates would meet to re-vote on Congress' offer for statehood."

Not everyone agreed about having a second convention. Most Whigs were against the election and did not join in. Some counties did not send delegates. On December 14th eighty-two delegates met in Ann Arbor. The delegates voted for Michigan to become a state and agree that Toledo was in Ohio.

Source: *Michigan: Visions of our Past,* edited by Richard J. Hathaway, Published 1989

Investigation Question:

Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?

What does this piece of evidence tell you about the investigation question?

teacher's write a magazine called "michigen history" to prsude democratic's.

Figure 15: Group 5 — History Source 5.

Using the Text

I begin by discussing the students' overall approach to reading the texts for the source evaluation portion of the science and history investigations. In general, the students did much more active reading and referencing of the texts during the history investigation than during the science investigation. As Table 23 shows, the groups read, referenced, or made a bid to read or reference the text 47 times during the history investigation and only 11 times during the science investigation.

Groups 1 and 3, in particular, seemed to place a good deal of emphasis on reading the text during the history investigation. Group 3 referenced or reread portions of the sources 14 times. This included three instances when they began reading the text before I began reading to the

Table 23: Frequency of Reading and Referring to Evidence.[6]

Subject	Reading	Reference/ Quotation	Bid	Read to Self	TOTAL
History	27	5	3	12	47
Group 1	9	2	1	2	14
Group 3	13	1	0	1	15
Group 5	5	2	2	9	18
Science	5	3	1	2	11
Group 1	1	0	1	2	4
Group 3	2	0	0	0	2
Group 5	2	3	0	0	5

class. This suggests they understood the important role of the texts in this part of the task. Pre-reading the texts, in particular, show that the students in Group 3 had a strong sense that their work and ideas needed to be rooted in the source.

In one case Group 3 explicitly discussed rereading the text and implied that it should be the source of their ideas. After I read Source 5 to the class, Group 3 was having a hard time coming up with ideas. Brandon told the group, "Come on, think of something" (G3HD2:648). The group then decided to read the text in what seemed like an attempt to jumpstart the discussion with Taylor asking, "Yeah, what does that say?" and Brandon conceding, "Fine, I guess I'll read it over again" (G3HD2:663, 666). As he was rereading the source, the other students drifted off and began having a side conversation. Brandon became frustrated and told them, "I'm not going to read it if you guys keep on being rude" (G3HD2:717). While his main frustration seems to be

6 These data come from the analysis of the videos of students' small group work. As noted in Chapter 3, Group 2 was excluded because it was comprised of ELLs and students reading significantly below grade level who got a good deal of support from the classroom teacher. Group 4 is not included in the video data because some students did not consent to being recorded.

that they're not paying attention to the work he's doing, it also indicates that he knows reading the text is important to the task and that if they do not listen to the text they cannot complete it.

At another point, some of the students in Group 3 directly tied the texts to their role as evidence in an argument. While examining Source 2, Aisha quoted directly from the source in suggesting something to write down on the worksheet. In looking at Source 3 Aisha told Yasmin, "No, you have to tell me a piece of evidence" and Yasmin responded, "I evidenced right here looking at these because they're the same [referring to two numbers on the table in the source]" (G3HD2:281-281). Aisha demonstrated that she understood that the purpose of the task was to find evidence to support their answer to the investigation question. Yasmin's response, in turn, suggests that she shared this understanding and had been taking that approach throughout the task. Furthermore, this exchange between Yasmin and Aisha, seemed to set a precedent for Brandon. He tried to add to the discussion but had to tell Yasmin, "Move your finger. I need to look at something. I can't read what it says" (G3HD2:285), showing that he also saw the importance of rooting the discussion in the texts.

In contrast, the students rarely made explicit reference to the sources during the science task, and only read or reread the texts 5 times total across three groups. This may be related to their beliefs about the science investigation in general. As I discuss below and in the following chapter, the students had stronger initial ideas about the answer to the science investigation in comparison to the history investigation, and many seemed to decide on an answer before even looking at the sources. If the students already had ideas about the investigation question, they may not have placed as much importance on the texts as the source of information necessary to answer the question.

While their beliefs about the answer to the science investigation likely explains their lack of attention to the texts in the science task, Group 5's failure to reference the texts is notable for two reasons. First, Jason, who took a strong leadership role in his group, was responsible for 6 of the 9 instances of students in Group 5 reading the text to themselves during the history task. He was fairly conscientious about ensuring that he responded directly to the sources in this part of the task. In the science task, however, he only reread a source one time. The difference in his behavior during the science task indicates a possible shift in attitude or beliefs about the task. Second, during the science task Daniel expressed a sentiment similar to Yasmin's above. While suggesting a response, he told the group, "Lemme show you. Lemme show you. Capisce? This is this proof right here [indicating text]" and later, "We need to read—fine. This is the proof, and then you have to read this to confirm the proof and then this to—" (G5SD1:244; 254). While he clearly understands that "proof" needs to come directly from the evidence, this did not translate into reading or referencing the text in the science task very often.

Purposes of Reading

When reading and interpreting the texts, the students took two main approaches to reading the text. Their reading strategies either appeared to focus on a literal comprehension of the text or made inferential/evaluative readings of the text. By *literal* reading strategy, I mean a reading that focuses on basic understanding of what the text is saying and does not look beyond the scope of the text in the reading. Literal readings were evident in responses that provided a summary or key details from the source. An *inferential/evaluative* reading strategy, on the other hand, means reading the text with an eye towards the investigation question or making inferences that go beyond the literal meaning of the text. Inferential readings of the sources resulted in responses that went beyond restating what the text said and made connections between the text and the investigation question or drew conclusions not explicitly stated in the text.

In general, the students took different approaches when reading the evidence in the two tasks. During the history task they were more likely to make literal readings than they were to use an inferential strategy. As Table 24 shows, the majority (60% total) of their written responses provided a detail or summary of the text (i.e., literal reading) and only 38% provided an inferential/evaluative reading of the text. In the science task however, just over half of the written responses indicated an inferential reading of the text.

Table 24: Details, Summaries, and Inferences in Written Responses.

Discipline	Literal Reading		Inference / Evaluation	Unclear	TOTAL
	Detail	Summary			
History	21 (34%)	16 (26%)	23 (38%)	1 (2%)	61 (100%)
Science	13 (20%)	36 (23%)	15 (55%)	1 (2%)	65 (100%)
Total	34 (27%)	52 (24%)	38 (46%)	2 (2%)	126 (100%)

Reading for Literal Meaning in History

The focus on literal comprehension during the history task was evident in the group discussions as well as their written responses on the source evaluation worksheets. Group 1's work provides examples of written work and discussions that were centered on literal interpretations of the sources. First, their answers on the source evaluation sheet mostly provided details or summaries of the sources. See Table 25 for their responses.

Table 25: Group 1 – Purpose of History Source Evaluation Responses.

Source Number	Response Number	Response	Purpose of Response
1	1	The whigs wanted to keep Toledo.	Summary
1	2	How many people wanted to keep Toledo.	Summary
2	3	71 people wanted to hear and be where they are at.	Detail
3	4	That more people wanted to become a state then not give up Toledo	Summary
3	5	That they work so hard to be a state.	Inference / Evaluation
4	6	That it is Democrats wanted to make Michigan a stat	Inference / Evaluation
5	7	Whigs were against the re-vote.	Summary
5	8	It was on December 14th eighty-two delegats wer ther	Detail
6	9	That the leader was a democrat too	Detail
6	10	that its elected by the people of the state	Detail

Second, much of their discussion revolved around deciding the literal meaning of the text. Such a discussion occurred when they were deciding what to write for Source 3 and recorded response 4 in the table above. After reading Source 3, Eric stated “I know what I’m gonna write” (G1HistD2:226) and began writing “That more people wanted to become a state then not give up Toledo” without input from the rest of the group. Eventually, the other students began reading aloud what he was writing and gave suggestions about spelling. As this was happening Angela interjected and said:

Angela: “Toledo. No, not give up.”

Eric: “To give up.”

Solomon: “Yes.”

Eric: “’Cause they have to give it up.”

Anna: “They didn’t want to—”

Angela: "It says not give up. Oh, never mind." (G1HistD2:250-256)

Here, they were discussing both the correct interpretation of the table in the source and how to correctly apply the facts they learned during the introduction to the text (specifically, the Whigs' stance). When Angela said "No, not give up" she is challenging how Eric is recording the facts of the story. She seems to be confused by Eric's phrasing and believed that he had gotten the conditions of Michigan's statehood incorrect. While the answer that Eric recorded could be applied to the investigation question and interpreted to mean that the source supports the claim that popular opinion was behind the second vote, this discussion was about the literal interpretation of the text. In other words, Angela seemed more interested in how well Eric comprehended the text than whether or not he connected it to the investigation question appropriately. At the same time, however, while they were not making inferences about the investigation question based on the text, this reading of the text did focus on details that were important to understanding and answering the investigation question.

This approach focused on literal meaning was also evident in some of the group members' strategies for recording their responses. When preparing to write response 9 in the table above, Caitlin suggested, "Can't we just write the whole entire thing? [indicating a portion of the source]" Solomon agreed and told her to "Write this whole thing." She then seemed to change her mind saying, "No, that's just copying it," but Angela and Anna tried to convince her that it was appropriate because, "It's a piece of evidence" (G1HistD2:915-925). By considering just copying the text, the students showed that they were focused on finding what they considered to be important details, and not necessarily drawing inferences from the text.

The other groups also took a largely literal comprehension focused approach to reading the sources in history. In one case, Group 5 even steered their initial inferential reading of the text away from drawing conclusions and toward demonstrating their basic understanding of the source. At the beginning of the exchange, it seemed as if the students were going to have a discussion about the relevance of the source to the investigation question. Jason began by saying “Because of the Whig party…” and Zahra picked up his idea and repeated “Because of the Whig party” a number of times (G5HistD2:288-296). Because Zahra was not echoing Jason’s words to write them down, the repetition indicates that both students were thinking about the same idea. Specifically, they were attempting to make inferences from and evaluate the source by assigning causality to the results of the second convention and thus tie the source to the question of who was mostly responsible for the vote.

Quickly, however, the students moved from an inferential discussion, to one focused on literal meaning, and then finally a single student dictating ideas. As students were looking back at the source after the above exchange, I came to the group and asked them, “What are you thinking?” Continuing his thinking above, Jason responded, “I’m thinking about that Whig Party that had to do something, maybe, about it, or it’s because of the popular opinion, because everyone thinks—” Amira, however, then suggested, “It tells us about the number of people.” Jason then picked up on this train of thought and offered, “It tells us about how much countries attended and— Counties attended to be a state.” With this statement, the discussion effectively ended as Amira said, “I’m writing down everything he says” and the other students begin an off-task discussion as Jason dictated the response (G5HistD2:305-318).

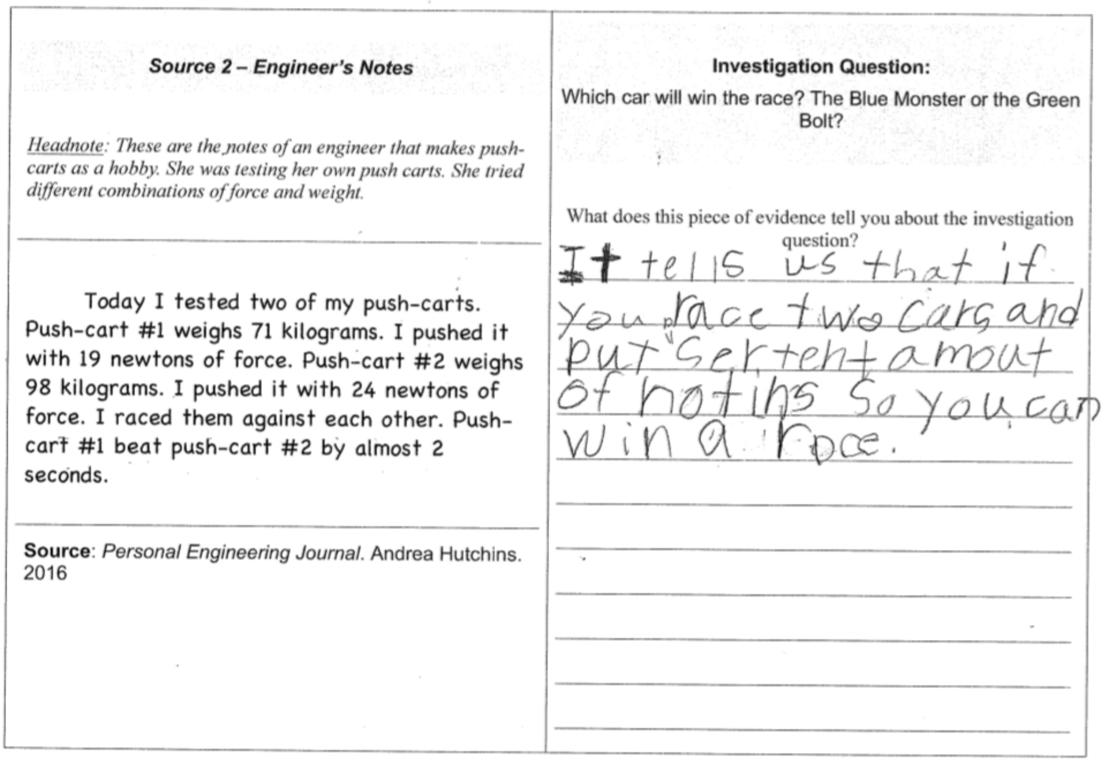
Reading for Inference and Evaluation in Science

In comparison to the history task, the students read with the goal of inference and evaluation more frequently in the science investigation. Group 5, in particular, had the greatest number of inferential responses of any group (7 versus 3 and 4 for the other two groups). As Table 26 shows, the majority of their written responses to the sources in the science task were inferential. Two of these responses, numbers 1 and 7, directly answer the investigation question based on the students' reading of the sources. The remaining inferential responses did not directly answer the investigation question but did provide general principles that could be useful to support an answer. "The lighter the faster," "the People Who has the lighter cart will WIN!" and something that is "heavyer so it is going to be slower" all result from inferences based on the sources and could be used to support the conclusion that the Green Bolt, the lighter of the two carts, will win.

Table 26: Group 5 – Purpose of Science Source Evaluation Responses.

Source Number	Response Number	Response	Purpose of Response
1	1	The green bolt [will win], because the ligher the faster even though there not as strong but it's light. so They can push it a lot faster.	Inference / Evaluation
2a	2	It tells us that Cart #2 is lighter and the lighter the faster.	Inference / Evaluation
2a	3	And cart # 1 is heavyer so it is going to be slower.	Inference / Evaluation
2b	4	It tells us that the person who tested #1 and #2 was trying to test them for the race.	Detail
3	5	the person who has the ligher cart [will win]	Inference / Evaluation
3	6	it's light but the air pressure goes on the cart and it slows it down.	Inference / Evaluation
4	7	That Asia's [the Green Bolt] is going to win.	Inference / Evaluation
4	8	Also Edgar [the Blue Monster team] is worried she [Asia] is going to win.	Summary
5	9	Their testing all of their carts so they can see what cart is the best for the race.	Summary
6	10	The People that have more force and heavier cart will LOSE! But the People Who has the lighter cart will WIN!	Inference / Evaluation

One notable example of inferential/evaluative reading strategies during students' discussions in the science investigation comes from Group 3's reading of Source 2b, Figure 16. In this exchange, the students actually transitioned from attempting to summarize the source into making an inference based on the source.



Source 2 – Engineer's Notes

Headnote: These are the notes of an engineer that makes push-carts as a hobby. She was testing her own push carts. She tried different combinations of force and weight.

Today I tested two of my push-carts. Push-cart #1 weighs 71 kilograms. I pushed it with 19 newtons of force. Push-cart #2 weighs 98 kilograms. I pushed it with 24 newtons of force. I raced them against each other. Push-cart #1 beat push-cart #2 by almost 2 seconds.

Source: *Personal Engineering Journal*. Andrea Hutchins. 2016

Investigation Question:

Which car will win the race? The Blue Monster or the Green Bolt?

What does this piece of evidence tell you about the investigation question?

It tells us that if you race two cars and put serten amout of notins so you can win a race.

Figure 16: Group 3 — Science Source 2b.

Marino: What does this tell you about the investigation question?

…

Taylor: That it tells us that the girl. It tested two of her push-carts, and then the first one won.

Marino: Okay, so how does that help you answer your investigation question?

Brandon: It tells you which one is faster. And the length of time.

Taylor: It helps with our investigation question because it tells us—

Brandon: Which one is faster.

Taylor: Yeah, it tells us—

Brandon: Which one of the cars is faster.

Taylor: - which one—

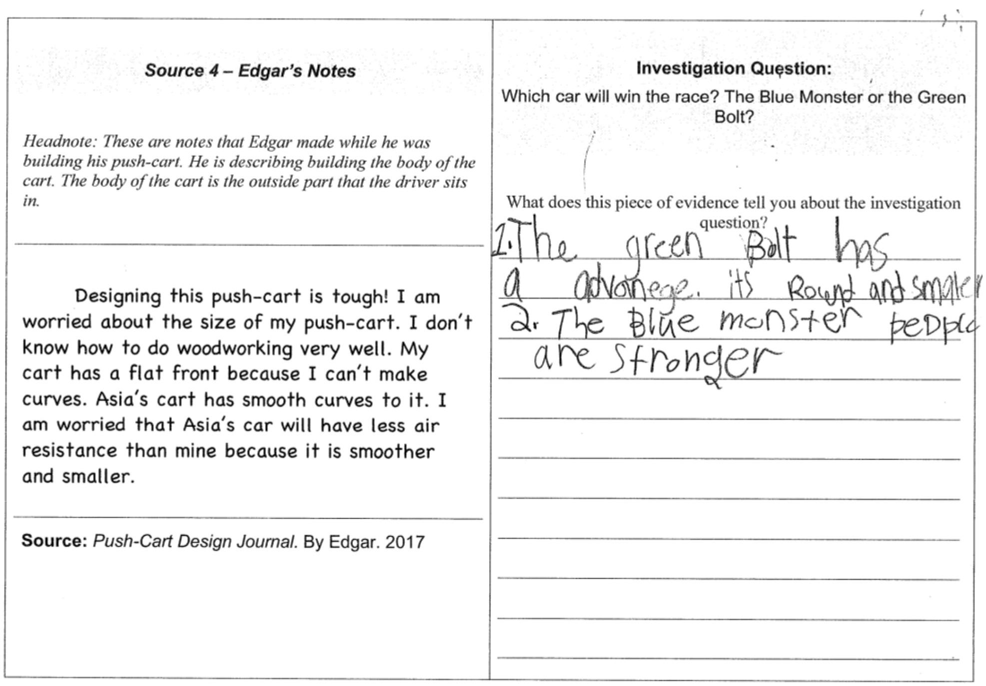
Brandon: So, then we can know—

Taylor: - which One is faster, but it also tells us—it also tells us how much newtons you can put in a car, and then when you race them against each other, that One might win. (G3SD2:229-244)

Taylor and Brandon both began by identifying important details from the text and summarizing it as it tells "which one of the cars is faster." It is important to note, however, that the source is reporting the results of a trial using push-carts that are *different* than the ones in the investigation question. Although one is heavier and stronger and the other is lighter and weaker, push-carts #1 and #2 from Source 2b are not the Blue Monster or Green Bolt. In Taylor and Brandon's summary, then, was the potential that the students were only providing a summary and misunderstanding the text. Taylor's addition, however, made a generalization based on the information in the text, and showed that she understood the source was not about the Green Bolt and Blue Monster. Admittedly, Taylor's statement did not directly connect the source to the investigation question and was thus of limited use in her argument. She still, however, was attempting to go beyond the face value of the text and draw a generalizable conclusion based on its information. Although this exchange was instigated by a prompt asking the students to connect the evidence to the investigation question, this should not diminish the significance of Taylor's conclusion and transition away from summary because my prompts did not necessarily lead to inferential responses from the students.

Answering the Investigation Question

Another important element of how the students read and evaluated the evidence during the investigations is the extent to which they addressed the investigation question in their responses and discussion. Each worksheet included the prompt, "What does this piece of evidence tell you about the investigation question?" Ideal responses, then, would tend to be more inferential, reference specific elements of the investigation question, and might even suggest an answer to the investigation question. Group 4 provided such a response in their reading of Source 4, Figure 17. They said, "1. The green Bolt has a advanege. Its Round and smaler." This directly connects to the investigation question by naming the Green Bolt, indicating a possible answer to the question, "the green Bolt has the advantage," and including a detail from the source, the Green Bolt is "round and smaller" connected to the inference.



Source 4 – Edgar's Notes

Headnote: These are notes that Edgar made while he was building his push-cart. He is describing building the body of the cart. The body of the cart is the outside part that the driver sits in.

Designing this push-cart is tough! I am worried about the size of my push-cart. I don't know how to do woodworking very well. My cart has a flat front because I can't make curves. Asia's cart has smooth curves to it. I am worried that Asia's car will have less air resistance than mine because it is smoother and smaller.

Source: *Push-Cart Design Journal.* By Edgar. 2017

Investigation Question:
Which car will win the race? The Blue Monster or the Green Bolt?

What does this piece of evidence tell you about the investigation question?

1. The green Bolt has a advanege. its Round and smaler
2. The Blue monster peple are Stronger

Figure 17: Group 4 — Science Source 4.

This response, however, was atypical of the answers students provided on their source evaluation worksheets and their discussions. On the source evaluation worksheets, only 16% of their responses suggested an answer to the investigation question based on the sources. In addition, there were relatively few instances in the discussions in which students answered the investigation question. Instead, when they discussed the texts or their responses the students often focused on identifying details from the text that should be included in their answers.

Stronger Ideas about the Science Task

Although the students answered the investigation question relatively infrequently in both investigations, there do appear to have been differences in the strengths of their beliefs between the two tasks. This is indicated by the difference in the number of times they answered the investigation question in each content area. In history, the students only answered the investigation question in 10% of their responses. In contrast, they answered the question in 21% of their responses in the science task. This suggests that the students had stronger beliefs about the answer to the investigation in science.

Additionally, some students appeared to have had ideas about the answer to the science investigation prior to reading any of the sources. Group 5, for instance, started discussing which push-cart would win the race as soon as they got to their table and before reading any of the sources. Jason initiated the conversation with, “I already know the answer,” then, “Which one? The lighter one or the heavier one?” (G5SD1:5; 26). In response, Daniel suggested “The lighter one, but the heavier one has more—” (G5SD1:27) and Zahra and Jason both agreed, with Daniel concluding, “I think it’s the lighter one, I’ll just go with that” (G5SD1:41). Similarly, before reading the sources, Caitlin in Group 1 asked her partners, “I know that we don’t have the piece of evidence yet, but who votes?” and Eric emphatically responded, “Green” (G1SD1:21; 26).

As with Group 5 above, most of the students thought that the Green Bolt, the lighter of the two push-carts, would win the race. They seemed convinced that the lightness of the cart mattered more than the fact that the Blue Monster, the heavier push-cart, was pushed with greater strength. Their written responses often included this justification, such as: “also the cars that way [weigh] light go faster and that the small one could win,” [G1] “The green bolt, because the ligher the faster even though there not as strong but it’s light. so They can push it a lot faster,” [G5] and “The green Bolt is going to win becus it has les fors [less force] and it it liter [lighter] so you can push it frther.” [G3]

Another indication that students may have held stronger beliefs about the science investigation was the number of opportunities they took to address possible answers to the investigation during their discussions. Table 27 presents the number of source discussions in which the answer to the investigation question was discussed (i.e., Group 1 discussed the answer to the investigation question for 3 of the sources from the history task). Groups 1 and 5 included the answer to the investigation more frequently in their discussions of the science sources than the history sources.

Table 27: Source Discussions Answering the Investigation Question.

Group	History	Science	Both
1	3	6	9
3	4	4	8
5	1	5	6
Total	8	15	23

The Role of Teacher Prompts

Teacher prompting seems to be one factor that played a role in how much the students discussed answers to the investigation question during their small group work. As the students worked in their small groups to evaluate the sources, I circulated the room monitoring their work, encouraging them to add more detail, and providing prompts when they seemed stuck.

Typically, my prompt was, “What does this tell you about the investigation question?” which echoed the prompt on the source evaluation worksheet. I repeated this prompt after completing reading each source aloud to the group. Although there were times that the students spontaneously suggested answers to the investigation question or interpretations of the text that would directly support a claim, there were also a number of instances when they needed my prompt to begin their discussions.

During the history task, for instance, each of Group 1’s three discussions involving the investigation question were the result of one of my prompts to the group. While reading Source 3, for instance, I asked their group “Okay, what else does this tell you about the answer to the investigation question?” Two students responded that they did not know, so I followed up with, “Well, remember, the investigation question is, did the second vote happen mostly because of popular opinion or mostly Democrats?” This added guidance seemed to help because Eric immediately answered, “Democratic” and Solomon said, “That it was because - they worked as hard - to make the Democratic party” (G1HD2:288-303). Based on their responses it was clear the students had ideas about the source and the answer to the investigation question. It appears, however, that they needed the guidance to understand that making that link was what was expected of them in the task.

Reading to Support Existing Ideas

The students’ answers to the investigation question also seemed to have an effect on their reading of the sources. When the students did express beliefs about the answers to the investigation question, they often read the sources in ways that were favorable to their ideas and supported their existing claims. This occurred with several students in both the science and history investigations. Group 3, for instance, thought that the Democrats were largely responsible for the second convention on Michigan's statehood, not popular opinion. After coming to that conclusion, they ignored or did not look for evidence that would support the other possible answers to the question. Their inferential written responses supported the claim that Democrats were responsible for the second vote. In addition, whenever they discussed possible interpretations of the texts, it was always to support the idea that the Democrats were responsible for the second vote. Some students, for instance, even used Source 2, which was the strongest source supporting the claim that popular opinion was behind the vote, to support the claim that the Democrats were responsible. In response to one of my prompts to the group Brandon said, "It was more popular at the second–" and Aisha interjected, "Yeah, it was mostly what the Democratics say" (G3HD2:68-69). Even though Brandon's response suggests that popular opinion was behind the vote, Aisha picks up his idea and interprets "people" to mean the Democrats. Their written response also indicates that they may have been ignoring evidence that countered their claim. They recorded "Many people wanted to change there votes" on the worksheet.

In another case, one of the students included her existing ideas about the motivations for the second convention into her reading of the source.

> Marino: What does this [Source 4] tell you about if it was popular opinion or if it was Democrats did?
>
> Taylor: I think it's what the Democrats did.
>
> Marino: You think it's what the Democrats did? What does it tell you?
>
> Brandon: Yeah, I'm sayin'—
>
> Aisha: It's me too. I think it's because the Democrats because the people of Michigan probably wanted to become a state and let Ohio have it [Toledo]. (G3HD2:447-451)

On one hand, Aisha's reading of Source 4 is appropriate because it did support the claim that the Democrats were largely responsible for the re-vote. The source, however, did not mention anything about Toledo and the dispute between Ohio and Michigan over the city. It actually described a meeting of Democrats that resolved to call for public meetings that would petition for a second vote on statehood. The conflict over Toledo had been mentioned in other sources and the introduction to the investigation. Aisha's response, then, appears to be unconnected to the source itself because she is making claims based on details not contained in the source. Instead, she seems to simply be repeating her existing ideas about the investigation and applying that to the source at hand.

In science, Group 5 had strong beliefs about the answer to the investigation question that seemed to shape their reading of the texts. First, their responses did not always directly relate to the text of the sources and were often just a statement of their answer to the question. In addition, when they did directly address the text they were likely to ignore evidence countering their claim and interpret it only in favorable ways. The clearest example comes from their reading of Source 2a, which was the source that had a typo and reported that a push cart a heavier, weaker cart went faster than a lighter, stronger cart.[7] As with many of the groups, Group 5 did not notice this error. Instead, they interpreted it as supporting their belief that the Green Bolt would win. On the worksheets, they recorded: "It tells us that Cart #2 is lighter and the lighter the faster" and "And cart #1 is heavyer so it is going to be slower." While Cart #2 was the lighter of the two, according to the source it lost the race. The students did not pay attention to this detail in the text, however, and instead repeated ideas they had expressed earlier.

[7] Students were given a revised version of the source to reevaluate on the following day, and when they created their arguments with the SenseMaker board, they only had access to their notes on the second, corrected source.

In another case, the students ignored their own thinking to come to a reading that supported the Green Bolt. Source 6 had details that would support both the Green Bolt and the Blue Monster as the likely winners of the race. When discussing source 6, Jason initially placed a lot of value on a cart's push force during this exchange:

> Jason: –the people that are pushing with the most Newtons will have a better chance of getting there. the car, with the weight, it can be slow. But they have–
>
> Daniel: But they can push it.
>
> Jason: The air will just–
>
> Daniel: It'll go slower, but–
>
> Jason: Yeah, it goes slower, but–
>
> Daniel: They can push it–
>
> Jason: They can push it with a lot more strength.
>
> Daniel: Yeah. (G5SD2:401-408)

Jason's ideas that with more force they "will have a better chance of getting there" and "they can push it with a lot more strength" seemed to support the Blue Monster (the heavier cart pushed with a greater force) as the winner of the race. He later clarified, however, that he still believed the heavier cart would lose and that "The people who are not as strong but have the lighter car are going to win" (G5SD2:433).

The idea that the Green Bolt was going to win because it weighed less came up a number of times while Group 5 read the evidence, including times when the cart's weight was irrelevant to the text. While reading Source 5 Daniel commented that, "it's [the Green Bolt] lighter so it can go faster" (G5SD2:187) even though the source was reporting on push force. When reading Source 4, which was primarily about aerodynamics, Jason noted "They're [the Green Bolt] only five newtons down on the force [than the Blue Monster]. … It's [the Green Bolt] gonna go faster than–" (G5SD2:94-96). As with Daniel, he seemed fixated on the idea that weight was the most important factor of the push-carts' speed, which appears to have influenced his reading of the texts.

Ambiguity of Responses

Although relatively few responses directly answered the investigation question, there were a number of responses that could arguably have been used to support an answer to the investigation question. In the history task, for instance, responses like "Whigs were against the re-vote" could be taken to mean that the vote did not reflect popular opinion, or "A lot of people wanted to vote again!" could support the claim that popular opinion was largely responsible for the second convention on Michigan's statehood. Unfortunately, however, it is difficult to determine the intent behind these responses. First, many of these ambiguous responses are details reported from the source. Because there was no clear inference or evaluation evident, it is possible that the students were only recording it because it seemed important to the text, not necessarily to the investigation question. Second, there were no indications in the written responses or discussions that this is how the students intended for responses such as these to be used. As I discuss later, there were relatively few exchanges in which students articulated their reasoning or justification for their responses. More typically, they simply suggested an answer which was then recorded.

The ambiguity of these responses means that the students may have made more connections between the sources and investigation question than the previous analysis suggests. In fact, two groups did use the Whigs' opposition to support their claim. Unfortunately, the only way to confirm that responses that provided a detail or summary were actually connected to a claim is if the response was in their final argument. This means that attempting to identify when those types of responses were connected to the investigation question was too difficult and too subjective a task. Because a handful of these responses did make it into the final arguments, however, this indicates that at least some of the time the students responded with specific details from the text they were considering how those details supported their claim.

Group Discussion and Collaboration

One hallmark of the groups' work during the source evaluation task was relative lack of meaningful discussion. While they read and evaluated sources the students rarely talked about the texts in a way that constituted a true discussion. This means that they did not typically respond directly to one another's ideas by agreeing or disagreeing or in ways that indicated they were engaging with their partners' ideas. They also did not usually include any kind of justification for their ideas when sharing what they thought with the group. Instead, the students were more likely to simply take turns stating their ideas, dictate what to write to another student, or narrate what they were writing, all without explanation. So even though they were involved in an argumentation-focused activity, finding evidence to support a claim to the investigation question, they were not engaging in argumentation amongst themselves.

Group 1's work on the science task was fairly representative of the groups' "discussions" during the source evaluation portion of the investigation. Most of their written responses were the result of a single student recording their own idea or a second student dictating what to write. Of their 18 written responses, 10 were the result of this kind of independent work without any input from the rest of the group. During these long stretches of independent writing, the rest of the group was either off task or waiting for their turn to write. There was a point, for instance, where Solomon was writing a great deal about Source 2b on his own and the other students in the group were having an off-task discussion. As this was happening, I came by their group and gave them a second copy of the source and told them, “While he’s writing, you three can get this one” and then came back and asked, “Ok, what did you all decide?” (G1SD2:759; 824). Even with the explicit direction to collaborate on the source, however, Angela recorded her own idea on the second sheet without any input from the group.

In the cases when there were meaningful discussions about the sources, they were almost all prompted by me. Of the 7 exchanges identified as discussions, 5 were clearly initiated by a question or prompt from me to the group[8]. After these prompts, the students were often able to engage in a more meaningful discussion about the text, such as in this exchange:

> Marino: What is this one telling you?
>
> Solomon: That if there’s enough light on the car—
>
> Anna: —that either way, if the cart is heavy or light, it could still go faster.
>
> Marino: What does that tell you about who you think is gonna win?
>
> Anna: It tells me that the heavy one—the light one could win.
>
> Marino: It tells you that the light one could win?
>
> Solomon: If it has enough light, it goes faster. (G1SD2:531-540)

Here, the students built upon each other’s ideas to come to a conclusion that the lighter of the two push-carts, the Green Bolt, would win the race. Even though their attention was directed at me and they did not explicitly acknowledge other’s ideas, they were still in agreement about the

[8] “Prompts” in this case are slightly different than the prompts described above. In that case, I was limiting my analysis to prompts that directly invited students to make connections between the source and investigation question, such as “What does this source tell you about the investigation question?” Here, “prompts” also includes open-ended prompts and questions that do not explicitly guide the students to make that link. This includes prompts like, “So, what are you thinking?” and “Why do you think that?”

conclusion and were able to articulate their thinking. This demonstrates that while the students could engage in meaningful discussions, they were unlikely to do so independently.

Although the students did not frequently engage with one another's ideas, there were indications that they were at least occasionally attentive to what other students thought. In some instances when one student was writing independently, other students watched as they wrote. They did not typically provide feedback however, and when they did it was often on superficial features such as spelling or handwriting. On one hand, this suggests that the students were thinking about the task as school assignment necessitating the typical teacher requirements of legibility and correct spelling. It also, however, indicates a tacit agreement about their partners' ideas. This implied agreement was also evident in some of the other ways the students interacted with their partners' responses. There were times, for example, where students read what someone else had written out loud, or confirmed what had been written (e.g., "One hundred meters?" "Mm-hmm" G1SD1:108-109). This general agreement may also account for the lack of discussion about the sources.

Collaboration and Student Roles

The students' perception of what constitutes group work may be one possible explanation for the relative lack of discussion during the source evaluation task. This was evident in the ways that students talked about and assigned roles during their small group work. Group 5 provides the clearest example of these beliefs, because they frequently made explicit reference to whose turn it was to do different tasks and who was responsible for different parts of the activity. There were many discussions about whose turn it was to read or write about a source. This often resulted in only that student doing the work of developing or recording an idea. In fact, some of the students seemed to think that each person should have a turn at being the "thinker" and that the other students were not responsible for coming up with ideas during that turn. During the history investigation, for instance, Jason asked the group, “How come you guys never share your ideas?” and Amira responded, “‘Cause it’s not my turn to write. I’m the last one" (G5HD2:112-114). Daniel, on the other hand, seemed to think that the person writing was only responsible for recording ideas, not coming up with them and said, “Just tell me what to write. Okay, tell me what to write” (G5SD1:410). Similarly, when Zahra was the one recording the answers, she said, “Because what? Think of a shorter idea” (G5HD2:36-37) indicating that she thought only one person at a time was responsible for coming up with responses to record. While the students in Group 5 were the most explicit about delineating different roles and their responsibilities, the other groups’ activities suggest they had similar ideas. As with Group 5, the other groups to tended to simply take turns sharing their ideas rather than discussing them.

Genuine Discussion

Despite the overall lack of genuine discussion among the students, there were some indications that they were capable of engaging in meaningful dialogue without prompting. Although brief, the following exchange demonstrates such an instance:

> Aisha: Oh, many people were excited about having a new vote. That’s why. That’s a piece of evidence.
>
> Yasmin: That is not—
>
> Brandon: You literally just copied that. No, you have to make your own sentence.
>
> Aisha: No, you can figure it out by reading
>
> Taylor: So a democrat—
>
> Aisha: People decided to change their vote. Okay. (G3HD2:103-108)

Here, Brandon responded directly to Aisha's suggestion with a substantive criticism. He did not seem to disagree with the content of her statement but did take issue with the form that it took, echoing the stereotypical teacher's direction to "put it in your own words" because she had suggested a direct quote from the source. Aisha initially disagreed with Brandon, but then rephrased her initial suggestion to satisfy Brandon's demand. Not only were the students all on topic during this exchange, but the listened to and responded directly to each other's ideas in meaningful ways.

Disciplinary Differences

In addition to using inferential/evaluative reading strategies more frequently in science than in history, there were some other differences between the students' responses in the science and history tasks. First, students' responses in science were much more detailed than in history. Although the number of responses for the evidence in each field was nearly the same, see Table 28. Students used considerably more words in their evaluations of the evidence in science. The total word count of their responses in science was 796 compared to 520 in history. The average length of the responses in science was 13.3 words compared to 8.5 in history. As an example, Group 5's longest response in science was 25 words, compared with only 10 words for their longest response in history.

Table 28: Total Responses and Word Count for Evidence Worksheets.

Subject	Total Responses	Total Word Count	Average words per response
Science	60	796	13.3
History	61	520	8.5

Students were also more likely to bring in outside information into their response in science than in history. Table 29 presents where the information in students' responses came from. The outside information code indicates when students brought information from outside the source into their response evaluating a source. Source 1 in history, for example, was a table presenting the results of the first convention for deciding Michigan's statehood. Group 1 responded that "The whigs wanted to keep Toledo." While this is true, the table only indicated how many people voted for or against the proposal, not what political party they came from.

Table 29: Evidence Source Code Count.

Element of Source	Science		History	
Title	0	0%	0	0%
Headnote	4	7%	16	26%
Main Text	47	78%	41	67%
Attribution	0	0%	0	0%
Outside information	11	18%	3	5%

As Table 29 shows, there were many more instances of students using outside information when they were evaluating the science evidence than when evaluating the history evidence (18% versus 5%). Despite the indications from the pre-test that students had more background knowledge of science than history, this does not seem to be related to their prior knowledge of science. Rather than bringing general scientific knowledge into their evaluations, the students were bringing in details from other sources, the investigation question, and the introductory lesson. Instead, the larger amount of outside information in their responses is related to the amount of inferring and evaluating that students were doing. Of the 11 responses coded as outside information, 9 were also coded as inferential/evaluative. This is because in order to interpret what the source was saying, often in light of the investigation question, students needed to relate what the source was telling them to information from other parts of the investigation. As an example, take Source 5 which presented students with a table comparing the force used to push a cart and the time it took to travel 100 meters. Group 4 recorded, “The blue monster because The people are stonger.” In this case, the source did not say anything about the Blue Monster push-cart or how strongly it would be pushed. But because the students were using the evidence to draw a conclusion about the investigation question, they needed to bring in details from other sources into their analysis of this piece of evidence.

Taken together with the students’ tendency to make inferences from the texts more frequently in the science task, these findings suggest that students had a better understanding of the science task and content overall. First, by providing more detailed answers that often incorporated outside knowledge, the students are indicating that they have a better understanding of the content. Or, at the very least, that they believe they have a better understanding since just because their answers are more comprehensive does not necessarily mean they were always more accurate. Second, the trend toward inference suggests that the students felt they had a basic comprehension of the texts and were thus able to focus more on the implications of the information in the sources, not just what information each source contained.

Table 29 also shows that students were more likely to use the headnote for information in the history task than in the science task. One possible explanation is that there was often more detail in the headnotes of the history task than the science task. This was because so much context was necessary to understand the sources in the history task. As a result, the headnotes in the history task provided more information that was relevant to the investigation.

Discipline-Specific Practices

In contrast with the above differences between the students’ activities in the history and science investigations, one similarity was that students rarely engaged in disciplinary-specific ways of engaging with the sources. In the history investigation, for instance, there were few indications that students used anything resembling the sourcing, contextualizing, or corroborating heuristics. There was only one instance, for example, in which students referred to the attribution included in the sources, when Brandon read aloud, “By John-Carlos Marino” (G3HD1:32) from Source 1 of the history investigation. He did not seem to do anything with that information, however, such as taking it as an indication of the source’s relative trustworthiness. Similarly, there were other instances where students mentioned details about the source that might be useful for the sourcing heuristic but did not use it in such a way. In the science investigations, for instance, some of the responses to Source 4 noted that it was Edgar (from the Blue Monster team) who had the ideas about the other team’s cart. But as with Brandon, they did not consider how Edgar’s point of view might affect their interpretation of the source. Similarly, the students did not seem to use science-specific strategies for evaluating the sources, such as considering the quality of the first-hand data, when engaging in the science investigation.

Early Skills in Analyzing Sources

Although the students appeared to face a number of challenges in interpreting the sources, their responses and discussions during the task indicated that, even with no experience and minimal preparation, they were able to engage in the practice successfully. Many inferential responses, for instance, were relevant to the investigation question. While the description of their reading strategy above indicates that students were less likely to make inferences based on the text, when they did, their responses frequently suggested an answer to the question or provided support for one of the possible claims. In the history task, for example, several responses inferred that the Whigs were against the revote, such as “Whigs were against the re-vote” and “Wighs realy realy did not want to do this.” In these inferences from the text, the students were beginning to build the case that there were groups against the re-vote, and thus providing evidence to counter the claim that the second vote was the result of popular opinion.

In addition to successfully drawing inferences from sources, the students also handled the difficulty of the texts in the history investigation well. Despite attempts to make them roughly similar, the sources in the history investigation were more challenging than the sources in the science investigation. They required a good deal of background knowledge that the students were unlikely to have, and the complexity of the issue relative to the science task was reflected in the texts. Despite this, many of the responses to the history sources were accurate and provided reasonable interpretations. The students were generally able to identify, for example, that political parties were groups of people with different opinions who, in this case, were at odds with each other. Some groups also saw that the beliefs of a political party, such as the Democrats, did not necessarily represent the beliefs of the public at large. Their relative success at using these texts suggests that students of this age may be able to handle, and learn from, complex and novel texts even as they engage in unfamiliar practices like analyzing evidence in support of a claim.

Another indication that students of this age are ready to practice argumentation is the degree to which they relied on the sources and their apparent ideas about evidence. The class received only minor instruction in analyzing sources and constructing argument. Their preparation comprised briefly modeling how to use the materials in the “Who stole the cookie from the cookie jar?” example, being guided to look at each piece of evidence one by one, and the prompt, “What does this piece of evidence tell you about the investigation question?” repeated on each source evaluation worksheet. Even with this relatively minor preparation, several students foregrounded the importance of basing their ideas on the sources and the role of evidence in the activity. As described above, Aisha's comment that "you have to tell me a piece of evidence" (G3HD2:666) was one example of a student picking up the practice of looking for evidence from the text rather than just any arbitrary detail.

Conclusion

This chapter described how students read and evaluated sources in preparation for answering the investigation questions and constructing arguments in the science and history tasks. Figure 18 presents the landscape of findings related to how students evaluated evidence in preparation for answering the investigation questions and constructing supporting arguments. In general, the students seemed to have a stronger understanding of the science task and content when compared to the history investigation. This was indicated by their more detailed responses and greater likelihood of incorporating outside information, and stronger beliefs about the answer to the investigation question. In terms of reading strategies, the students were more likely to make inferences from sources in the science task and more likely to read for literal comprehension in the history task. In the figure, literal readings are represented as a minor challenge and interpretive readings as a minor success because overall, their interpretive readings were more relevant to the investigation than the details from the texts they chose to highlight. There were also indications that the students' existing ideas about the investigation question affected their reading of the sources. Some students' readings of the sources were confirmatory where they ignored counter-evidence in the sources and focused heavily on details that supported their answer to the question, in particular for the science investigation. The social aspects of argumentation posed a significant challenge to students in both investigations. The students generally did not engage in meaningful discussions about the text. Although there were instances where they responded to and built on one another's ideas, they were more likely to simply take turns stating their ideas or allow one person to be responsible for coming up with a response.

Figure 18: Landscape of Findings Related to Source Evaluation.

Despite the challenges the students seemed to face, their work did suggest a nascent ability to evaluate sources. In both investigations, when the students made inferences from the text, their inferences were often reasonable and relevant to the investigation question. Second, several students appeared to grasp the primacy of evidence in developing an argument. Finally, even with the challenging nature of the history texts the students were generally able to interpret them and make sense of new concepts as well as identify ideas connected to the investigation question. They made inferences in both investigations but were slightly more likely to do so in the science investigation. In the next chapter, I describe how students used the sources they evaluated to construct arguments supporting their answers to the investigation question.

CHAPTER 6

Students' Use of Evidence to Construct Arguments

In this chapter I present findings related to the ways that students used the evidence from the previous part of the investigation to develop claims and construct supporting arguments. This analysis is based on the portion of the science and history investigations in which students worked in small groups to use their notes on the source evaluation sheets to complete the SenseMaker boards. Figure 19 provides an example of a completed SenseMaker board.

Based on the students' written responses and discussions during small group work, I make five major assertions, which I address in the sections that follow:

A) Students did not often engage in substantive discussions about their arguments or conclusions.

B) Some students used the sources to come to a conclusion about the investigation, while other students fit the sources to support their existing ideas about the answer.

C) There was considerable variability in the quality of students' arguments, and the evidence cited did not always support their claims.

D) There was only one instance of a response that included explicit reasoning tying evidence to the claim.

E) Given the limitations above, some of the supports built into the task and materials appeared to help students construct arguments with some features of high-quality arguments.

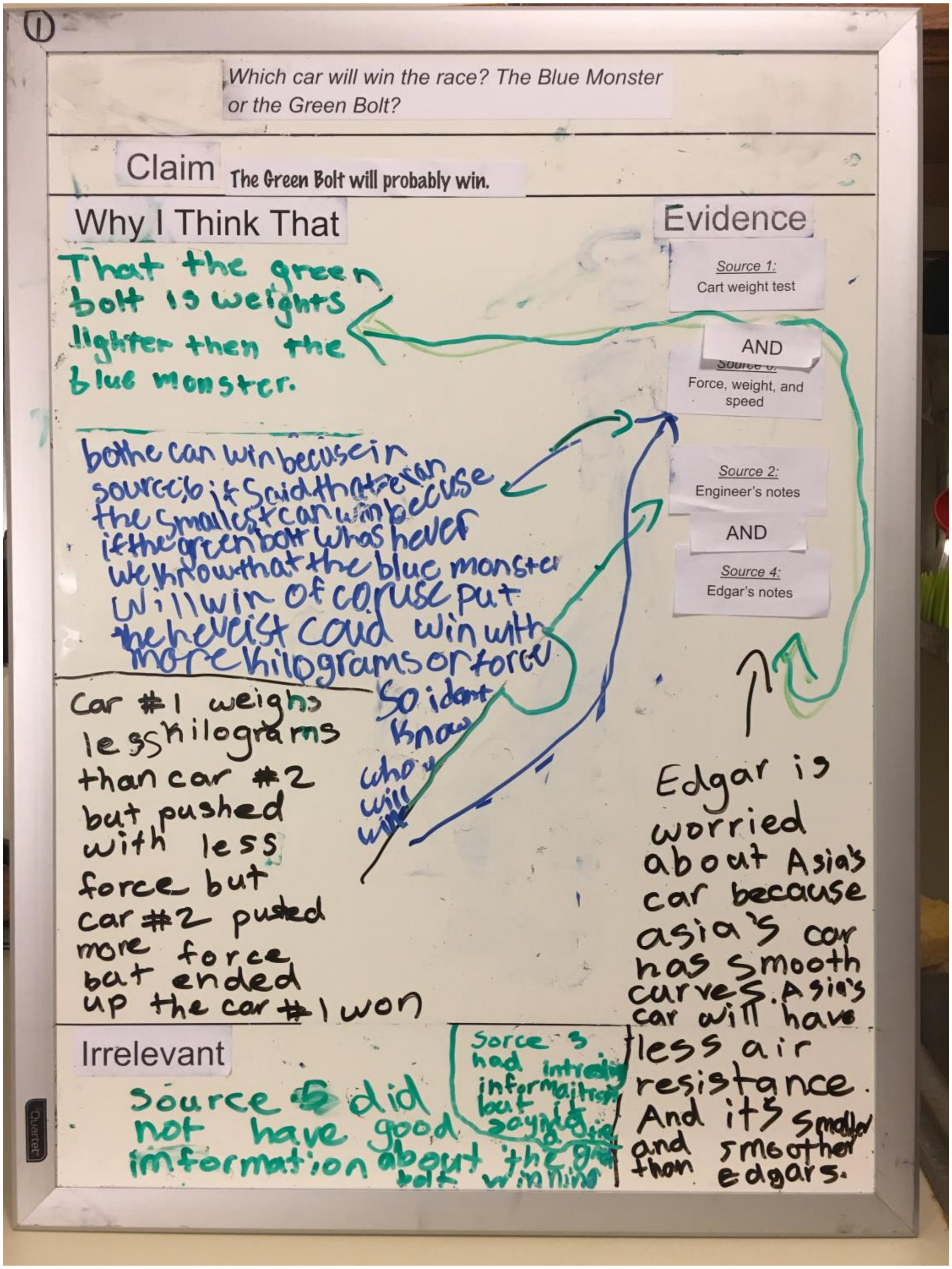


Figure 19: Group 1 — Science SenseMaker Board.

Lack of Discussion

As with the source evaluation task (see Chapter 5), there were relatively few meaningful discussions about the task among the students in their small groups. This was fairly consistent across the science and history tasks, and there did not seem to be any differences in the quantity or quality of their discussions between the two content areas.

Group 3's work during the history task, for instance, is representative of the kinds of interactions that students had during the small group work. In general, when one student made a suggestion to the group, that suggestion was taken up or rejected without any discussion about its strengths, weaknesses, or why other students in the group agreed or disagreed.

This pattern was evident from the beginning of the task when the students began sharing their ideas about the answer to the investigation question. Two students thought the answer was "both" and the other two thought it was "the Democrats." Aisha told the group that, "We have to settle this" (G3HD3:14) and later suggested that the group take a vote (G3HD3:20). The vote did not occur, however, because I arrived to check-in with the table and prompted them to give reasons for their respective answers to the investigation question. Taylor, Brandon, and Aisha all shared their thinking with me but did not respond to one another's ideas or appear to otherwise be engaged with what their group mates thought. Then, when I left their group, they ultimately decided to settle on an answer through Rock, Paper, Scissors. Taylor won both rounds and so the group went with the Democrats as their answer to the investigation question.

In addition, the group's recorded responses on the SenseMaker board were not the result of deliberations. The first two supporting claims came from one student dictating the response to another student who wrote it down. The third and fourth responses were the result of students independently writing their own ideas without input from the group. This pattern was fairly typical among the rest of the groups. In both the science and history tasks, the majority of Group 1's responses were the result of a student working independently or dictating a response to a second student. Only one of Group 1's responses in science appeared to be the result of the students working cooperatively to develop a justification for their claim. In Group 5, the majority of the responses on the SenseMaker board were Jason's. He either dictated a response to another student or recorded the response himself. In the science task there was one instance of the group appearing to collaborate, and one instance of Amira writing independently, but these were atypical.

Student Roles

As in the source evaluation portion of the investigation, the lack of discussion may be attributed to the students' understanding of group work and the different roles that they might play during the work. As with Group 5's work during the history source evaluation task, the students in Group 3 had some discussions about the task's different responsibilities during the history investigation. In particular, they separated the tasks of "writing" and "thinking." Early in the task, for instance, Aisha proposed that, "The person that doesn't have the marker is gonna tell us and stick on the stuff [source label sticky notes]" (G3HD3:65). Here, she seems to say that only one student is responsible for developing ideas. Later, Brandon tells Yasmin that she "can help us with that– with the ideas" (G3HD3:433), showing that he believes each person should get an opportunity to provide the idea with the group. In addition, there were the instances described above of the students either only dictating their idea to the "recorder" or independently writing down their idea.

Taken together, these statements and behaviors seem to indicate that the students did not necessarily see this as a deliberative activity. That is, they did not see it as an activity that required a discussion where they shared their ideas back and forth and collaboratively worked out a solution they all agreed on. Instead, they appeared to see it as an activity in which they each got to take turns sharing their ideas in much the same way that they had to take turns to use the marker or glue things on to the SenseMaker board. Rather than working together to develop their ideas, they simply accumulated a set of ideas that were inconsistently connected.

In addition to assigning roles such as “recorder” and “thinker,” there was also the tendency of some students to take on a leadership role. In Group 1, Angela led much of the group’s work and in Group 5, Jason was responsible for much of the group’s writing and discussion. No single leader emerged from Group 3’s work. As a result of Jason and Angela’s leadership, much of the activity described here reflects their work and thinking. They were typically the ones prompting the group for things like evidence or suggesting and recording responses. Furthermore, the other students in the group often allowed Jason and Angela to take the lead, much in the same way as the students in Group 3 often seemed comfortable with one person being responsible for the thinking or recording of responses and only occasionally interjected when one student took the lead.

On and Off Task Talk

While there were relatively few examples of back and forth discussions about the students’ claims and how they would support them, this did not mean that all of the students’ talk was off-task. There were many instances where the students were talking about something relevant to the activity, but they were often about superficial features of the answers, such as spelling, or the mechanics of the activity, such as filling out the SenseMaker board.

Group 1’s work in science, for instance, provides a good example of this kind of talk. They often talked to decide whose turn it was for a particular job or to manage the materials. Caitlin, for instance, seemed very interested in making the board look good and tried to manage how it was assembled by telling the group, “No. Don’t put it [an evidence title slip] that close. Don’t put it that close. Put it right there” (G1SD3:291).

There were instances when the students discussed the contents of the sources, but, again, the students were largely addressing how to complete the SenseMaker board, not how the source related to the argument. In one case when Caitlin was drawing a line between one of the supporting claims and the “Source 2” slip, Angela said, “No. It isn’t source two. Solomon said it wasn’t source two. Why would you connect it to source two if this wasn’t even source two?” (G1SD3:313). Although she was addressing Solomon’s claim and supporting evidence, it was only on a superficial level in order to determine if the SenseMaker board was accurately depicting what Solomon wanted to say. Angela was not indicating agreement or disagreement, nor responding to Solomon’s claim in any meaningful way and was thus not engaged in a discussion about their argument. In another case, Caitlin was trying to place an evidence slip on the board in the correct place and asked, “What do I do with this one?” and Anna responded, “This one was supposed to be right here ’cause this is telling him that his weight’s lower than the other” (G1SD3:289-290). Here, Anna indicates that she is actively engaged in constructing their argument by identifying what source is linked to her supporting claim. In terms of a discussion, however, she and Caitlin are not having a substantive exchange about the source or the supporting claim. Instead, Anna is just directing Caitlin how to complete the SenseMaker board according to her (Anna’s) ideas.

In addition to on task talk that addressed superficial elements of the tasks, there was also a good deal of off task talk. Some of the students were often interested in playing games, being silly, or discussing things that were probably more interesting to them than school work. Aisha, for example, was very intent on playing the “jinx” game with the other members in her group. As a result, Yasmin was “jinxed” at several times during the session and did not talk. Group 1 got into a long discussion about the length of their hair, the color of their hair, and getting a haircut. This extended conversation was sparked by a question about who had glued down a particular piece of paper on the SenseMaker board, because “It was someone with short hair” (G1HD3:419), presumably because a hair had gotten stuck to the board as well. In the history investigation, Amira, Daniel, and Zahra got into a heated discussion about whether or not Daniel was being mean, while keeping on the lookout for me as I circulated through the room.

Strategies for Developing Claims

In this section I describe the students’ strategies for developing and supporting claims. By strategy, I mean the ways they used evidence in relation to their claims, and I consider two possible strategies. First, students can use the evidence to support their existing ideas about the answer to the investigation question. Second, students can use their evaluation of the evidence to develop their claim. In this second scenario, their opinions about the investigation question are shaped by their interpretations of the evidence. Figure 20 visualizes these two strategies. In some instances, it was possible to evaluate students’ strategies based on their discussions and recorded answers on the SenseMaker boards and source evaluation sheets. In many cases, however, the lack of discussion among the students made it difficult to identify how students were developing their claims and coming to their conclusions. Consequently, the descriptions below do not necessarily mean that one strategy was more prevalent than another. Instead, they are only meant to present the possible ways that students might approach the task.

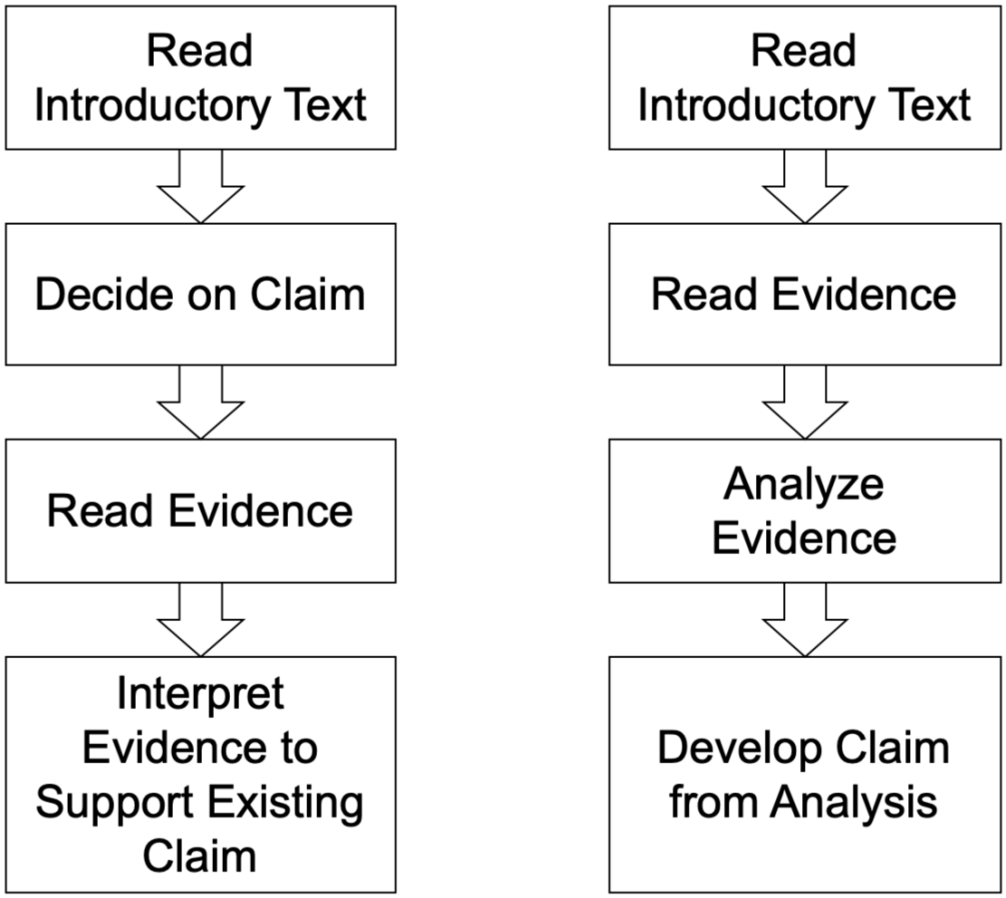


Figure 20: Strategies for developing claims and using evidence.

The clearest example of students coming to a conclusion based on the initial information of the investigation question comes from Group 1's activity during the science investigation, and Eric in particular. In the students' discussions it was clear that they had developed beliefs about the investigation question early on. In fact, it was the first thing they considered on the first day of the science investigation in which their task was to evaluate the evidence, not develop an answer to the investigation question (students did not use the SenseMaker board to develop their argument until the third day). After being introduced to the investigation and given the background information they went to their seats and Caitlin immediately asked:

Caitlin: Who thinks the Blue Monster will win?

Angela: We don't have the evidence yet.

…

Caitlin: I know that we don't have the piece of evidence yet, but who votes?

Solomon: I vote.

Caitlin: Who votes that the Blue Monster's gonna win? Who votes that the Green Bolt—

Angela: We are just wasting our time.

Caitlin: No! We're not.

Eric: Green. (G1SD1:6-26)

This exchange demonstrates that some of the students had already developed ideas about the answer to the investigation question, or that they thought they could have an answer prior to looking at the evidence. Caitlin seemed to think that it was possible to answer the investigation question with only the information contained in the question and some of the students had already formed opinions. Eric thought the Green Bolt would win and Solomon, though he does not give an answer, shows that he has one in mind by his willingness to take part in the vote.

It is possible that even though the students did not use the word "prediction," they were actually making predictions and not proposing a definitive answer to the investigation question. Making this determination is difficult, especially in light of the fact that the answer to the investigation question would itself be a prediction. I believe, however, that given some of the students' later comments in which they held to these beliefs, they were attempting to answer the question itself as opposed to making a prediction about the possible answer.

The above exchange also shows that Eric was convinced from the beginning that the Green Bolt would win the race. Eric not only repeated this claim later on in the day and on the second day of source evaluation, he also read the evidence in ways that supported his belief that the

Green Bolt would win, further demonstrating that he developed his claim first and then found evidence to support it later. This included ignoring evidence that favored the Blue Monster winning and interpreting ambiguous evidence in favor of the Green Bolt. When reading source 2a, for instance, Eric twice asked, “Which car is the lightest?” followed by the prediction that the “Green Bolt will win” (G1SD1:389-399). Source 2a was “The notes of an engineer that makes push-carts … [who] tried different combinations of force and weight.” With the typo in source 2a, however, the source reported that in a race between a heavier car with a small amount of force went faster than a lighter car pushed with a great deal of force. By prefacing with “Which car is the lightest?” Eric indicated that he believed weight to be an important factor in the results. As source 2a is written, however, the heavier car won. Eric ignored this and predicted that the Green Bolt (the lighter car) would win. In addition, he did not attempt to make sense of what should have been a surprising result, and instead jumped to supporting his initial idea of the winner.

In his reading of Source 3, Eric ignored another student’s counterargument and made a conclusion about the role of weight that is unsupported by the source. Source 3 explained air resistance and Anna concluded that the source explained, “That if it’s light, it could go slow still, ’cause the air pushes it, and it could push back.” Eric responded with, “No. If it’s lighter then it’s easier to push” (G1SD1:532-535). Here, he disagreed with Anna’s reasonable conclusion about the source, and restated his idea about lightweight cars. The source, however, did not mention the weight of the cars, only how this shape can affect the speed. Eric seemed fixated on his idea and read or ignored the evidence in ways to support that idea.

Eric also actively ignored evidence in the sources when reading Source 5 which reported the results of trials with a single cart pushed with different levels of force. This source mostly supported the Blue Monster (the heavier of the cars) as the winner because it is pushed with more force and the source shows that the greater the force the faster the car. Anna concluded that, “The one that has the more power [is going to win], ’cause this is 30 [newtons], and it has less time [i.e., traveled a certain distance faster than other cars pushed with less force].” Eric countered with, “No, ’cause maybe the lighter one is easier for the one that has less newtons” (G1SD2:355-362).

Eric’s insistence that the Green Bolt would win because it was lighter continued through the third day, when the group constructed the SenseMaker board. Eric adamantly claimed the Green Bolt would win, saying, “Green bolt, green bolt, green bolt. Green, green, green bolt!” (G1SD3:466) but did not cite any evidence. The only time that Eric cited evidence while his group completed the SenseMaker board was when he suggested that source 4 would support his claim because “it says Ebony and Asia have more chances of winning” (G1SD3:100). Taken with his previous tendency to focus on favorable evidence and ignore counter evidence, this suggests that Eric’s strategy for developing and supporting an argument was to come to a conclusion based on his initial impressions of the scenario and then find evidence that would support that existing view.

In contrast, Angela, also in Group 1, appeared to develop her claim by starting with the evidence, analyzing it, and coming to a conclusion based on that analysis. Her strategy was evident on the first day of the science investigation when the students began reading evidence. In the above exchange, after Caitlin prompted her peers for an answer to the investigation question, Angela responded, “We don’t have the evidence yet” (G1SD1:7). Unlike Eric, Angela did not suggest any answers to the investigation question while looking at the evidence. It wasn’t until looking at the final piece of evidence on the second day of the science investigation that Angela offered a possible answer. She told the other students in her group, “I think it’s the green bolt … This [source 2b] really actually proves it by saying that this guy has less newtons and it weighed less [but went faster], so I’m guessing it’s the green bolt” (G1SD2:852-854). Angela’s withholding her judgement until the final piece of evidence, suggests that she may have based her conclusions on the evidence itself rather than her initial impressions.

Although she did not discuss her thinking as clearly on the third day of the investigation, when students constructed their arguments and completed the SenseMaker board, some of her behavior suggested that she was starting with the sources rather than an existing belief about the claim. Before recording her ideas on the SenseMaker board or telling the group her answer to the investigation question, Angela re-read through several of the sources on her own. Because she did not express her ideas before looking to the sources, I believe this shows that she was using the sources to develop an answer to the investigation question. Another indication of this comes from an exchange with Solomon. At this point in the activity, Solomon seemed to be undecided about the answer to the question, but Angela believed that the Green Bolt would win. To convince him, Angela tells him to “Keep reading this [Source 2b]” and then to “Think about the kilograms and the force [unintelligible] push-cart number one, push-cart number two” (G1SD3:471-479). Rather than relying on an argument such as Eric’s rooted in her background knowledge that light things can go faster than heavy things, Angela directly used the sources as her backing. This suggests that she began with the sources rather than a conclusion.

On the other hand, when Angela listed sources in the “Irrelevant” area of the SenseMaker board she listed them because they did not support her conclusion. For Source 5 she wrote that it “did not have good information about the Green Bolt winning,” and for Source 3 she wrote that it “had interesting information but it saying it a tie.” These are not reasons that the sources were not relevant to their argument but are instead reasons why they did not support her claim that the Green Bolt would win. This possibly indicates that she was only looking for evidence which supported her claim. Based on her other behaviors, however, I do not believe this is the case. Instead, I believe it is more likely that she was interpreting "irrelevant" as "why I didn't include it." In addition, the students were largely unfamiliar with argumentation and the formal elements of argument. Consequently, Angela likely did not know that evidence supporting rebuttals and counter-arguments would be relevant despite not supporting her final claim.

Taken together, this indicates that Angela's strategy for developing a claim may have been different than her peers in the group, and Eric in particular. Angela expressed a belief that evidence plays a key role in answering scientific questions, or at the least this specific question. It is unclear where she got this opinion. Her behavior during the history investigation suggests that she held this view prior to this study. When making the history SenseMaker board she tried to orchestrate the students to read all the evidence and prompted students to provide evidence, saying things like, "We need to get our evidence out … Take the evidence out, so we know what the evidence are" (G1SH3:57-60).[9] It is also possible, however, that she was not necessarily expressing a belief about the importance of evidence, and just enforcing the parameters of the assignment.

Although the examples above demonstrate the two strategies for developing and supporting claims relatively well, more often, the students' approaches were obscured by their lack of discussion. Group 5's work, especially Jason's, on the history investigation provides a good illustration of the difficulty of determining the students' strategies. During the source evaluation

[9] Angela's use of the sources in the history task is discussed in more detail in *Using the Sources* below.

portion of the investigation, the students in Group 5 did not make any attempts to answer the investigation question. Instead, they seemed to focus on how individual pieces of evidence supported one claim versus another, which was the goal of this part of the activity. This is demonstrated in the following exchange where they discuss Source 5 (Figure 15).

> Marino: What is this telling you about if it was popular opinion or if it was the Democrats?
>
> Jason: I'm looking at the headnote, but—
>
> Daniel: Democrats, but they controlled—
>
> Jason: – it was—I'm pretty sure it was Democrats. (G5HD2:656-659)

Here, Jason suggested that the Democrats were responsible for the second vote on Michigan statehood, but it is clearly in response to the source itself and an analysis of what conclusion the source supports. He cited specific elements of the source itself, "I'm looking at the headnote" and used that to come to his idea that he was "pretty sure it was the Democrats." Rather than reading the headnote to support his existing ideas, it appears that Jason's conclusion came out of his reading of the source.

On the third day of the investigation when they completed their SenseMaker boards, however, Jason appeared to jump to a conclusion about the investigation question without seeming to take into account the evidence. As soon as the students got to their seats, Jason said, "Decide. Guys, we have to decide," and then, "The second vote happened because of popular opinion" (G5HD3:2-6). Prior to suggesting this answer, however, Jason did not review the evidence or look back at their notes. The speed of his response suggests that Jason may have had this idea going into the argument construction portion of the task but there were no indications of where he got this idea. Jason did not elaborate on his thinking with the group and began writing responses on the SenseMaker board independently. Based on the how quickly he gave his answer without looking at any of the sources, it would seem that Jason's strategy was to start with a claim about the investigation and then work backwards to find evidence that supported his claim.

A further complication in analyzing Jason and Group 5's strategy for developing a claim and using evidence is that two of the three supporting claims they provided actually favored *Democrats* as the answer, and not popular opinion as their claim stated. As Table 34 shows, their first two responses emphasize the role that Democrats played in convening the second convention. On one hand these supporting claims were consistent with their initial evaluation of the sources. As these sources did, in fact, generally support the Democrats as the answer to the investigation question, this suggests that they may have been reading the sources objectively. On the other hand, they then used these details to support the competing claim. This indicates that they may have been doing something like Eric and were interpreting counter-evidence in ways that would favor their existing ideas.

The group's final response on the SenseMaker board also introduces the possibility that the difficulty of the texts played a role in their inconsistent approach. Their final supporting claim was linked to Source 5. In Jason's initial reading of the source, as shown in the above exchange, he appropriately determined that it supported "Democrats" as the answer to the investigation question. When it came to constructing the SenseMaker board, however, he used the source to support his claim that popular opinion was responsible for the second convention. In his written response, he wrote that "(By new articals) leader's let citizen's decide to become Michigan." In this response, he seemed to believe that the source told about an article being distributed to the general public after the first vote. In fact, the headnote explained that source itself was an article from a magazine called “Michigan History.” The text of the source described how Democrats’ put out a notice that there would be a second election.

Somewhere in Jason’s rereading of the source he seems to have gotten confused and changed his interpretation. To read this source accurately, the students needed to understand that while the source was an article, it was not contemporaneous with the events described and that the author was presenting their own interpretation of events. Jason’s attempts to make sense of the source in light of the investigation question and as part of constructing an argument demonstrate the difficulty of the task and provide a possible explanation for the group’s seemingly inconsistent strategies.

Quality of Students’ Arguments

The quality of students’ arguments can be evaluated along a number of criteria. The following section breaks down the students’ arguments, supporting evidence, and other justifications to describe how well students constructed arguments supporting their answers to the investigation questions. I begin by describing the two ways that students justified their claims, either through statements of fact or with interpretive statements. Next, I describe how accurate and reasonable the students’ justifications were and whether or not they supported the claim. Finally, I review the extent to which the students provided reasoning in their arguments and highlight a few arguments that included sophisticated elements of argumentation.

The Types of Justification Provided

I begin by describing the ways that the students justified their claims. While many studies evaluate the quality of students’ arguments based partly on whether students provide any justification at all, such an analysis is not possible in this study. Because the students were using the SenseMaker board, they had to provide some kind of justification for their argument. In contrast, a prompt eliciting a more traditional written response would leave open the possibility of students simply repeating their claim. Consequently, I describe the two main types of justification the students used to support their evidence: “facts” reported from the sources and “interpretive statements” based on the sources.

To justify their arguments the students tended to provide details from the text, rather than inferential or evaluative readings of the text. This is similar to the work they did during the source evaluation task. Across the tasks in science and history, students gave more answers that consisted of a straightforward detail than answers that included some kind of inference. Table 30 presents the number and proportion of responses consisting of “facts” (i.e., details) compared to interpretive responses. The details that students provided included paraphrasing or summarizing the text such as, “Egar is worried that his team might lose because Asia push-cart is smaller” (G5SenS), directly quoting the text as in “The governor gave advice to the democrats” (G1SenH), or providing a detail from outside of the source, such as, “Because if they [the Green Bolt] have 75 kilograms so it will be lighter” (G4SenS; the source did not mention the weight of the Green Bolt).

Table 30: “Fact” vs. Inferential Responses.

Type of Response	Science	History
“Fact” Responses	63% (10)	64% (14)
Paraphrase	*50% (8)*	*36% (5)*
Quote	*0% (0)*	*23% (5)*
External Detail	*13% (2)*	*5% (1)*
Inferential Responses	38% (6)	36% (8)

Answers that went beyond the basic meaning of the text were considered to be interpretive. In these responses, the students made inferences based on the text and sometimes included a conclusion about the investigation question based on their interpretation of the source. Group 4’s

SenseMaker board for the history investigation, for example, wrote, “they really wanted to become a state and they had different belifes!” (G4SenH) and attributed it to Source 3. Source 3 presented the table of the results of the second convention on Michigan statehood and included the following headnote:

> This chart tells about the people who went to the second convention to vote on Michigan’s statehood on **December 14, 1836**. There weren’t as many counties at this convention as there were in the first one. Many in the Whig party did not think the vote should happen. So some Whig counties did not attend. Almost all of the delegates at this convention were Democrats.

Based on the unanimous results of the vote, the students in Group 4 concluded that “they really wanted to become a state” (although it is somewhat unclear who “they” are) and using the details about the Whig party’s opinions about the vote concluded that “they had different belifes!” The students went beyond a simple paraphrase or quotation of the text, and instead provided an inference of what the text meant.

Other responses were more clearly tied to the investigation question. In the science investigation, Group 5 wrote, “even though the blue monster is bigger and heavier and the people who are controling still won’t win because its heavier” (G5SenS). Here, they provide their prediction that the Blue Monster “still won’t win” based on the text which at one point states that, “Lighter cars often go faster than heavier cars.” Here, the students are not only going beyond the basic meaning of the text but using the text to support their claim. In addition, they deliberately connected the inference to the investigation question by using “because” and including a reference to the investigation question (i.e., the Blue Monster) in their response.

Accuracy

Another feature of high-quality arguments is that the supporting evidence is accurate and the justifications are reasonable. The accuracy and reasonableness of groups' responses in the two tasks was somewhat different between the science and the history task. Table 8 in Chapter 3 presents the codebook for assessing the accuracy and reasonableness of the students' responses on the SenseMaker boards. Responses that only consisted of a detail from a source were coded for how accurately they represented the information from the text. Responses that included an inference that went beyond the basic meaning of the text were coded for whether or not the inference was reasonable based on the text and the other information from the investigation.

As Table 31 shows, a greater proportion of the students' responses in history were either accurate or reasonable when compared to their responses in science. Only 50% of the students' answers that consisted of just a fact were accurate in the science task in contrast with 71% of the responses being accurate in the history task. Sixty-seven percent of the interpretive responses in science were reasonable compared to 75% of the interpretive responses being reasonable in the history investigation.

Table 31: Accuracy and Reasonableness of SenseMaker Responses.

Type of Response	Science	History
"Fact" Responses		
Accurate	50% (5)	71% (10)
Inaccurate	10% (1)	22% (3)
Both	10% (1)	0% (0)
Unclear	10% (1)	0% (0)
Not Applicable	20% (2)	7% (1)
Inferential Responses		
Reasonable	67% (4)	75% (6)
Unreasonable	0% (0)	25% (2)
Unclear	33% (2)	0% (0)

This result was surprising, as the texts in the history investigation were likely more challenging to read. The classroom teacher, for instance, remarked that the texts would likely be very difficult for the students. In addition, the sources in the history task necessitated background knowledge, such as the meanings of “party” and “representative,” that the students did not seem to have going in to the task. Despite these challenges, the students were able to do a fairly good job of accurately identifying details from the sources and making reasonable conclusions.

Degree of Support for Claim

Although the students’ SenseMaker responses were generally accurate (for statements of facts from the evidence) or reasonable (for inferences from sources), this does not mean that students always supported the claims they were making in their argument. Table 32 presents the number of responses that support or did not support the claim in each investigation. In general, the students had a harder time providing responses that supported their claim in the history task when compared with the science task. This is despite the fact that a greater proportion of their responses in history were accurate and reasonable compared to the proportion of accurate responses in the science task. See Table 31.

Table 32: Number of SenseMaker Responses that Supported the Claim.

Degree of Support	Science	History
Supports Claim	63% (10)	50% (11)
Does not Support Claim	25% (4)	32% (7)
Unclear	13% (2)	18% (4)

It is worth noting that I was fairly generous when coding for the degree of support the responses provided for their claims. This was due to the fact that the responses, in particular those only providing facts, were rarely explicitly connected to the investigation question. Responses such as “lighter cars often go faster than heavier cars” (G6SenS) and “the governor was a democrat" (G1SenH, G3SenH, G4SenH) did not include an explanation as to how those facts supported the claim. In cases like these, then, I made a judgement as to whether or not the response could *reasonably* be taken to support the claim.

As the Table 32 indicates, a several of the responses did not actually support the claims that students were making. There were two main ways in which the responses did not support the claims. In some cases, the students' responses simply supported the wrong conclusion. In the history investigation, for example, Group 5's claim was that popular opinion was the cause of the second vote on Michigan's statehood. One of their responses, however, was that "Because Democrats tried to convince others to vote for becoming a state" (G5SenH). On one hand, this was a reasonable interpretation of the source they cited. Unfortunately, it supported the opposite claim that the group was making. In fact, Group 5's second supporting reason, "In the letter the government gives advice to the democratics," also does a better job of supporting the answer to the investigation question that the students did *not* give. It is unclear as to why this is happening. One possible explanation is that there was disagreement in the group about the answer. This appeared to happen with Group 1 in the science investigation when Solomon thought differently than the rest of the group and recorded an answer that supported the claim that it would be a tie, when the rest of the group thought it would be the Green Bolt.[10] In Group 5, however, Jason stated to the group that he thought popular opinion was the answer, and then was largely responsible for recording both responses the supported a different claim.

Second, there were also a number of cases where the students' responses were unclear enough to make it difficult to assess if they were actually supporting the claim. In the history

[10] Solomon, incidentally, changed his mind about the answer. While he ultimately agreed with the group, they did not go back and revise the supporting claim on the SenseMaker board.

investigation, this was because of unclear references to "they" and "people." Group 3 thought that the Democrats were responsible for the second elements and wrote that "People wanted a new vote!" (G3SenH); which people, however, is ambiguous. The remaining three unclear responses in the history task, had a "they" with an unclear referent: "they really wanted to become a state and they had different belifes!" (G4SenH), "because they work so hard to become a state" (G1SenH), and "they wory about not having statehood" (G4SenH).

Providing Reasoning

Another way to evaluate the quality of an argument is to look for how, or if, the evidence is tied to the claim. This is often referred to as "reasoning" in the models of argumentation in education that rely on the Toulmin argumentation pattern. As the number of responses that simply provided a detail suggests, many responses only provided a restatement of a detail from a source without any direct connection to the investigation question. Responses consisting solely of a "fact" only reported information from the sources and included no explanation of how the detail provided supported the claim. This includes responses such as "lighter cars often go faster than heavier cars" (G6SenS), "the governor was a democrat" (G1SenH, G3SenH, G4SenH), and "That the green bolt is weights lighter then the blue monster" (G2SenS) were not explicitly tied to the investigation. In responses such as these, students did not use logical language, such as "because" or "so," that would clearly link the evidence to the claim. Nor did they include the reasoning that tied the evidence to the claim. In the responses above, for instance, responses with reasoning may have looked like, "Lighter cars often go faster than heavier cars *and the Green Bolt is lighter than the Blue Monster*" or "The governor was a Democrat *and had power to make a vote happen*." Consequently, none of the responses that only provided a detail without any further inference were considered to have included reasoning or any coordination between the evidence and the claim.

The “reasoning” element of arguments also has disciplinary implications. In science, one way to warrant arguments is through scientific principle which can tie evidence to claims and explain how the evidence supports a claim. In history, reasoning also includes considerations or evaluations of evidence and whether it is trustworthy, reliable, or relevant. There were no clear examples of either types of reasoning in the students’ arguments. With two minor exceptions, there were no clear examples of students including wither kind of reasoning in their arguments. Group 6 came close to citing a scientific principle when they said that “lighter cars often go faster than heavier cars” (G6SenS). This was a quote from one of the sources, however, so it is not clear if they saw it as a principle that could explain their phenomenon or a “fact” that supported their claim. Group 3, namely Brandon, repeatedly said that “slow and steely [steady] wins the race.” He appeared to be using a general principle as a way to support his claim, but it was neither accurate nor used in a way that tied evidence to his claim.

Identifying when there was reasoning present in the remaining responses proved difficult. The responses were often confusingly written, included inaccurate details from the sources, or unreasonable interpretations of the sources. Despite these challenges, I was able to identify what elements of an argument (e.g., reasoning, evidence) were present in the students’ responses. This analysis also only focuses on whether or not reasoning is present. It does not take into account whether or not the statement is accurate or if it reasonably supports the claim that students are making on the SenseMaker board. First, the appropriateness of the justifications, including interpretive statements, is discussed above. Second, my goal here is to determine whether or not students understand what reasoning is and how to provide it. In this analysis, the accuracy of reasoning is irrelevant.

Of the 27 interpretive responses across both investigations only one clearly included reasoning that tied the evidence to the claim. In the science task, Group 4 wrote, “Green Bolt would win because it would be close to 71 kilograms wich was fast” (G4SSense), and cited Source 2, Figure 21, as the supporting source. This supporting claim actually consists of a complete argument, and includes a claim, evidence, and reasoning. First, they stated their claim, “Green Bolt would win.” This, however, was not strictly necessary as they had already stated the claim at the top of the SenseMaker board. Second, they provided evidence to support the claim, “it would be close to 71 kilograms,” and used “because” to indicate that this explained and supported their claim. Their evidence is that the Green Bolt weighed 75 kilograms, which “would be close” to the weight of the cart that was faster in Source 2. Finally, they connected the evidence to claim (i.e., reasoning) with, “wich was fast.” This explains that the weight of the Green Bolt is relevant because it was similar to the fast car in the source.

While this response demonstrates that students are capable of developing complete arguments independently and without prior instruction, it was not typical of the rest of the answers students gave, even within Group 4.[11] Of the remaining responses, there was not a single trend that describes them all, but in general they left some element of an argument implicit and did not state them directly.

Some responses, even though they were interpretive, left the reasoning unsaid. Group 5’s response in the science investigation provides a good example. They wrote that, “even though

11 Although as is shown below, Group 4 did demonstrate other sophisticated ways of engaging in argument and evaluating data.

Source 2 – Engineer's Notes	Investigation Question: Which car will win the race? The Blue Monster or the Green Bolt?
Headnote: These are the notes of an engineer that makes push-carts as a hobby. She was testing her own push carts. She tried different combinations of force and weight.	What does this piece of evidence tell you about the investigation question?
Today I tested two of my push-carts. Push-cart #1 weighs 71 kilograms. I pushed it with 19 newtons of force. Push-cart #2 weighs 98 kilograms. I pushed it with 24 newtons of force. I raced them against each other. Push-cart #1 beat push-cart #2 by almost 2 seconds.	
Source: *Personal Engineering Journal*. Andrea Hutchins. 2016	

Figure 21: Science, Source 2.

the blue monster is bigger and heavier and the people who are controling still won't win because its heavier" (G5SSense). Here, the evidence supporting their claim that the Green Bolt will win is that the Blue Monster is heavier. The students provide the evidence, restate their claim, link the evidence to the claim using "because," and then restate their evidence. What is missing, however, is an explanation of why the relative weight of the Blue Monster is relevant. While the students do not state this, they imply it through the source they cited, Source 6. In the source, it states that "lighter cars often go faster than heavier cars," a scientific principle that provides the warrant for their evidence. By citing this source, the students appear to have identified the usefulness of what it says but left it out of their final response. The implicit reasoning behind responses like this make these answers similar to those that just reported a detail from the sources. In many cases the evidence did, in fact support the claim, but the students did not articulate the link between the claim and evidence.

Other responses, while making inferences that went beyond the basic meaning of the text, simply restated the groups' claims. Group 4, for example, wrote "The democratics gave them a nuther chance to vote again" (G4SenH) on their SenseMaker board in the history investigation. This is an interpretive answer, because the source they cited did not explicitly say that the Democrats gave a second opportunity to vote. It leaves out, however, the specific detail from the source that made them draw that conclusion. Additionally, the response is basically a restatement of the group's answer that the Democrats were responsible for the second vote. So, although the response is interpretive, it includes neither the evidence nor reasoning necessary to support their main claim.

Using the vocabulary of argument. Despite the lack of explicit warrants in the students' arguments, it was clear that they were still attempting to provide support for their claims. One indication the students were at least attempting to justify their use of particular pieces of evidence was their use of language associated with providing warrants and backing (i.e., reasoning) for an argument. Words such as "because," "since," and "so" can all potentially be used as logical connectors that coordinate evidence with a claim. Looking for this kind of language in the students' responses helps reveal where there may at least have been the possibility of reasoning tying the evidence to the claim. Table 33 presents the number of responses in each investigation that included logical connectors.

Table 33: Responses Using Logical Connectors.

Logical Connectors	Science	History
Connectors Present	56% (9)	23% (5)
No connectors	44% (7)	77% (17)

"Because" was the most commonly used logical connector, and all of the responses that used logical connectors included it. Three of those responses also included "so," showing that some students had a broader vocabulary of the kinds of language that could be used to indicate causality or argument. Table 33 also shows that the students were more likely to use "because" in the science investigation than in the history investigation. Two of the uses of "because" were due to the students directly quoting or paraphrasing a source, but that still leaves seven instances where the students were using it to express an independent thought.

While the use of "because" indicates that there is a link between the cited evidence and claim that it is supporting, this does not mean that students explicitly named that link. Despite using "because" in these responses, elements of the argument remained implicit, even in responses that could be considered sophisticated. Take, for example, one of Group 5's responses to the history investigation, "Because Democrats tried to convince others to vote for becoming a state" (G5SenH). Here, the reasoning is present, but the specific evidence is not. The source they cited, source 4, is a report of a meeting in which the Democrats recommended holding meetings about holding a second convention. This detail is the data, or evidence, that the students appeared to be using to support their claim. Their written response is the reasoning which ties that evidence to the claim but does not include that evidence. A complete response, in other words, might look like, "Because *the Democrats recommended meetings about another vote* they tried to convince others to vote for becoming a state."

Other responses, such as Group 3's response that "The green bolt Is [going] to win because it has less Kilograms and it is smaller so you can push the green bolt" (G3SenS) left out the reasoning that tied their evidence or interpretation of the evidence to the claim. In this case, the students did not provide a warrant for why the fact that the Green Bolt "has less Kilograms" supports their conclusion that it would win. Here, a complete answer might have been, "The green bolt Is [going] to win because it has less Kilograms and it is smaller so you can push the green bolt *faster, because light things go faster than heavier things.*"

Lack of coordination of claims and evidence. There are a few reasons why students' responses may not have included warrants or backing (i.e., reasoning). First, the reasoning tying evidence to a claim is not always explicit, even in sophisticated arguments. *Enthymemes*, or arguments with an unstated premise, are not uncommon because sometimes the link between a claim its supporting data is self-evident. This could arguably be the case for some of the evidence the students gave in their responses. Two of the groups, for instance, included some variation on "The whigs Did not agre" (G3SenH) as evidence supporting the claim that Democrats were mostly responsible for the second vote (Group 3 included this twice in their argument). The implied reasoning here is that, "…so it could not have been popular opinion since a large group of people was against it." It is debatable whether or not such explicit reasoning is absolutely necessary for the argument to make sense or be convincing. Thus, students may not have included it because it seemed obvious to them.

There were also times when the inference was present, but not the underlying fact. For example, Group 4 wrote, "The democratics gave them a nuther chance to vote again" (G4SenH). Here, the source reports that "Almost all of the delegates at this convention were Democrats." The student leaves that detail out of their response and essentially jumps straight to the reasoning. The implied argument being, "*Since it was mostly Democrats at the second convention,* the democratics gave them a nuther chance to vote again." In this example, "the democratics gave them a nuther chance to vote again" serves as the reasoning and "it was mostly Democrats at the second convention" is the implied evidence.

Another explanation is that articulating one's thinking is a difficult and sophisticated task, and as third-graders, these students naturally struggled with it. This is clear in one of Group 4's responses to the history task. They wrote, "it was mostly the demacrats because the demacrats might have annonced the election but the people voted but it was still mostly the democrats" (G4SenH). As described in *Arguments with rebuttals* below, this was one of the most sophisticated responses provided by the students. Not only did it clearly tie the source to the investigation question, but it also included a rebuttal to a counter-argument. Despite this, the students clearly had difficulty putting their thinking into words. Their use of "might have," for example, weakens the evidence that Democrats announced the election and implies that it does not actually account for the reasons of the vote.

There is also the likelihood that students were not accustomed to explicitly providing reasoning tying claims with evidence, or simply did not know how to do so. As I discussed in the previous chapter, they may just not be used to reading a text for a purpose other than proving they have read the text.

Finally, reading comprehension may have also played a role, particularly in the history investigation where the texts were more challenging. If students were having a difficult time making sense of the texts it follows that they would also have a hard time explaining their ideas about them and how they fit into their larger argument.

Other High-Quality Features

While many students struggled with accuracy, providing evidence that appropriately supports the claim, and providing reasoning, there were some responses that indicated the students were capable of sophisticated thinking and argumentation practices. These answers showed that students sometimes considered counterarguments and rebuttals, and that they were able to engage in high-level data analysis.

Arguments with rebuttals. While many of the students' responses left things unsaid or indicated limited attempts to connect the evidence to the claim, there were two responses that indicated some students had sophisticated ideas about what constitutes an argument. In addition to using "because" to indicate that their evidence was in support of a claim, these examples also included language that showed the students were attempting to address counter-arguments. In the science investigation, Group 5 wrote that, "even though the blue monster is bigger and heavier and the people who are controling still won't win because its heavier" (G5SenS). In this response, the students are acknowledging a possible counter argument and rebut it with evidence. They do this by using "even though" and "still." "Even though," indicates that there is evidence that might contradict their conclusion that the Green Bolt is going to win. They go on to say that the Blue Monster "still won't win" and provide a supporting piece of evidence. The "still" signals a rebuttal and indicates that the students believe that counter-argument they acknowledge is not sufficient to overcome the evidence they have. A second group, Group 4, also included a counter-argument and rebuttal in one of their responses on the history investigation. They wrote, "it was mostly the demacrats because the demacrats might have annonced the election but the people voted but it was still mostly the democrats" (G4SenH). Although the language is somewhat confusing, the students still signaled that they were considering a counter-argument by using "might have" and that they had a rebuttal in mind with "but" and "still."

While these two responses demonstrate a sophisticated understanding of argument, they also illustrate the difficulty that students had articulating their ideas. Group 5's counter-evidence was accurate but did not actually support the counter-argument. They correctly stated that "the Blue

Monster is bigger and heavier," but this did not support the counter-argument that the Green Bolt will lose. It was the fact that the Blue Monster was pushed with greater force that suggested it might beat the Green Bolt. Given the difficulty of the task, the fact that students had not engaged in this kind of work before, and the relatively large number of texts, it might be expected that students would make errors when recording their responses. Additionally, as they were recording the response, Jason told his group mates that "They have stronger, stronger, stronger people, stronger people" (G5SD3:264), which suggests he knew the information which appropriately supported the counter-argument they were rebutting.

Group 4's response had appropriate details and language that signaled counter-arguments and rebuttals, but they seemed to be in the wrong order. They wrote, "it was mostly the demacrats because the demacrats might have annonced the election but the people voted but it was still mostly the democrats." By saying "the demacrats might have annonced the election" the students signal that this is a counter-argument, but it is in fact evidence supporting their claim that it was mostly because of the Democrats. It is actually the fact that "the people voted" that supports the counter-claim. The students may have been trying to say, "it was mostly the Democrats, because the people might have voted, but it was still the Democrats that announced the election." Unfortunately, I did not have permission to record the students in Group 4, so I cannot look at their group's conversation for an indication of their actual thinking on the subject

Analysis of source. Another of Group 4's responses in the science investigation demonstrated a surprisingly sophisticated analysis and interpretation of a source. The response was based on Source 1, Figure 22, and stated, "Because the Blue Monster will take about 32 seconds and the green 27 seconds."

Source 1 – Cart Weight Test

Headnote: To get ready for building their push-cart each team did tests with different weights of push-carts. They got 4 cars with different weights and then pushed them with the 20 newtons of force. They timed how long it took each push-cart to travel 100 meters

Push-Cart Weight	Time
80 kilograms	28 seconds
85 kilograms	29 seconds
90 kilograms	30 seconds
95 kilograms	31 seconds

Source: 4th *Grade Engineering Journal*, by Karl, Edgar, Ebony, and Asia. 2017

Investigation Question:

Which car will win the race? The Blue Monster or the Green Bolt?

What does this piece of evidence tell you about the investigation question?

.that 80 kilograms is faster ~~85~~
85 is the 2nd fastes.
30 is the 3rd Fastes.
31 is the 4th fastes.
95 Kilograms is Slower.

Figure 22: Group 4 – Science, Source 1.

At first glance, this answer does not appear connected to the source because the times the students provided were not on the table. Upon closer inspection, however, it seems that the students were extrapolating the carts' possible speeds based on the trend in the data they were given. The students in Group 4 saw that the weight of the carts in the table went up by five kilogram increments, and that for each five kilogram increase the time it took the carts to travel 100 meters went up by one second. They took this information and applied it to what they knew about the Green Bolt and Blue Monster. In the investigation, they were told that the Green Bolt weighed 75 kilograms and the Blue Monster weighed 100 kilograms. They then inferred that since the Green Bolt was five kilograms less than the lightest cart on the table, it would go one second faster, coming to the conclusion that "the green [would take] 27 seconds." They applied the same logic to the Blue Monster, which was five kilograms heavier than the heaviest cart in the source and concluded that “the Blue Monster will take about 32 seconds,” one second slower than the slowest cart in the table.

Interpreting data and sources is a key element to argumentation, since that forms the basis of one’s claims. Furthermore, to develop a claim one must often go beyond what is explicitly stated in sources of the data available. This response shows that, in this case, the students understood they needed to make inferences about the data and not just report what was stated in the sources. While the statements lack the reasoning explaining their conclusion (e.g., “the Blue Monster will take about 32 seconds *since it is 5 kg heavier than the biggest cart*”), it still demonstrates a careful reading of the source and a consideration that the answer to the claim may not be explicitly stated in the sources.

This answer also includes language that is important to high-quality arguments. Namely, the students use hedging language to indicate that their conclusion about the carts’ speeds is not absolutely accurate. Instead of saying “the Blue Monster will take 32 seconds,” they wrote, “the Blue Monster will take *about* 32 seconds.” This is important for two reasons. First, in terms of the argument it functions as a defense against counterarguments by acknowledging it is an inference rather than statement of absolute fact. Second, it indicates a good deal about the students’ thinking. Using “about” shows that the students themselves understand that this is only a possibility, and that they are arguing for a scenario that more likely rather than totally certain.

Students’ Use of Supports

There were a number of supports built into the design of the activities in the investigations and the materials the students used. In addition to modeling how to use the materials and create an argument at the beginning of the study, there were also features of the worksheets and SenseMaker board designed to help students construct complete arguments. First, the

SenseMaker board explicitly laid out the elements of an argument that were required. The column labeled "Why I think that" provided a space for students to justify their claim, and the "Evidence" column indicated they needed to support their claim with a source. Additionally, the SenseMaker boards limited the amount of writing that students needed to do, by enabling them to structure their arguments graphically. Second, the students had the source evaluation worksheets to guide the supporting evidence they provided on the SenseMaker board. Finally, including slips of paper for each source and slips with "AND" written on them had the potential to encourage them to use multiple sources. In this section I report how students used those supports.

Providing Justification

As the findings in the previous section (*Quality of Students' Arguments*) indicated, the SenseMaker boards were successful in supporting the students to provide some kind of justification for their claims. While there were a few instances where students simply repeated their claims, they also frequently included specific details and interpretive statements to back up their answers to the investigation questions. Modeling providing justification on SenseMaker board with the "Who stole the cookie from the cookie jar?" example, and including a space dedicated to "Why I think that" seems to have resulted in arguments that included justifications and not just restatements.

The SenseMaker board, however, was not completely successful in helping students make complete arguments. The initial intent was to have students provide their reasoning along the arrows connecting the source citation to their justification in the "Why I think that" column. Figure 23 provides an example of what that might have looked like in the science investigation.

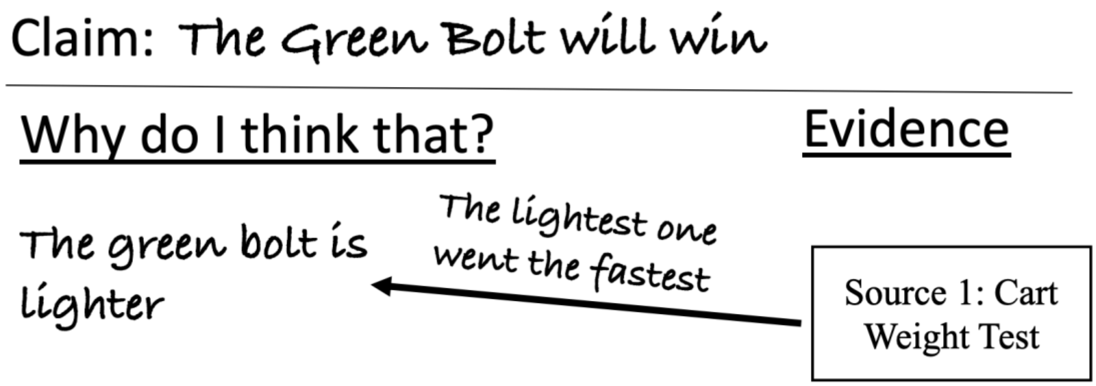


Figure 23: Example of Possible Reasoning on SenseMaker Board.

Although I modeled writing their reasoning along the arrows in my introduction and encouraged students to include it when they completed the boards in the history investigation, no groups included this. While unfortunate, this is not unsurprising. Reasoning is an aspect of argumentation that is typically difficult for students, even those who have had direct instruction in building arguments. These students had received no prior instruction on what "reasoning" was and had to rely on my brief introduction for their understanding of what to put along the arrows.

Using the Sources

One way the scaffolds provided by the materials and structure of the activity seemed to support students was in the use and prioritizing of evidence. The SenseMaker board included an area for students to provide citations for their reasons supporting their claims. They were given small slips of paper with each source number and title on them and were told to glue them in the right-hand column of the SenseMaker board. They then had to draw a line from the source title to the supporting claim they had written to indicate where they got the evidence from. All of the groups were able to successfully follow these instructions, and all of their supporting claims included a citation to a source. In addition, by making the citations central to the SenseMaker board, the students may have been influenced to foreground the importance of the sources and getting their ideas from the texts. Angela, Brandon, and Jason seemed to place great importance on carefully going through the sources and making sure that their ideas came directly from the sources. Since these students took such a leadership role during parts of the SenseMaker board activity, it is unclear whether or not the other students in the group shared these ideas about the importance of evidence.

Angela from Group 1 provides the clearest example of a student prioritizing getting evidence and justification from the sources themselves rather than just from background knowledge or the details of the investigation question. First, as described above in *Strategies for Developing Claims*, she seemed to come up with an answer to the science investigation question based on the evidence, suggesting she saw the texts as the appropriate basis of her claims. In addition, she recognized that it was important to use all of the sources in her argument. Almost as soon as they began the activity, Angela laid out all of the sources on the table for the members of her group to see, and then told the group "We're going in order" (G1SD3), meaning that they would go through the sources from 1 to 6 in order. This systematic approach to considering the sources indicates that she understood that they needed to consider all of the sources in their argument. Several times during their small group work, Angela also returned to the sources on her own and read them independently.

In addition, the evidence played an important role in the way that Angela interacted with the SenseMaker board. She repeatedly asked Solomon which source his supporting claim came from so she could provide an attribution in the "Evidence" column of the board, saying things like, "You have to find where the source—can you just glue this and give it back to me 'cause I wanna—," "It's because Solomon wrote here. I need to wait until that's [the source citation] there. Find what source it is," and "Then where did this come from? Where did this come from? [referring to a supporting claim written on the board]" (G1SD3: 198, 221, 366). She also worked directly from the sources to record her supporting claims. Prior to writing her two supporting claims independently on the SenseMaker board, she read from the source quietly to herself. Then, as she was writing, she held the source in her hand and repeatedly referenced it as she wrote.

Angela behaved similarly during the history investigation where she immediately prompted her peers to give “a reason why” they thought the vote occurred due to the Democratic party (G1HD3:12). She continued to press her group mates to give supporting evidence and return to sources for more information. Throughout their discussion, there were 10 instances when Angela made some sort of bid or prompt for evidence from the rest of the group. These bids generally fell into two categories: telling her group mates to support their claim or orchestrating the group’s use of the evidence. In addition to her first prompt telling Eric and Solomon to “give a reason why [it was both the Democrats and popular opinion]” she also asked, “Why do you think it’s both?” and “Anybody answer why it’s [both the Democratic party and popular opinion]” (G1H3:13; 292; 336).

Another set of her prompts for evidence were part of her attempts to orchestrate how the group worked their way through the evidence, so that they looked at each source. First, she asked the group, “Does anybody know what source they want [to read]?”, later repeated the guidance, “Okay, just put that paper right there, and then we all can read it. Each one reads one source. Anna, what source do you wanna read?” and then finally tried a more direct approach, “Source 2. You guys read source two and see if there’s anything that you wanna write about” (G1H3:138; 277; 365). Her other prompts for specific sources, such as “Can I see the source 3?” or “Let’s go to source 6” (G1H3:261; 502) were instructions to move from one source on to the next.

Angela's use of evidence in the task showed that she believed that it was important to back up claims and that the more support an argument had the better it would be. Her instruction to the group that, "We need to get our evidence out … Take the evidence out, so we know what the evidence are" (G1H3:57-60) combined with her numerous prompts for supporting evidence and reasoning showed that she held evidence in high regard. In addition, her attempt at a systematic approach to using the sources indicated that she wanted to have as much support for her argument as possible and did not want to miss any details that may have helped her strengthen their claim.

In Group 3, Brandon seemed to place similar importance on getting evidence directly from the sources and systematically going through each source to find evidence in support of their claim. He began the argument construction portion of the science investigation, for instance, by quickly going to the folder of sources and reading through them carefully. His close rereading of the texts was evidenced by his question, "Who wrote this?" (G3SD3:35) about one of the responses on a source evaluation sheet and his observation that "We don't have a source one" (G3SD3:40) after looking through the folder.

There were also instances in the history task that demonstrated the Brandon was sometimes thoughtful about not just using the sources for justifying the claims, but also making sure their claims were accurate to the source. Two such examples occurred while he and Taylor were examining source 2:

> Taylor: The Whigs voted against the Democratic Party. I'm gonna write, the Whigs were not happy, and they wanted—and there was a revote.
>
> Brandon: Okay. That was in the second—that was in the second convention.
>
> Taylor: I know.

Brandon: We're not on the second convention.

Taylor: Oh.

Brandon: At the first convention.

Taylor: Okay, then I should take that out, then.

Brandon: Yeah. (G3HD3:337-344)

Here, Brandon saw that Taylor was getting ahead of herself in the analysis of the documents. Source 2 (Figure 24) describes the community meetings that took place after the first convention. Taylor, however, offered a detail that was from Source 3, or as Brandon put it, "in the second convention." Having clearly read the source, Brandon pointed out the error and Taylor accepted his idea and removed her idea.

Source 2 – Biography of Michigan's Governor	Investigation Question:
Headnote: This comes from a biography of the first Governor of Michigan, Stevens T. Mason. The author describes what happened after the first convention voted for Michigan not to become a state because of the fight over Toledo.	Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?
After the first convention, it seemed like many people really wanted to have a new vote about becoming a state. Soon after, there were public meetings all through Michigan. These meetings were an example of democracy. People were meeting to decide what the State would do. In 1836 these citizens met in barns, schools, town halls, and village stores. They chose five men to get delegates for a new convention. They were successful. On December 14th another convention met and seventy-one people were there.	What does this piece of evidence tell you about the investigation question? ________________________________
Source: *Stephens Thomas Mason: Misunderstood Patriot*, by Kent Sagendorph, pages 242-243, Published in 1947.	

Figure 24: History Source 2.

Later, as Taylor was recording their revised response for source 2, Brandon was looking at the arrow linking the "Source 2" sticky note with the comment that Taylor was writing and said, "Wait. This is the—wait. No, no, no, no, no. Wait. No, that's correct. That's correct. Never mind. That's correct." (G3HD3:426). Although he did not articulate his thinking, he indicated that he thought Taylor had made a mistake in her attribution of the comment to the source. Again, this shows that Brandon wanted to be accurate in his argument and that he was looking to the texts as the source of support for their claims, as opposed to their own ideas.

Similar to Angela, Jason took a leadership role during the history investigation and was responsible for much of the group's work. Table 34 presents the group's written responses from the source evaluation task and their responses on the SenseMaker board. As it shows, the details Jason provided were largely consistent with the group's initial reading of the sources. This suggests that besides knowing justification was important, Jason knew it needed to come from the sources they had read.

Table 34: Group 5 – Source Evaluation Responses and SenseMaker Responses.

Source	Source Evaluation Response	SenseMaker Response	Consistent
4	Democrat's try to convince other's to become Michigan.	Because Democrats tried to convince others to vote for becoming a state	Yes
6	The govoner gave advice to democrat's to become a state.	In the letter the government gives advice to the democratics	Yes
5	teachers write a magazine called "michigan history" to pursude democratic's.	Explain: because of the Michigan State university when they make the article the[y] gave it to the public to get others to vote for them.	No
3	it tell us about how many people attended to become Michigan.	Because it only told us how many people attended (*Labelled "irrelevant")*	Yes

In addition, Jason also directly quoted from the sources when providing his reasoning on the SenseMaker board, suggesting that he knew the evidence should play a primary role in their argument. For the second supporting claim, “In the letter the government gives advice to the democratics,” Jason referred directly to the text of the source. He first read from the source, “In the letter, the governor gave advice to the Democrats,” then told Zahra, “Copy this whole entire letter until it goes all the way down—to like— … And this line. This one and this one” (G5HD3:162-169), and then underlined the text that she should copy onto the board. In the third response, Jason originally told Amira to write “The university of the article— … of the newspaper article” (G5HD3:320-321). After consulting the source text, however, he revised that to “Michigan State University,” directly quoting from the headnote to source 5.

Finally, Jason came very close to explicitly saying that support for their argument had to come directly from the sources. When looking through the source evaluation worksheets, he stopped at one source and mused aloud, " — and see where—let’s see. ’Cause this will probably be a claim" (G5HD3:84). While his use of “claim” is somewhat different than how it was being used on the SenseMaker board, he is essentially saying that the support for his answer needs to come directly from the source.

In contrast, Group 5 seemed much less intent on carefully using the sources in the science investigation. At the beginning of the activity, Zahra and Amira coordinated who would review which source, with Zahra saying, “Okay. You read it. You read this” and Amira telling Daniel, “You read that. I’ll do this” (G5SD3:51; 54). After this initial reading, however, the students made only one mention of finding or reading a particular source.

Citing Multiple Sources

Another way the design of the activity and materials supported students' argumentation was in the way the students were encouraged to use multiple pieces of evidence. All of the groups provided at least three, and as many as six, supporting reasons and citations for the claim. There were several elements of the activity that may have contributed to this. First, when modeling how to use the SenseMaker board with the "Who stole the cookie from the cookie jar?" example investigation, I modeled using multiple pieces of evidence to support my claim and reinforced this idea to the students.

Second, the SenseMaker boards came with multiple slips of paper with "AND" written on them. The students were to use these to link together multiple supporting reasons and citations. In a few instances the students made direct references to them. Aisha commented that, "There's so many 'AND,' 'AND,' 'AND, 'AND'" (G3SD3:171) when reviewing the materials. Angela looked at another group's work and said that, "They only wrote one 'AND'" (G1SD3:487). Jason noted that "We still have another 'AND'" (G5HD3:230) and proceeded to add another supporting claim to the SenseMaker board. This demonstrates that the students understood that using the "AND" slips and including multiple pieces of evidence were important to the task and a necessary feature of their arguments. Building this into the activity, then, may have influenced the students to include more pieces of evidence than they might have without those slips.

Including multiple pieces of evidence is generally a feature of high-quality arguments, but this evidence must also appropriately support the claim. While using the "AND" slips as a support may have encouraged students to use a greater quantity of evidence, it did not necessarily ensure that the evidence they cited was of high quality. In fact, as the above excerpts suggest, some of the students appeared to add evidence only because there were more "AND" slips to use. This support, then, does not help students understand what it means for evidence to adequately support a claim. As a first step to more sophisticated practice, however, the “AND” slips did appear to help students understand that a single piece of evidence was likely insufficient for a good argument.

Ignoring Prior Work

One support that was used inconsistently among the groups was the source evaluation worksheets. Ideally, the students would have been able to use what they had recorded about the evidence on the source evaluation worksheets when completing the SenseMaker board, and simply transfer some of their responses. This did not happen consistently across the groups, however, and some students seemed to redo the work of reading and evaluating the sources on the final day of the investigation.

Group 3, for example, did not appear to use the source evaluation sheets in the history investigation. Table 35 compares their responses on the source evaluation worksheets and the supporting claims they recorded on the SenseMaker board. As it shows, the students did not use their responses on the source evaluation sheets on the SenseMaker boards. The SenseMaker response “The whigs did not agre” seems to line up with the source evaluation response “The Wigs disigreade at the meting.” These responses, however, were based on different sources and each source does support the claim that the Whigs disagreed, showing that the students were likely not using the evidence sheet as a guide. In addition, the students not only left out support from sources 4 and 5 but also labeled them as “irrelevant” to their argument even though their reading of both sources included details that supported their answer of “the Democrats.”

Table 35: Group 3 – History Source Evaluation Responses and SenseMaker Responses.

Source Number	Source Evaluation Response	SenseMaker Response	Consistent
1	There were 21 people egestied [against] 28 people in the mater of all the votes.	The whigs voted agents [against] the democratic party	No
2	Many people wanted to change their votes.	The governor made another convention	No
3	The evidence was that there were 82 delegates and 82 to become a state. because were more to be a state.	The whigs did not Agre	No
4	Democratic Party convince michigan to become a state	Labeled “Irrelevant”	
5	The Wigs disigreade at the meting. It also telles us were it happened at. Not everyone agreed about having a second convention. On December 14th eighty-two delegatas met in Ann Arbor.	Labeled “Irrelevant”	
6	The Democrats because they tried there best to pswaed [persuade] them.	The governor was a Democrat	

Along with leaving out details from their source evaluation sheets that supported the Democratic party as the answer to the investigation, Aisha and Brandon did not use the evidence sheets to support their claim that both public opinion and the Democrats were responsible for the second vote. Two of their responses, “Many people wanted to change their votes” and “because were more to be a state,” supported the idea that public opinion played a role in the decision. Aisha and Brandon, however, did not include this in their justification of their claim.

Group 5’s use of the source evaluation worksheets during the history task was also inconsistent. Jason was responsible for most of the responses the group recorded. Although his responses were rooted in the sources, he did not actually look at the sources when coming up with the first and third responses. In both cases, Jason only looked at the sources after I prompted him to provide a citation for his idea. In addition, Jason did not appear to use what the group had recorded on the source evaluation sheets as a guide. For their second response Jason did use the source evaluation sheet, but instead of repeating what his group has recorded, he had Zahra record a direct quote from the source. On the other hand, although Jason did not work directly from the source evaluation sheets, his responses on the SenseMaker board were very close to what they had written about the sources on the two previous days. This suggests that even though he didn't use the tool directly, the activity of analyzing the sources did support his development of an argument.

Similarly, Group 1 made little reference to their source evaluation sheets when constructing their SenseMaker board for science. Although I adapted the activity to include an opportunity for the groups to review their source evaluation sheets and decide on an answer to the investigation, this group did not make much use of that time. They did not attempt to reach consensus on their claim until well after the activity had started. And, despite Angela's attempt to organize the activity, they did not go through the sources to look for the evidence they found compelling or refresh their memories.

All of this suggests that some of the students were using their memory of the texts in combination with a small amount of rereading, and thus reinterpreting, of the sources on the day of the argument construction. This may indicate that they did not see any connection between the two parts of the investigation. On the other hand, it may show that they simply did not realize how they were supposed to use the source evaluation sheet. If that were the case, more scaffolding demonstrating to students how to apply their evaluation of evidence to their argument may have been necessary.

In contrast with the other groups, Group 3 appeared to make more use of their source evaluation worksheets in the science investigation. After being prompted to take a moment to review their evidence before completing the SenseMaker board, Brandon picked up the folder of sources and began reviewing them independently as he did in the history investigation. Then, he tried to assign students to be responsible for reading different pieces of evidence, saying, “So, you read one, you read one—” (G3SD3:93). Taylor and Aisha followed his lead and Taylor said, “Okay,” and Aisha began reading aloud their written response on Source 1. Brandon noticed what Aisha was reading, and they began a discussion in which they debated what part of the sheets they should be looking at, either their own written responses or the original text of the source:

> Brandon: No, not that [students’ response]. This [original source text].
>
> Aisha: Oh.
>
> Aisha: No, we have to read what we wrote. He [Marino] said we have to look back on that.
>
> Taylor: [Reading students’ response] It tells us that if we race two cars and put—
>
> Yasmin: This part? [students’ response] Brandon, he said we have to read this part?
>
> Brandon: No. That part [original source text].
>
> Aisha: We have to read this part [students’ response].
>
> …
>
> Taylor: We’re reading our writing. (G3SD3:100-111)

While Brandon brought the group’s attention to the sources, Aisha and Taylor recognized that they also needed to be looking at what they had written about the sources, not just the original text. Aisha’s statement that “He said we have to look back on that” suggests that my instruction to review their responses at the beginning of the task was responsible for her attention to their previous work. The students in Group 3 did not pay similar attention to their prior responses in the history task, where they quoted directly from the source, rather than their own writing. This indicates that they needed to be explicitly told to use their prior work evaluating the sources when creating their argument.

Discipline-Specific Practices

As in the source evaluation portion of the investigation, there was little evidence that students were using the evidence to construct arguments in ways that reflected disciplinary practice. Although the students' arguments cited multiple sources, for instance, none of the students made an attempt to corroborate details across sources or even explicitly link details from one source to another. The closest the students appeared to come to applying disciplinary-specific strategies for developing and supporting arguments came in the science investigation. As described above, a number of students used the idea that "lighter things go faster than heavier things" as support for their conclusion that the Green Bolt would win. While somewhat simplistic, this amounts to a scientific principle that students were applying to an argument, and scientific principles are often considered part of the reasoning which can tie evidence to a claim. It was unclear, however, if students were using this idea as a way to warrant the use of their evidence or as evidence itself. As reported in the section titled *Providing Reasoning*, the students appeared to have difficulty expressing the reasoning which connected their evidence to their answer to the investigation question.

Early Skills in Argumentation

Even with the difficulties the students had with constructing arguments from evidence described above, there were several indications that they were either ready to begin to engage in the practice of argumentation or had some nascent abilities with the practice. Perhaps most significantly, their use of logical connectors in both investigations (e.g., "because") suggests that they understood that arguments needed to be justified, even if they had difficulty articulating other aspects of their arguments such as the reasoning. The students also appeared to readily understand the importance of basing their arguments on evidence and use multiple pieces of evidence to support their claims. In both investigations, there were students who pressed their group mates to base their support in the evidence found in the sources. As described above, for instance, Angela pressed her group for citations in the science investigation and Brandon made sure that his peers were reading the source accurately. Furthermore, all of the groups successfully incorporated multiple pieces of evidence into their arguments. While the design of the task and materials heavily supported the students to use multiple sources, none of the groups used all of the sources. They clearly had ideas that some sources supported their claim whereas others didn't. Instead of just mechanically adding one detail from each source they were mindful about their use of the sources even if their reading was not always accurate.

In addition to the above trends across groups that demonstrated the beginnings of argumentative practice, there were also indications from individual students that students at this age can meaningfully engage in the practice. Angela, for instance, consistently led her group to engage in substantive conversation about the investigations. Even without practice or support in this kind of discourse, Angela was able to do so independently. As a result of this leadership, her peers often responded appropriately, and were able to provide the evidence she prompted them for. Angela's contributions together with her peers' appropriate responses suggest that with more deliberate support they would be able to engage in more "real" discussions about the investigations and evidence.

Group 4's reasoning about the likely outcomes of the science investigation (described above) provide another example of work that was not necessarily representative of the class, but still indicates that students at this age are capable of sophisticated work in argumentation and evidence analysis. In developing a prediction of future outcomes extrapolated from analyzing patterns in data, these students demonstrated a disciplinary practice of science. So, they were not only able to evaluate and make inferences from evidence, but also able to do so in disciplinary appropriate ways.

Conclusion

This chapter described how students used the evidence they evaluated in the first part of the investigation in order to construct an argument supporting their answers to the investigation question. Figure 25 presents the landscape of findings related to the students' work constructing arguments. As it shows, the students had many of the same successes and challenges across the two investigations.

In both investigations, students used logical language, such as "because," to indicate they were providing justification for their claims answering the investigation questions. This was done without prompting or support from the materials. Additionally, several students appeared to understand the importance of using sources to provide evidence in support of their claim and were able to incorporate multiple pieces of evidence in their arguments.

They also struggled in ways that were similar across the two content areas and related to their work evaluating sources as reported in Chapter 5. First, there was relatively little substantive discussion between students about their ideas, and they mostly simply took turns sharing their ideas. Second, with one exception, students left out the reasoning that tied their evidence to their claim, even though a number of their responses included "because," as a way to signal the evidence was supporting a claim. Finally, many of the students' responses were the result of literal readings and consisted of simple restatements of details and did not include inferential or evaluative readings of the sources. This is presented as a slight challenge in the figure because of the lack of reasoning that tied the details to the source, so the link between the detail and the conclusion was not necessarily clear.

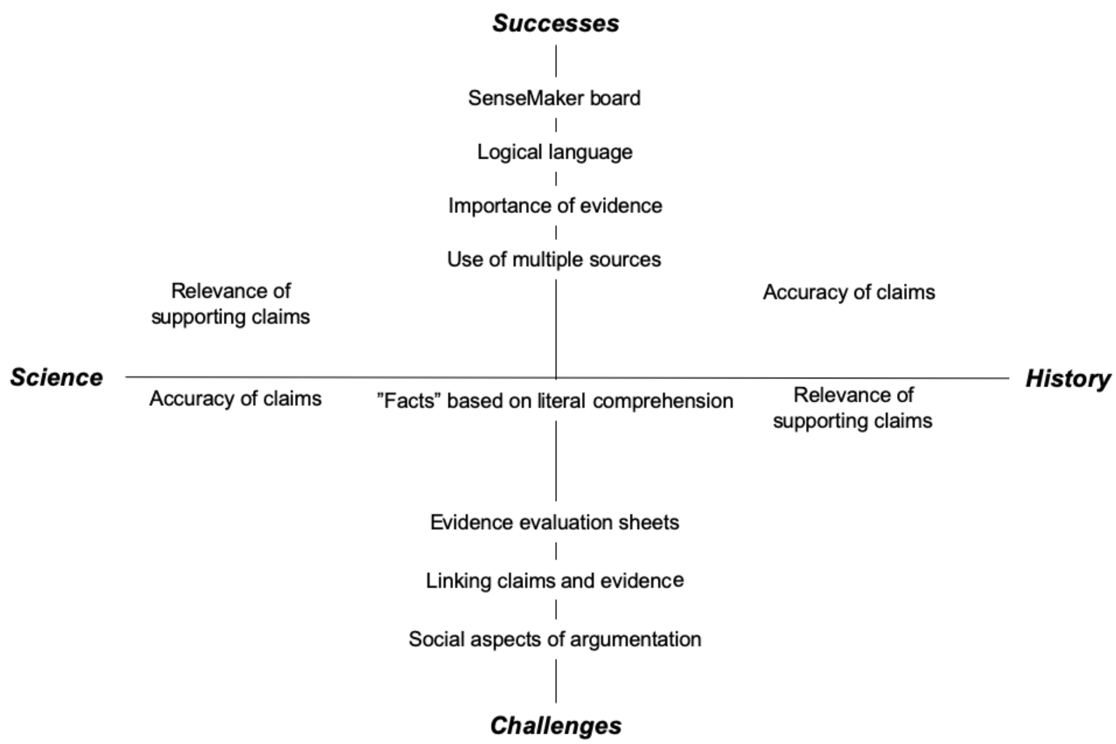


Figure 25: Landscape of Findings Related to Argument Construction.

The students also used the supports built into the materials in similar ways across the investigations. The design of the SenseMaker board supported all of the groups to include some kind of justification for their claim and provide multiple pieces of evidence. The source evaluation sheets, however, did not ensure that students relied on their prior reading of evidence to support their claim or that the evidence they provided was appropriate to their claim.

The main difference between the science and history investigations was in the quality of students' arguments, which varied. The accuracy of their supporting claims, for instance, was higher in the history investigation (71%) than in the science investigation (50%). The degree to which the evidence students provided supported their claim also varied. In the history investigation, only 50% of their responses reasonably supported their claim, compared to 63% in the science investigation.

Besides just identifying the challenges younger students new to argumentation may face, this chapter also demonstrated the potential students have in this practice. Despite the struggles described above, there were a number of indications that these students were likely ready to engage in argumentation and would have been generally successful with more preparation. In the following chapter, I discuss the ideas students had about the disciplines in comparison with their own work after the investigations.

CHAPTER 7

Students' Beliefs about Their Work in Relation to the Disciplines

In this chapter I present findings related to students' ideas about the investigations they completed, their relationship to the disciplines of science and history, and the relationship between the two tasks across content areas. Knowing about students' ideas about these relationships may help us understand differences in their approaches to argumentation in each field. This analysis is based on the post-survey students completed in their small groups after completing both investigations. The post survey asked the following questions:

1. Think about the investigation you did in **Social Studies** how was it similar and different from what **historians** do?
    A. What **we** did:
    B. What we **both** do: C: What **historians** do:
2. Think about the investigation you did in **Science** how was it similar and different from what **scientists** do?
    A. What **we** did:
    B. What we **both** do: C: What **scientists** do:
3. Think about the investigations you did in science and social studies. Were they the same in any ways? Were they different in any ways?
    A. Ways our social studies and science investigations were the **SAME:**
    B. Ways our social studies and science investigations were **DIFFERENT:**

The findings in this chapter are also based on the video recordings of their small group work, and a focus group interview in which I followed up on students' responses to the post-survey. Based on this analysis, I make five major assertions:

A) Students expressed their ideas in inconsistent ways, making analysis of this kind of data challenging.

B) "Research, learning, and thinking" were characteristics the students saw in common between their own work and the work of scientists and historians.

C) Students recognized that "experimentation and data collection" are important aspects of scientific practice.

D) The students appeared to have inconsistent ideas about the role of evidence when comparing students to practitioners.

E) Students appeared to have a stronger conception of scientific practice than historical practice, consistent with their ideas prior to the investigations.

Analytical Challenges

The groups' responses to the post-survey presented a number of challenges that made analysis difficult. In general, the groups' responses were inconsistent both within and across groups. The number of responses to each question, for instance, varied widely. As Table 36 shows, the total number of groups' responses on these surveys ranged from 13 to 25. In addition to this large range, there was also great variability in the number of responses given for each question. For example, in the questions about science, Group 1 gave eight answers whereas Group 5 only recorded three. This variability also occurred within some of the groups, as in Group 1 who gave 12 responses comparing their work with historians but only 5 responses when comparing the investigations to each other. The reason for this disparity in the number of responses is unclear. There did not appear to be patterns, for instance, related to the order of the questions or the difference in disciplines that accounted for the wide range of students' responses. I did not see similar disparities in the students' responses to the pre-survey.

Table 36: Post-Survey: Number of Responses per Group.

	Group 1	Group 3	Group 4	Group 5
History	*12*	*4*	*4*	*6*
Kids	2	2	1	4
Both	6	1	2	1
Historians	4	1	1	1
Science	*8*	*5*	*7*	*3*
Kids	0	2	1	1
Both	4	2	3	1
Scientists	4	1	2	1
Investigations	*5*	*4*	*10*	*5*
Similar	4	2	5	3
Different	1	2	5	2
TOTAL	25	13	20	14

Because of this range of numbers of responses to the questions, it is difficult to make generalizations about the students' ideas about the activities in the investigations and in the disciplines, and in particular, how they prioritized different themes. Sometimes when a theme such as *research/learning* was present in a large number of responses, many of those responses were the responsibility of one or two groups (repeating related ideas) and not the whole class. In cases such as this, the large number of responses addressing *research/learning* does not necessarily reflect consensus among all the students. As a result, frequency counts are unhelpful in identifying important themes across groups.

Another difficulty in analyzing students' beliefs expressed in the post-survey is the inconsistency in their responses. Group 4, for instance, wrote that "We use evadanse [evidence]" (G4Post) when describing what they did in both the history and science investigations. In comparing their own work with the work of practitioners, however, they only mentioned evidence in comparison to historians, but not to scientists. If they believed that they used evidence in both science and history investigations, it should follow that using evidence would be relevant to comparing their own work with that of scientists. This is also inconsistent with their responses on the pre-survey. In response to the question, “What do scientists use as evidence?” Group 4 provided 16 responses. While not every answer was an accurate example of evidence, the number of responses indicates that the students had some idea that scientists used evidence. Taking their responses on the pre- and post-survey together, it seems likely to me that if asked directly whether or not using evidence was important to both scientists and students, Group 4 would say it was. For some reason, however, they did not articulate this idea in the post-survey. Because of inconsistencies such as this, it is difficult to determine if students deliberately left out ideas because they actually believed they were not relevant or if they left them out for some other reason, such as getting distracted or being fatigued from answering so many questions.

Confusion about the Question Comparing the Investigations

In addition to the difficulties just described, there were also indications that the students were confused about the questions comparing the investigations with each other (questions 3a and 3b). It appeared that students may have been comparing their work in the investigations with the work of practitioners in the field, rather than comparing the investigations with one another. One indication of this is that the majority of the students’ responses were in the present tense, rather than the past tense. In fact, only six of the 25 responses were written in the past tense. Using the past tense showed that the students were considering events in the past and comparing them with one another, as in, “We ansered the queshton” (G3Post). Answers like, “we look back at our reasearch” (G1Post) in the present tense imply that the difference is ongoing, thus suggesting the students were comparing themselves to scientists and historians as opposed to comparing to discrete events in the past.

The students also frequently used “we both…” when answering “Ways our social studies and science investigations were similar.” Six of the 14 answers to this question started with “We both” as in, “We both read a lot” (G4Post). As opposed to “in both [investigations]” or simply “we”, “we both” suggests that the students are comparing themselves to another person, rather than making a comparison between themselves at two different times.

Finally, the content of some of the responses indicated the students may not have been comparing the investigations to one another as I had hoped. “Sciencetest [Scientists] do we didn’t make eny eplotion [any explosion]” (G1Post), for instance, clearly compares what the students did in the investigation with what they think scientists do. Group 5 made a similar comparison when they wrote, “science investagations and history investagations science they go exploreing and history investagations they don’t.” (G5Post). Here, the students were comparing investigations in science and history in general, rather than the specific investigations they completed in class.

There are a few possible explanations for this confusion. First, the students may have simply had a difficult time articulating their thinking, as they appeared to do in other parts of the study. When they wrote answers in present tense, for instance, there is the possibility that they did not realize the important difference between writing in the past and present tense. There is also the fact that this was the last question on the survey. Prior to answering these two questions, they had been asked to compare what they did in the science investigation with what scientists do, and what they did in the history investigation with what historians do. It is very possible that they still had those questions in mind when answering the final set of questions and misinterpreted the prompt.

Addressing Analytical Challenges

I used three main strategies to address the analytical challenges explained above. First, I did not rely solely on frequency counts to identify the trends in the students' beliefs about their own work, the work of practitioners, and the investigations. As I described previously, high frequency counts were often the result of a single group giving a larger number of responses related to a theme. To mitigate this bias, I also took into account the number of groups that shared an idea. Responses coded as *experimentation and data collection* provide a good example. There was a total of 12 responses expressing this idea, but half of them were provided by Group 1. Instead of using the frequency count, then, I looked at how many groups gave a response related to this theme. In this case, four groups provided answers related to experimentation and data collection. I then used the fact that all of the groups had at least one answer expressing this theme to infer this was something important to many students, and not the total number of responses.

I did, however, use frequency counts to identify trends within groups, which allowed me to create profiles of some groups and what they prioritized in these comparisons. In the above example, for instance, since Group 1 provided so many responses related to experimentation and data collection, I was able to see that this was something important to this set of students. While it was possible to identify trends within some of the groups, it was not possible for all of the groups. Some groups were simply less likely to be repetitive in their response than others. Group 4, for instance, only had one instance where more than two responses two a question repeated the same theme. In general, they had fewer total responses and their responses represented a wider range of themes. Thus, it was not possible to infer which themes they felt were more important than others.

My second strategy of dealing with the inconsistencies in students' responses was to compare students' ideas about their own activity during the investigation to practitioners in general, and not specifically to historians and scientists. In other words, in most cases I did not attempt to determine how disciplinarily-specific their responses were. I did this because of the many instances where students' responses were not consistent across questions, such as the example above where Group 4 did not include anything about evidence in their response to the questions about historians. The one exception to this was examining the students' ideas about experimentation and data collection. In that case, the ideas the students were expressing were necessarily limited to scientific practice, so it was possible to look at that theme in terms of scientists in particular as opposed to practitioners in general.

Finally, I did not analyze the students' answers to questions 3a and 3b to characterize their ideas comparing the two investigations. As I described above, there were many indications that students misinterpreted the question, so analyzing these responses did not seem appropriate. I did, however, use the responses in identifying the within-group trends, extending the analysis of the previous two questions. Group 5 for instance, included responses related to evidence when comparing the two investigations. While it was unclear whether or not they were answering the question appropriately, these responses did still suggest that evidence was something the students believed was important in general.

Comparing Investigations and Disciplines

When comparing the work that they did in the investigation with what practitioners of each discipline do in the field, the students responded that they had slightly more in common with practitioners than not. Out of 48 responses, the students provided 20 responses for “What we both (i.e., students and practitioners) do,” compared to 13 for “What we did” and 15 for what scientists/historians do." Although this is a strong indication that students saw investigations as generally similar to the disciplines, their ideas about specific elements of practice were less clear.

There were three aspects of disciplinary practice and the students’ work in the investigation where students’ ideas were clear. First, there was a strong trend among the students that research, learning, and thinking were common to the work of historians and scientists as well as students. Second, they differentiated between learning and discovery, and saw the latter as being the work of practitioners and not students. Third, the clearest difference students articulated between practitioners in the disciplines and students was that scientists engage in experimentation and specialized data collection.

The students’ ideas about the remaining themes and where they fell on the spectrum between “only kids” and “only historians/scientists” was much less clear. This was partly because of the inconsistencies in the students’ responses and the range in the number of responses, as described above. *Evidence,* however, was one theme that received a large number of responses across groups, suggesting it was something students felt was important. The ideas about evidence, however, were varied.

Research, Learning, and Thinking

When comparing the work of practitioners with students, the most commonly occurring theme was that of “research, learning, and thinking.” Of the 48 total responses comparing the work of scientists and historians with students, 14 addressed the research, learning, or thinking involved in doing the work of the disciplines. In general, the groups saw this as something that was in common between practitioners and students, and three groups cited it as a similarity.

While the greatest number of responses describing similarities were coded as "research/learning," the majority of these answers actually came from Group 1, which was responsible for 6 of the 9 responses naming research, learning, and thinking as a similarity between practitioners and students. This relatively large number of responses does two things. First, it suggests that research, learning, and thinking were things the students in Group 1 thought were particularly important to the work of practitioners and their own activity during the investigations. Second, it provides insight into how the students conceptualized these themes in the different disciplines. They seemed to have different ideas about historical and scientific practice, and how learning is done in each field. In history, they emphasized the research aspect and said, "we write down what we see, find, and research," and "we look back at research" (G1Post) as things that both students and historians do. In science, on the other hand, they focused on data analysis, saying, "we compare things to find stuff that are similar" and "we use other ideas" (G1Post). Rather than a focus on research, these responses suggest that in science they were paying more attention to the methods of developing a claim and argument.

The fact that Group 1 had the largest number of responses with this theme does not necessarily mean that it was less important to Groups 3 and 4. In Group 4's case, their answers were much more varied, and in general they were less likely to provide answers that repeated themes in the same way that Group 1 did. Thus, the relatively low number of responses indicating research, learning, and thinking may only be a reflection of Group 4's tendency towards brevity. Additionally, looking at their responses across the questions suggests they saw thinking as something especially important. When comparing practitioners to students they wrote, "We [students and historians] thing about what would happen" and "We [students and scientists] both think about what whould [would] happen" (G4Post). Similarly, when comparing the two investigations, they wrote, “We think about what would happen” and “We both think about what would happen” (G4Post). The consistency of the phrasing indicates that they saw this as an especially important aspect of work in science and history, and that it was something that students did as well as practitioners.

As with Groups 1 and 4, Group 3 also saw research, learning, and thinking as a similarity between historians and students, but their perspective was unique. They wrote that “They [students and historians] talk about both sides of the queshtin [question]” (G3Post), which shows that they saw historical events as having multiple interpretations. Their use of “both sides” may indicate that they see historical events as only having two sides, but they may also be referring to the two “sides” of the investigation question, the Whigs and the Democrats. In either case, the response does show that they picked up on the importance of looking at an event from different perspectives, an essential aspect of historical practice.

While Group 3 reported that analyzing multiple perspectives was common to students and historians, they also seemed to think that there were differences related to research, learning, and thinking between practitioners and students. They said that only students “investigate” answering that, “We [and not historians] did investigations” and “we [and not scientists] investigate” (G3Post). Here, it seems that they were influenced by the vocabulary of the activities, which were referred to in the survey as “investigations.” The other difference they saw between themselves and practitioners is that “Historians [but not students] do research on the past” (G3Post). On one hand, it shows that they understand what kind of work historians do. But on the other, it raises the question of what the students thought they were doing in the investigation. They may have seen it as something different from “research,” or perhaps were making some distinction between “the past” that historians study and what they as students were learning about. Unfortunately, however, neither the audio recordings of the group's discussion, nor focus group interview shed light on their response.

Discovery

In contrast with their ideas about research, learning, and thinking, one difference that some of the groups saw between practitioners and students was that practitioners were involved in making discoveries and generating new knowledge, whereas students were not. This was present in responses from three of the four groups. Responses such as, "they [historians] make more discovery" (G1Post), "They [historians] find out about lost or unknown history about the Presidents from a long time ago" (G5Post), and "they [scientists] creat new things" (G2Post) suggest that the students believed that historians and scientists found out new things about the world. These responses are different than the ones coded as "research/learning" because in that case, the implication is that what is being learned is already established knowledge. The students, then, may have seen the role of practitioners as the generators of new knowledge and different from their roles as students, where their primary goal was to learn.

This was an idea that seemed important to Group 1. In addition to giving three responses related to this theme, it also came up in their discussions. They said that only historians "make more discovery" whereas students "don't have a inport [important] job" and "discovered a little ideas" (G1Post). These answers seemed to have been recorded by Solomon, who during the group work told me that, "We're mini historians … We only discovered a little stuff" (G4Post: 268; 273). Here, he seems to see that the students were engaging in similar work as historians, but that historians do more consequential work compared to what the students were learning.

Discovery also came up during the interview. Two of the students in the interview expressed the idea that finding new things was part of scientists' work. When asked, "What do you think a scientist is?” Anna said that, “They discover new things that can make chemical reactions, and they kinda discovered dinosaurs,” and Jason later added, “…and they discover new species” (IntPost: 213; 230).

Experimentation and Data Collection

While the inconsistencies in the students’ responses made it generally difficult to parse how disciplinarily specific the students’ ideas about practitioners in the discipline were (i.e., the extent to which their ideas about historians were specific to historians and did not apply to scientists), there was one clear theme that applied to only one discipline. A number of the students’ responses made the distinction between scientists and students by naming ways that scientists engage in data collection or experimentation. All four of the groups gave at least one response related to data collection or experimentation. As with other themes, one group, Group 1 in this case, was responsible for the majority of the responses (5 out of 8), but the idea was present in four groups’ answers.

Some of these answers focused on experimentation, and chemistry in particular. These answers included, “make cemacel reacshon [chemical reaction]” (G1Post), “expiriment” (G3Post), and “They use cemiclels [chemicals]” (G4Post). As in the pre-survey, students seem to have a stronger vision of chemistry as a scientific practice. This was also reflected in the focus-group interview where Anna defined scientists as, “They discover new things that can make chemical reactions, and they kinda discovered dinosaurs” (IntPost:213). Taylor said that scientists, “Use a room to work in” and clarified that she meant “A science lab” that had chemicals in it (IntPost:215-219). Another group, as discussed later, pointed out that scientists collect data in different ways, writing, “They go on aventures to collect other data about animals and others” (G5Post). Here, they identify field work as something specific to scientists. Finally, Group 1 continued the trend of associating the historical sciences with history. They not only wrote that scientists but not students "put bones together," but also responded that historians but not students "make skeletons" (G1Post). As in the pre-survey where Group 1 wrote that historians do their jobs "by digging" and "comparing other animals" and use evidence "to know about other animals/fossils," they again seem to be confusing a historical science with history.

Using Sources and Evidence

Given the importance of sources and evidence in the investigations and the prominence of that vocabulary in the materials, it might be expected that they would appear frequently in the students' responses comparing the investigations with the disciplines. This, however, was only partly true. While 3 groups did agree that the use of sources and evidence overlapped with the work of practitioners, their responses were somewhat inconsistent. Groups 1 and 5, for instance, listed something like "use sarces [sources]" (G1Post) or "we both [historians and students] collect evidence" (G5Post) as something students had in common with both scientists and historians. Group 4, on the other hand, only listed "We use evadance" (G4Post) as something that was shared between students and historians and did not mention evidence at all in their comparison of students with scientists. This is inconsistent, however, with an answer they provided about the investigations. There they said that "We use evadanse" (G4Post) was something that happened in both the science and history investigations. It is uncertain why Group 4 left out any mention of evidence in comparing scientists with students, but given their later response about the investigations, this does not necessarily mean they did not see sources and evidence as something that was unimportant to either students or scientists.

Another complicating factor in the students' beliefs about sources, evidence, and scientists and historians is that half of the responses related to this theme came from Group 5 (5 out of 10 responses). In these responses, the students provided contradictory answers, in particular when comparing historians with students. At the same time, they wrote that "we both [historians and students] collect evidence" (G5Post), they also said that only students "We solve by using our data," "We read to collect data," and "We collect evidence" (G5Post). Because the students did not discuss their responses, the reasoning behind these answers is unclear. Their answers comparing scientists to students, however, had a much clearer distinction between disciplinary practice in the "real world" and in school. First, they acknowledged that, "We both [scientists and students] read (collect)[12] for data" (G5Post). Then, they went on to say that, "They [scientists and not students] go on aventures to collect other data about animals and others" (G5Post). Here the students make a clear contrast between students and scientists, namely in the way that scientists collect data in comparison with students. While students and scientists both read to collect evidence, typically only scientists go out in the field, or on "aventures," to gather new data.

Despite the inconsistencies in Group 5's answers related to evidence, their responses do suggest two points. First, their responses comparing scientists and students show that the students in Group 5 saw a particular difference between scientists and students. Their ability to name a specific practice (i.e., collecting data in the field) for scientists contrasts with their responses in history. There, they did not name a specific way that historians collect evidence for their claims. There are two possible explanations. First, they may not have believed there was any particular difference in the ways that historians and students collect evidence. The other, and I believe more likely, explanation is that they were less familiar with the ways that historians

[12] This parenthetical was written by the students.

collect evidence and thus could not name them. This is supported by their responses that identify generic ways of collecting evidence in history (e.g., "collect evidence") and the relatively stronger understanding of scientific practice they displayed in the pre-survey.

Secondly, while Group 5 may have had inconsistent ideas about evidence and its use by students, historians, and scientists, the number of responses indicate that it was something they found important. This may be related to the repeated use of "source" and "evidence" during the investigation. Looking at these responses in comparison to their responses on the pre-survey also suggest this. Their responses there suggest they did not hold evidence in as high regard. First, when discussing historians, they referred to historians finding "information" but not "evidence," "data," or "sources." Additionally, they were only able to provide three reasons why scientists and historians use evidence. None of those responses suggested that evidence can be used to support claims or develop explanations. The response that came closest to this was, "To find out about it" (G5Pre), which appeared to focus on the learning or research that might come with looking at sources of information.

Although the students' responses about historical evidence in the pre-survey were limited, and they did not go into detail about evidence in the post-survey responses or discussions, there was one notable exchange during the post-survey interview that showed that at least three students had more robust ideas about historical evidence than the pre- and post-surveys suggested. I asked Taylor whether or not historians use the same kinds evidence as the students used in the investigations and she responded:

> Taylor: I think that it might be different because it was in the middle of the actual time. They wouldn't know how it would actually turn out.

> Marino: Oh, so the people back then, it was in the middle of the time for the people back then, so they didn't know how it would turn out? What were you gonna say, Anna?
>
> Anna: It's the same 'cause we all talk about it, and then they can—but instead, they get photos and things like that, instead of just paper and just reading it.
>
> Marino: A historian might look at photos? Would they look at any of the kinds of things that we looked at?
>
> Anna: Yeah.
>
> Marino: Like what?
>
> Anna: Like how it said that—how they wrote it down, and how the people were saying it.
>
> Marino: Which people were saying it?
>
> Anna: An example was the go kart, when—I don't remember his name, but he said that his car has less curves, and then that Asia's has more.
>
> Marino: Edgar? Oh, that was Edgar.
>
> Marino: A historian might look at what people say? What are you thinking, Lillian?
>
> Lillian: They would look back in time and see what people would say and look at the evidence they wrote. (IntPost:147-159)

Here, Taylor hints at the idea of historians using primary contemporaneous sources by saying they would use something that "was in the middle of the actual time." This seemed to spark Anna and Lillian to think about sources that might be collected from the time of the event in question. Anna suggested photographs and first-hand accounts, both valid forms of primary sources. Similarly, Lillian's response calling out "the evidence they wrote" names first-hand accounts as the kind of evidence that a historian might use.

Whether or not the students' ideas about primary sources were influenced by the investigations is unclear. On one hand, the students said that primary sources were something different used by historians. On the other, Anna recognized and named a primary source used in the history investigation. The source she named, however, most closely resembles the kind of source they were talking about, which was a record of what people said. Other primary sources in the investigations such as the tables or newspaper articles are less clearly "what people said," and so may have gone unnoticed by the students.

Regardless of the role of the investigation in the students' ideas about historical evidence, this response suggests that the students' ideas about this kind of evidence may have been more sophisticated than the pre- and post-surveys indicated. While Anna and Lillian did not independently offer examples of primary sources, the fact that they so readily agreed with Taylor and were able to immediately give examples suggests that they may have had these ideas already.

Connections to Pre-Survey

In addition to revealing students' ideas about the investigations and their relationship to disciplinary practice, the post-surveys also had some connections to the findings from the pre-survey. Namely, the difference between their understanding of scientific and historical practice appeared to persist.

Ideas About Historians

Some students continued to be unclear about the work of historians, which was the case in the pre-survey, as well. While their written responses for the work of historians in the post-survey were all appropriate to the discipline (unlike before the investigation), there were exchanges in the group work that demonstrated some students were confused. Yasmin in Group

3 asked, “What do historians do? What does historians do?” (G3Post:34) and then repeated the question three times, eventually telling Brandon that “I have no excellent clue” (G3Post:45) and later saying, “Historians do—I don’t know” (G3Post:58). Similarly, in Group 1, Eric and Anna were unsure about the work of historians, asking, “My question is what is an historian?” (G1Post:181) and “What are historians? That’s my question” (G1Post:253) respectively. These questions sparked a discussion among the students as to what historians did:

> Marino: Eric, you had a question. You keep asking what are historians. What do the three of you—what do the three of you think historians are?
>
> Eric: I don’t know. People?
>
> Anna: People who discover history?
>
> Solomon: Like us.
>
> Eric: Museumists?
>
> Marino: Museumists? What do you mean museumists?
>
> Solomon: People like us.
>
> Eric: I don’t know what those are. I’m guessing.
>
> Solomon: People like us. People like us.
>
> Marino: People like you?
>
> Solomon: We’re mini historians.
>
> Marino: You’re mini historians?
>
> …
>
> Solomon: We need to find out about history. … Different? Oh, I know. We discovered only a little stuff. (G1Post:258-273)

Anna and Eric's initial expressed confusion gives way, here, to a vague understanding of what the work of historians is. Anna understands that their work has to do with history, but still seems unsure of what that entails. Eric's connection to museums was unexpected but makes sense. Museums display artifacts from the past and may be students' primary encounter with the study of the past given that there are relatively few representations of historical practice in popular culture, and the social studies curriculum prior to 3rd grade focuses on civics and community rather than history. Meanwhile, Solomon's assertion that they were "mini-historians" suggests that he saw significant connections between the work of historians and what they had just done in the history investigation. By using the qualifier "mini" he is also indicating that he understands there are important differences between what they did and what historians do. This may simply be a matter of acknowledging the difference in age, but he also shows that he knows that historians do more involved or consequential work. As opposed to the students, Solomon sees historians as discovering "big stuff," which could mean events that are more important or previously unknown. This was also reflected in his group's responses. They wrote that historians "make more discovery" and "ancover more mystrys [uncover more mysteries]" and that students, in contrast, only "discovered little ideas" (G1Post)

During the interview some of the students also demonstrated confusion about what historians do. When asked what a historian was, Jason responded, "I kinda be a historian sometimes 'cause I look up on the Internet about how tornados and stuff, how they form. I like to collect history about them" (IntPost:47). Here, Jason seems to have conflated any kind of research with history. Later, he said that a topic that historians mostly study is "Well, mostly they're [historians' topics] animals" (IntPost:61). This was consistent with Jason's ideas during the pre-survey, where his group listed "animals" as an example of evidence used by historians and "they study animals to get information" as one way that historians use information. Although this suggests Jason does not know what historians do, there is the possibility that he was conflating historical sciences, namely paleontology, with history. Later in the interview, the subject of digging for dinosaurs came up, and Anna said that "We don't get to dig out dinosaurs" (IntPost:409). Jason agreed, and when I asked if that was something done by scientists or historians do, he said "Maybe a scientist and a historian" (IntPost:416), indicating he thought it was part of both of their practice, or possibly someone who was both a scientist *and* a historian.

Of the students in the interview, Taylor was the only one who expressed an accurate representation of what historians do. She said that, "I think a historian is something that looks up information or figures out information, but in a group." (IntPost:54). When asked what kind of information they look for, she responded "History" and later elaborated "Basically just history, the topic is history" (IntPost:56; 65). Taylor's understanding of historical practice is that is collaborative, research-based and examines the past. Her ideas about the topics of history, however, seem to be fairly generic and limited to "history" without elaborating on what kinds of events or specific topics that might entail.

Ideas About Scientists

The students also continued to demonstrate a stronger understanding of scientific practice. This was evident in the responses which detailed experimentation or data collection techniques. Answers like "they use cemiclels [chemicals]" (G4Post), "expiriment" (G3Post) "make cemacel reacshon [chemical reaction]" and "put bones together" (G1Post) show that many students had knowledge of the specific ways in which scientists do their work, as opposed to their generic idea that historians study history. This was apparent in the focus group interview where students were able to name particular topics that scientists study or elements of their practice as happened in the following exchange:

> Marino: Before we start talking about it, what do you think that a scientist is? Yes, Anna.
>
> Anna: They discover new things that can make chemical reactions, and they kinda discovered dinosaurs.
>
> Marino: What do you think, Taylor?
>
> Taylor: I think that they use a room to work in.
>
> Marino: What's special about that room?
>
> Taylor: A science lab.
>
> Marino: It's a science lab? What might be in a science lab?
>
> Taylor: Chemicals.
>
> Marino: Chemicals. Anything else?
>
> Taylor: Tables.
>
> Marino: Tables. What are you thinking, Lillian?
>
> Lillian: That they help the people tell the weather.
>
> Marino: Scientists help people tell the weather?
>
> Jason: Uh-huh, so like weather balloons and stuff. (IntPost:212-226)

Here, the students were able to list specific topics, such as paleontology and weather, that scientists might investigate. They were also able to name appropriate tools, like a laboratory or weather balloons, that are used in scientific practice. Furthermore, Taylor was able to generate an idea of work a scientist might do based on the science investigation. She suggested they might "study stuff" such as "Maybe study air resistance before they try and make their own car?" (IntPost:255-257).

Conclusion

This chapter described students' responses to the post-survey and their ideas about the differences and similarities between students, scientists, and historians. The students' responses were largely inconsistent, and the data collected proved difficult to analyze, suggesting that using surveys with such a small sample size may be problematic. It was still possible, however, to identify a few trends in the students' beliefs.

In general, the students saw that research, learning, and thinking were things that were in common between students and practitioners in the disciplines. This was the most commonly cited similarity and the idea was shared by a number of groups. Their knowledge of experimental and data collection techniques in science was evident, and they saw it as a major difference between students and scientists. The students seemed to think that evidence was an important theme, but their ideas about it were varied. Finally, some of their responses were consistent with my previous finding that they had more robust ideas about scientific practice compared to historical practice.

In the following chapter I discuss the implications of the findings presented here and in the previous findings chapters, as well as the findings' connection to what is already known about students' argumentation in science and history and the limitations of the study.

CHAPTER 8

Discussion and Implications

In this chapter I discuss the findings I have presented in light of the literature, and the findings' implications for research and instruction. This chapter is organized into three main sections. The first section synthesizes the landscape of findings from the pre- and post-surveys and history and science investigations, and what it tells us about students' argumentation and use of evidence. In the second section, I discuss the methodological implications of the study and revisions for future work. Finally, I explore the instructional implications and suggest ways this study's findings can support teachers in the field.

Throughout this chapter, it is important to keep in mind that the students in this study were not only relatively young, but also completely new to the practice of argumentation. This lack of experience and preparation meant that their argumentation typically lacks the traits of high-quality argumentation in science and history. Consequently, it is easy to describe the students' practices in terms of what they did *not* do as opposed to what they did. I believe that it is important to avoid this deficit perspective, however, for two reasons. First, it discounts the impressive work that many of the students were able to do considering the sophisticated nature of the task and their unfamiliarity with both the process and the content. Second, it obscures one of the main goals of this study, which was to gain a better understanding of the kinds of knowledge and skills that students bring to the table when they are just learning argumentation. An important thread in this discussion, therefore, will be on what the students were able to accomplish and what that tells us about how students begin to engage in argumentation across disciplinary lines. And, as the sections below explain, the students were able to demonstrate considerable abilities with evidence analysis and argumentation given their age and lack of prior school-based experiences with the practices.

Landscape of Findings

Ultimately, the central goal of this study was to compare how novices to argumentation engage in the practice across disciplines. To do this, I return to the figures presented at the end of Chapters 4, 5, and 6 and present a final figure overlaying my findings related to how students evaluate sources and construct arguments.

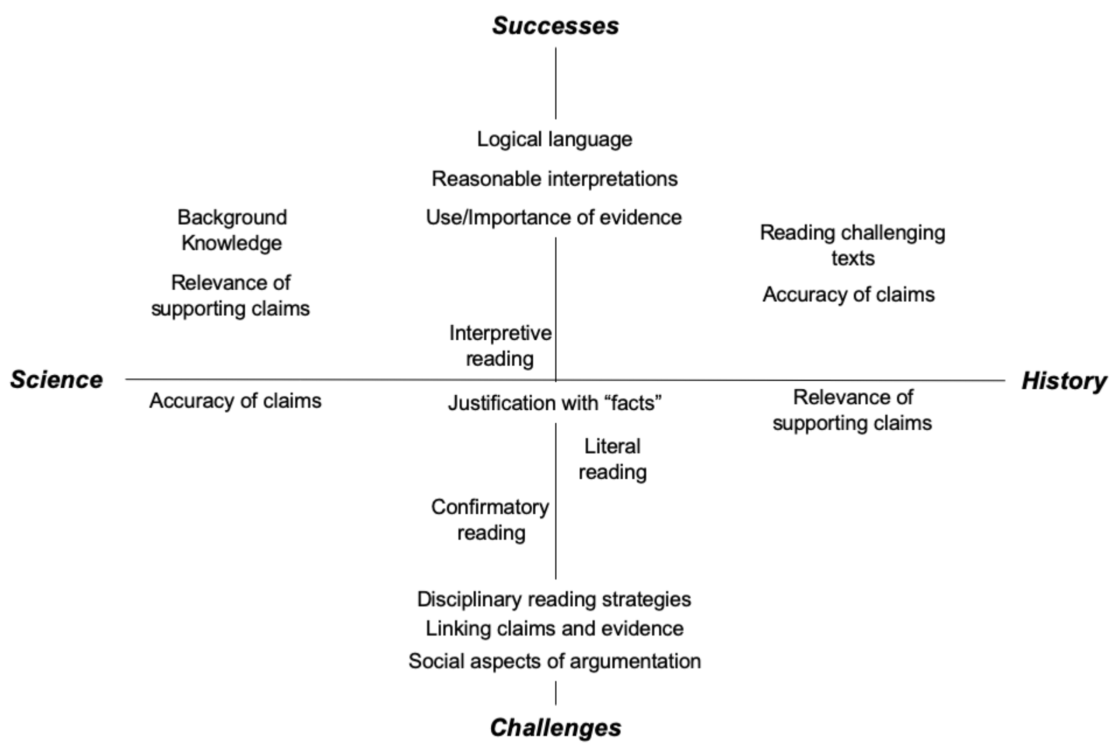


Figure 26: Landscape of Findings Across the Study.

Reviewing the landscape of findings represented in Figure 26 shows that their practices across disciplinary lines were generally similar. They faced challenges and had successes in many of the same aspects of argumentation in both disciplines. This suggests that for students new to the practice of argumentation, even if they have different ideas about the disciplines, they may not necessarily engage in the practice in significantly different ways. Instead, they appeared to struggle with many of the same issues in both cases. They read the evidence in ways that did not always support the ultimate goal of developing and supporting a claim (i.e., reading for literal comprehension), sometimes used the evidence to fit their existing ideas, and did not articulate the reasoning behind the evidence they selected. But there were also successes that were similar across the two domains. Much of the evidence they provided was relevant to their argument, they showed initial attempts to link evidence and claim through their use of "because," and they were able to look across multiple pieces of evidence to find support for their argument.

One major question of this study was whether or not students' ideas about the disciplines and evidence within the disciplines would have any relationship to how they actually engaged in constructing arguments in each discipline. Based on the findings, there did not appear to be a connection between their responses and discussions on the pre-survey and how they completed the investigations. While there were certainly disciplinary differences in the pre-survey, similar differences were not evident in their work on the investigations. Their apparently stronger ideas about science as a discipline, for instance, did not translate into particularly stronger arguments. Similarly, their relatively limited understanding of what constituted evidence in history did not seem to have any connection to their ability to use such evidence to support a claim, as their performance on the history and science investigations were roughly equal. In fact, in some cases the students even did a better job of interpreting the evidence in the history investigation than they did in the science investigation.

Students' Struggles to Articulate Reasoning

Figure 26 also shows that students had difficulty with the "reasoning" element of argumentation in both studies, which is consistent with previous studies on students' argumentation. Warranting the connection between data and claims has been shown to be especially difficult for students in both science (Clark & Sampson, 2007; Erduran et al., 2004; Jiménez-Aleixandre, Bugallo Rodríguez, & Duschl, 2000; Sandoval, 2003; Sandoval & Millwood, 2005) and history (De La Paz & Felton, 2010; Nystrand & Graff, 2001; Young & Leinhardt, 1998). Although I modeled providing the reasoning on the SenseMaker board, the students clearly needed more supports to be able to do this. It should not be assumed, however, that the fact that the students did not include their reasoning for how the evidence and claim were connected meant they did not have any connection in mind. Everyday argumentation and informal logic often include unstated premises or warrants (Simosi, 2003; Walton, 2001), and these may in fact feel superfluous when the reasoning feels self-evident. So rather than taking a deficit view of students' thinking and assuming they do not have a reason for using a particular piece of evidence to support a claim, it is important to keep in mind that it is simply an aspect of argumentation that they are unfamiliar with and thus do not include and have difficulty articulating.

Despite not articulating the reasoning that linked their claims and evidence, there were indications that students were at least attempting to do this. Their use of logical connectors, "because" in particular, showed that they were attempting to demonstrate that their evidence provided some justification for their claims and that there were reasons supporting the ideas they expressed. Keeping in mind that the students had not had prior instruction in argumentation or explicitly stating the reasoning portion of their arguments, this suggests that the students had some idea that justification was a necessary step.

Students Demonstrated the Building Blocks of Argumentation

Despite the clear challenges these students faced and the many possible ways to support them, Figure 26 also shows that there were several ways students demonstrated they were ready to engage in the practice. The findings revealed some of the building blocks of argumentation that students new to the practice can bring to the table and use as the foundation of future practice.

Although the students had mixed ideas about evidence, there were many indications that some of the students understood its importance to developing claims, argumentation, and justifying one's ideas. In the pre-survey students cited "discovery" and "proof" as reasons for using evidence and in their discussions, some students turned to the sources as they developed their claims and built their arguments. If students are to use evidence in appropriate ways, they need to understand the reasons for using evidence in the first place. The findings here suggest that even without extensive instruction in argumentation, students at this age are somewhat aware of the importance of evidence and its role in argumentation.

Relatedly, even though they did not articulate the reasoning behind their arguments, some students seemed to understand that arguments needed justification. This was evident in their use of logical connectors like "because" and in their discussions where some students pressed their classmates for reasons supporting their ideas. The fact that the students spontaneously used the word "because" when constructing their arguments suggests they are ready for more explicit instruction on how to justify one's claims, namely providing the reasoning linking claims and evidence.

The students also seemed open to the idea that multiple pieces of evidence were necessary to better support their arguments. The students appeared to be comfortable with the idea that they needed to include more than one piece of evidence and needed relatively few supports to do this. Furthermore, the students were discriminating in what evidence they included in their argument. While the quality of evidence was sometimes inconsistent and at times did not clearly support their claim, the students did appear to value some sources over others. This suggests the students had a basic idea that more evidence makes a stronger argument and that some pieces of evidence are better than others.

When considered together, these details about the students' nascent ideas about and skills in evaluating evidence and constructing arguments suggest that introducing the practice to students of this age is appropriate. These students appeared to be ready for instructional activities centered on source analysis and argumentation. Given more supports in evaluating evidence, constructing arguments, and having meaningful discussions, these students would have likely been able to produce more consistent arguments based on more accurate interpretations of the sources.

Thinking about and using evidence

Earlier research on students' ideas about evidence is consistent with my findings about students' relatively limited ideas of what constitutes evidence. Barton (2001) found that a group of elementary students in the U.S. believed that we find out about the past through word of mouth stories originating with people who were at the event. They did not mention other sources of information, and even when the students mentioned books, they were framed as simply a recording of an oral telling. The students in my study also had a limited understanding of what constituted historical evidence. And while they did not mention first-hand accounts passed down orally during the pre-survey, students in the interview expressed this idea in their discussion of the evidence.

There are two possible explanations for the students in my study not mentioning oral history and passing down first-hand accounts as a means of knowing about the past. First, they may have simply not believed that this was a way to know about the past. Second, given that the interviews suggested some students did have ideas about transmitting first-hand accounts via word of mouth, the wording of the question may have influenced their responses. Where Barton asked students, “how [do] people find out about the past?” this study asked, “What do historians use as evidence?” The students in my study may have left out first-hand accounts because of their inconsistent or unclear ideas about evidence, historical practice, or historians. Many students, for instance, conflated historical science with history. For these students a story from someone who was “there” may not be an obvious way of learning about the past. Other students did indicate that “discovery” was a purpose of evidence, but their ideas about what was being discovered was still unclear and vague, such as the group who said that historians use evidence “to get knowledge.” Since the students seemed to have such a tenuous understanding of what evidence was and its purpose, the phrasing of questions may be of particular importance in learning about students’ ideas about evidence.

The students’ ideas about evidence in science were similarly aligned with previous work. Some students did have a sense that proof was important but as with students in other studies, they rarely mentioned data, argument, or explanation in their discussion of evidence (McNeill, 2011). While McNeill (2011) was able to identify changes in students’ ideas about evidence and argumentation, because the questions on the post-survey were different from those on the pre-survey, I cannot say how much, if at all, the students’ ideas changed over the course of the investigations. While it is possible they may have developed a broader sense of what counts as evidence in the two fields, they did not pay any particular attention to the different types of evidence they used, nor did they receive any direct instruction on how the sources were different, where they came from, and what genre of text they were. It is uncertain, however, if this kind of experience was sufficient to help grow students' ideas about evidence and its uses. Further research on the types of activities, materials, and instruction that affect students' beliefs about what constitutes evidence and how it is used is warranted.

The students' use of evidence in this study also mirrors what has been found in other research on how students read and use evidence. Primarily, they did not use strategies consistent with disciplinary practice as they read the sources. The students did not use the heuristics of sourcing, contextualization, or corroboration (Wineburg, 1991) or pay much attention to source information (Britt & Aglinskas, 2002). In fact, there were only two times when they noted the sources. In one instance a student noted that I had authored the source but did not visibly apply that to her analysis. In another, a group noted that the source came from an article but misinterpreted that to mean that the article was a primary source. In addition, the students had difficulty moving across text types and did not attend to the differences between the type of sources they were reading (Afflerbach & VanSledright, 2001).

Ideas about the Disciplines

Another contribution of this study are the findings related to the students' understanding of the disciplines of science and history. The major difference between the students' ideas and practice in science and history was the apparent stronger understanding of science compared to history. In the pre-survey, they were able to name more examples of the kinds of evidence used in science, and their examples demonstrated a knowledge of several scientific sub-disciplines. In the source evaluation task, the students' reading of evidence in science was slightly more sophisticated in that they were more likely to make inferences based on the sources in comparison to the history task. They also gave longer and more detailed responses on the science task. In addition, the students were more likely to incorporate outside information in their reading of sources in the science task, suggesting their background knowledge was stronger. Taken together, these findings suggest that students' understanding of science in general was stronger than their understanding of history and historical practice.

One possible explanation for this is the fact that these students were probably much more likely to encounter representations of science and scientists than of historians and historical practice. Whereas the "mad scientist" trope is well trodden, and a number of popular television shows depict criminal forensic science, there are relatively few representations of historians. One exception might be the adventuring archeologist found such as Indiana Jones or the characters in *The Mummy*, but these still occur less frequently. The research on depictions of these practitioners is another indication of this difference. A search for "popular representations scientists" yielded many results addressing topics such as the mad scientist trope, representations of gender, and the link to science education. My search for "popular representation historians," however, did not provide any relevant results.

Additionally, history as a school subject does not enter in to the curriculum until third-grade in this school district. Prior to third-grade, social studies instruction focuses on issues related to the community and civics. Consequently, these students have likely had significantly less exposure to history as a topic of study than science. It should not be a surprise, then, that they appeared to have a better developed understanding of science than of history.

These findings suggest future directions for research on students' understanding of the disciplines, and for history in particular. There is a relatively large body of work on students' beliefs about science, such as research that focuses on students' beliefs on the nature of science (NOS) and its relationship to their work in school science (Deng, Chen, Tsai, & Chai, 2011). There have also been a number of studies using the "Draw-A-Scientist Test" in which students are asked to draw a scientist as a means of learning about what a scientist is, who can be a scientist, and other issues related to students' and practitioners' identities (Chambers, 1983; Losh, Wilke, & Pop, 2008; C. L. Mason, Kahle, & Gardner, 1991). Research on history education, however, does not appear to have analogous studies on students' beliefs about historians and the discipline of history. Based on the findings in this study, it may be worth learning more about what students think about the nature of history, who historians are, and how they develop these ideas.

Methodological Limitations and Implications

One of the unexpected consequences of this study was that it revealed important methodological issues related to studying elementary students, argumentation, and interdisciplinary practice. The findings highlight the importance of the kinds of data collected on students' argumentation and the need for multiple sources of data on students' thinking, particularly for children of this age who have difficulty articulating their ideas. The results of this study also suggest the importance of acknowledging the difficulty of working with younger students and the inconsistent quality of data that comes from such investigations.

There are also limitations to the findings of this study related to the methods employed. First, as with all case studies, the sample size was relatively small, making it impossible to generalize the findings. This was compounded by the fact that some students had to be excluded from the analysis. In addition, however, there were limitations related to the population of students that I chose to work with and the lack of pilot work to guide the final study design.

Multi-Modal Data is Key to Analyzing Argumentation

One key methodological implication of this study is the importance of having multiple sources of data on student thinking, particularly with younger students. Studies in argumentation often use students' writing as a primary source of data (Britt & Aglinskas, 2002; Ryu & Sandoval, 2012; Young & Leinhardt, 1998). As this study demonstrated, however, students can struggle to articulate their thinking and write clearly. This is especially true for students such as these who are both younger and unfamiliar with composing arguments. Consequently, much of the students' written work in this study would have been difficult to interpret without the accompanying video recordings. Even though they were limited, the students' conversations were many times essential to understanding the students' responses, such as in the case where the written response "record books" meant *The Guinness Book of World Records* and not historical records. In addition, the audio captured a great deal of thinking that was not evident in the students' written responses. Without the audio, for instance, I would not have been able to identify the different strategies that students used to develop their arguments or see how much confirmation bias played into their source analysis. The importance of the video data should not be overlooked either. Besides helping to clarify what students were talking about when they said things like, "Give me that one," the video recordings also provided insight into their work and thinking. I was able to observe, for instance, when students consulted sources before speaking, or when a student was working independently and generating ideas on their own.

Together, this suggests that students' written work may provide insufficient evidence of their thinking and approach to argumentation as the video and audio data often made it possible to make sense of the students' written work. In addition, the video and audio recordings provided data that gave insight into the students' thinking in ways that written artifacts on their own could not. Audio and video data, then, can clearly provide insight into the ways that students create and support arguments and should not be overlooked.

The Difficulty of Interpreting Argumentation without True Discussion

Perhaps the greatest analytical challenge I encountered was the consequence of working with relatively young students on a task requiring so much sophisticated thinking. Namely, the lack of meaningful discussion among students provided a significant challenge to analysis. Given how infrequently they spoke about their ideas about the evidence, investigation question, and their argument, it was often difficult to characterize their thinking. This lack of true discussion in the small group work highlights a challenge of studying students' argumentation in a setting such as this. One of the goals of this study was to establish how students engage in the practice of argumentation without any prior instruction in the practice. A problem with studying students who have not been prepared in the practice, however, is that they have not been prepared in the practice. Consequently, many students did not articulate their ideas or interact in ways that would have done a better job of revealing their thinking. The apparent solution would be to provide the students with some training in having a discussion. While this might yield better data about the students' thinking, it would also introduce the possibility that their thinking might be influenced by that training and thus produce results that do not reflect a "baseline" of students' thinking.

Another possible solution might be to engage students in an argument about an everyday topic that they were more familiar with and perhaps more interested in. Such work on students' everyday argumentation has been done before (Bricker, 2008; Hudicourt-Barnes, 2003) and shown that students do have competencies making reasoned justifications for a claim. This approach might provide more information about the initial ideas and practices students have about argumentation. The shortcoming, however, is that it would not provide any information relating to how students construct arguments in different academic disciplines.

The nature of students' talk during the small group work highlights another difficulty of this kind of research. Perhaps as a former teacher I should have expected it, but the students were often playing games, being silly, or discussing things that were probably more interesting to them than school work, such as when Aisha played "jinx" with her group. These are all behaviors that were totally normal but made analysis very difficult. Almost no research on children's work that I have found, however, makes any mention of this analytical challenge. Not acknowledging the "messy" nature of student work obscures how students actually work and behave. This can potentially misrepresent the data on which claims are made and make them appear more clear-cut than they may be. Second, it furthers the disconnect between academia and the practice of teaching. Teachers may feel that educational research is too theoretical and has little bearing on teaching in the "real world." Presenting results in a way that distorts the actual classroom experience may reinforce this idea that research does not study "real teaching" and thus alienate teachers further.

The Challenge of "Authentic" Contexts

In an issue related to the reasons behind students' lack of discussion, one of the biggest challenges to this study and a factor in its limitations (though also its strengths), was the study's context. It was important to me that the setting of the study be "authentic." I wanted to avoid an idealized setting in which students had access to better than typical resources, represented a single demographic, or were prepared in the practice of argumentation. To that end, I selected a school site that was demographically diverse, had a fair number of English language learners, and did not have better than average scores on standardized tests. In addition, the students were new to the practice of argumentation, and had not previously received instruction in it.

On one hand, I believe this provided a more realistic representation of how a typical class of third-grade students might learn and engage in argumentation in science and history. While there were clear leaders in some of the groups, such as Angela and Andrew, who appeared to have a stronger grasp of what to do, there were also students representing a broad spectrum of skills with argumentation. In addition, the picture of argumentative practice I have provided here truly presents the work of novices in the practice. Not only had these students not had prior school-based experiences with it, they were also relatively young. On the other hand, one consequence of using this setting was some of my data was “messy” or unusable for the purposes of establishing an understanding of how elementary students engage in argumentation in science and history. The data from Group 2, in particular, had to be excluded because of the significant amount of support the students ended up receiving from Ms. Fisher. Two of the students were English language learners and the other two read significantly below grade level.

Doing work with such a population, then, can present significant analytical challenges and limitations to the findings. Although I did not do an in-depth review of Group 2’s work, my initial analysis suggests that beyond their struggles to read and make sense of the sources, they had significant difficulty simply understanding the questions, including those on the pre- and post-survey. Given the students’ reading levels this should not be surprising, but it does not yield much useful information about how students engage in argumentation.

As implied in my previous discussion of possible solutions to working with students unaccustomed to “true” discussions, when working in a “realistic” setting one must balance the authenticity of the context with the challenges it presents. Working in a diverse setting with younger students who are unfamiliar with argumentation may have presented a more typical educational context and allowed me to identify challenges of teaching argumentation that many teachers will face. But it also made it more difficult to get a clear picture of specific elements of students' thinking.

Learning to Argue versus Learning through Argumentation

One limitation of this study is that it focused heavily on how students learn to engage in argumentation, but did not attend to whether or not they learned anything through the process of developing in arguments. My findings suggest the ways that students learn to use evidence and construct argument, but not if those activities strengthen their understanding of content. While exposing them to primary sources and data analysis in this way engaged them in academic activities they had not done before, there was little evidence to suggest they learned the content better or worse than through direct instruction or other means. While the disciplinary practices are important to learn, they should be the means to an end, and not the sole focus of instruction. Otherwise, students would simply be learning the procedures for their own sake. It is important, then, to understand the connections between engaging in argumentation and learning content. This study, however, was not designed to answer such questions. Future work comparing students' argumentation across disciplinary lines should not only examine the ways in which students evaluate evidence and construct arguments in different content areas but how the design of investigations or activities maximize learning about the practices and through the practices.

Evaluating sources, for instance, is one area where the differences in disciplinary practice may have consequences for the ways that students learn content. In history, one aspect of preparing students to look at sources in a disciplinary appropriate way would be to support them to use the heuristics of sourcing, contextualization, and corroboration (Wineburg, 1991). Using contextualization as an example, students would need to develop a base of background knowledge about events, actors, and other elements of the context surrounding a historical period. This kind of preparation would likely support their learning in ways that simply introducing them to the generic aspects of argumentation, as this study did, would not. In science, students might need to assess the quality of evidence or its relevance to the question at hand. In this case, they would need to understand both how the data was collected, thus improving their knowledge of scientific practice, and the various elements of the scientific phenomenon.

Pilot Work

Another methodological takeaway from this study is the importance of doing pilot work when expecting students with no experience to engage in such a sophisticated practice as argumentation. The complexity of developing and supporting a claim not only requires students to understand the structure of arguments and have the ability to articulate their thinking. In addition, they need to have a certain level of reading skill in order to make sense of sources which would likely be novel in terms of genre, content, and structure. To be successful, students also need enough background knowledge to appropriately interpret sources and draw conclusions from them. Unfortunately, I was unable to do pilot work for this study and I believe that some degree of pilot work would have benefitted the study. In moving forward with work on students' argumentation in science and history, however, I can think of this as the pilot for future work. In the sections below, I treat this study as a pilot describe the changes I would make for a second study based on my findings.

Identifying content. Many of the limitations of this study can be traced back to the topics and content of the investigations. Although I worked to make the investigations and materials as similar as possible, the history investigation was more likely challenging for the students than the science investigation. This was partly due to the students' unfamiliarity with both the content and the process of history investigation, especially in comparison with the science investigation. In the case of science, the investigation was built upon phenomena that students would have had first-hand, everyday experience with, namely the relationship between force, weight, and speed. Even if students did not know what a Newton was, or had difficulty grasping the effect of aerodynamics, they would have likely been able to make reasonable predictions about the outcome. In fact, this is what happened with some students as evidenced by my description above of the ways some students interpreted the sources to support their pre-existing ideas.

In contrast, the history investigation did not have any connection to students' everyday experience (besides being related to Michigan) and required a good deal of background knowledge that they likely did not have. For instance, they appeared to be unfamiliar with the concept of political parties and the process of representative democracy in general. Background knowledge is essential to the quality of arguments (von Aufschnaiter, Erduran, Osborne, & Simon, 2008), and many students had difficulty making sense of the sources because of their lack of background knowledge.

The disparity between the two topics was a result of my attempt to integrate the investigations into the classroom's existing curriculum and grade-level standards. Another example of the difficulties posed by conducting a study in an authentic teaching context. As a result of this constraint, the topics available for the history investigation were somewhat limited. The history standards for third grade in Michigan focus on Michigan's history. Within the content outlined in the standards there were very few topics that left room for multiple interpretations. Much of the content focused on historical narratives around which there was little debate. This left me with few topics to build an investigation with multiple possible answers around. This was compounded by the fact that, in science, the standards at that time of year focused on force and motion, topics that the students had a great deal of everyday experience with.

Although it was very important to me to integrate the content of the study with what students needed to learn, I believe in the next study I might set that consideration aside if it made it difficult to make the two investigations equally difficult. I would not abandon the appropriate standards or existing curriculum entirely, but I would work more closely with the classroom teacher to identify topics that included more equal amounts of students' everyday experience and existing background knowledge.

Text difficulty. Relatedly, the difficulty of the texts in the two investigations were mismatched, despite my efforts to make them roughly equivalent, and the texts from the history investigation were generally more complex than those in the science investigation. This was partly because so much background knowledge was necessary to understand the history investigation question and its possible answers. In addition to providing details that could serve as evidence to support a claim, the sources also had to provide some of the context necessary to understand the situation surrounding the votes for Michigan's statehood. In contrast, the science investigation required less background knowledge and had fewer factors to consider. Students only needed to understand the relationship between mass and force, and the effects of aerodynamics. For the history investigation they had to understand the process for admitting a state, what representatives are and how they are chosen, how political parties function, and the motivations of each political party. As a result, both the investigation overall and the texts themselves were more complex in the history investigation than the science investigation.

A follow-up study would address this by first calibrating the difficulty of the investigations overall by finding matching content as described above. Second, I would bring in more outside expertise writing and modifying the texts to use in the investigation. While I did work within the classroom context to develop the content, Ms. Fisher was not as involved in developing the materials as I would have liked. Bringing in the classroom teachers' knowledge of grade appropriate text difficulty, as well as of how their students read, would likely make the texts better matched to the students' skill level. Furthermore, I would also bring in a literacy expert into the process sooner. While I did have a scholar in the field of literacy review the texts and provide feedback, they were not involved in drafting the texts which I believe would have yielded better results.

Supports for students. Although the differences between the history and science investigations may have had implications for findings related to students' reading strategies or means of constructing and supporting claims, they do not appear to have affected the social aspects of students' argumentation. In both investigations, the students were equally likely to engage in superficial conversations or limit their interaction to taking turns sharing their own ideas without responding to one another. On one hand, this was an important finding that supported prior research that students benefit from preparation and support during discussions (Ashman & Gillies, 1997; Gillies & Ashman, 1998) and also suggested that students new to argumentation likely need explicit instruction in how to engage in the social aspects of argumentation. On the other hand, however, the lack of meaningful interaction limited the insight I was able to gain into their thinking around evaluating sources and constructing arguments.

Having established that students are not likely to independently engage in the social aspects of argumentation, in a follow-up study I would provide students with preparation in taking part in meaningful discussions. One danger of including such supports or preparation is that students' thinking might be influenced by the protocols or tools they are given, thus limiting the value of my findings. It is possible, however, to design such supports in a way that does not rely on the language or structure of argumentation or evaluation of evidence. Ashman and Gillies (1997), for instance, provided students with supports that focused on the interpersonal and collaborative aspects of group work such as stating ideas clearly, providing constructive criticism, and clarifying differences of opinion. Such supports would not include the language of argumentation but could still result in students expressing their ideas more clearly, prompting each other for their thinking, and responding to one another's ideas. Ideally, this would result in more insightful data without unduly influencing how they engage in argumentation.

I would also change the design of some of the materials to better support students and get more insight into how they treat different kinds of sources. In general, the students treated the different sources roughly the same, regardless if they were primary or secondary. In addition, some students seemed confused about who authored the different science sources and appeared to believe that all of the primary sources were written by one of the teams in the scenario. One possible explanation for the students' confusion and similar treatment of the sources is the design of the materials. The design of each worksheet was exactly the same and there were no visual cues to differentiate the nature of one source as opposed to another. A key modification to the worksheets would be to include a photograph or reproduction of the original document. Seeing the difference in appearance between, in the case of the history investigation, a newspaper article from 1836 and a book chapter from 1989 may have made it clearer to students that these documents were from different time periods. Such a change may also make the author of the texts more "visible" and encourage greater engagement with the texts (e.g., Paxton, 2002). While this would not necessarily result in students applying different reading strategies, it would allow me to eliminate confusion about the nature of the source as a possible explanation for their similar approaches across texts.

Methodological changes. In terms of the study's methods, I believe that a second study would benefit from more explicit think-aloud protocols, particularly for the pre- and post-surveys on students' ideas about evidence, argumentation, and the disciplines. One difficulty in analyzing these data was determining whether or not students' responses were or were not appropriate to the discipline. Having a simple list of what they thought counted as evidence was sometimes insufficient to characterize their ideas. When students listed "DNA" as an example of evidence in history, for instance, it was not clear if they had a use of DNA as evidence in mind that was consistent with historical practice. It is possible they were aware of ways that historians might use DNA to establish important familial relationships, but I believe it is more likely that they had a different idea in mind. Given their tendency to conflate historical sciences and the discipline of history, they may have been thinking of DNA as a way to identify animals from the past. Without a clearer idea of their thinking, as would be provided by a think-aloud, it is not possible to build a complete picture of that student's understanding of historical evidence.

One way to implement data collection through a think aloud would be to include think-aloud as part of the focus group interviews. That would provide me with the greatest opportunity to follow-up on students' thinking and probe them for more detail. Conversely, however, it is also a very time consuming way of collecting data that is difficult to implement logistically within the constraints of the school day. It also requires the students to recall what they were thinking as they recorded a response. In a second study, I would capture their thinking using the audio and video recordings of their small group work as a way to capture students' thinking in a think aloud protocol. Leaving the students to independently manage themselves in a think aloud protocol, however, leaves open the very good chance that they would become distracted or abandon the protocol part way through. One possible solution to this problem would be to leverage the students' interest in the recording equipment and their tendency to want to perform for the camera. The think aloud protocol would involve the students first recording their ideas independently on paper and then taking turns "presenting" their ideas to the camera and explaining their thinking behind their responses. The other students would also be encouraged to respond to their partners' ideas during this time as well, possibly through assigning roles related to eliciting and responding to ideas. I believe the excitement of performing for the camera would make students excited to present their ideas, and more likely to engage in the think-aloud protocol consistently.

Research questions. The findings from this "pilot" study also influence the research questions I would ask in a follow-up study. In the case of my first research question, I would only make minor adjustments to the main research question but would significantly change the supporting questions. The original set of research questions were:

1. How do students interpret and make sense of evidence in science and history?
    A. What are the similarities and differences between these strategies in science and history?

I would revise these questions as follows:

1. How do students make sense of and interpret sources in science and history when developing a claim?
    A. What kinds of details do they identify as important?
    B. Why do they prioritize some details over others?

C. What role does an investigation question play in how students read evidence?

First, I modified the overarching question by clarifying that students would be reading the sources in service of developing a claim, and changed the language from "evidence" to "sources" to align with the idea that sources do not become evidence until they are used to support a claim. Second, I generated a new list of sub-questions in response to the findings and shortcomings of this study. In particular, I am interested in learning more about the details that students identified as important in their analysis of the documents. While in this study I distinguished between responses that were the result of literal vs. inferential reading strategies, this does not necessarily mean that there was no analysis in students' thinking as they selected details or paraphrased the sources. Distinguishing between relevant and irrelevant details is in itself an analysis. It was unclear from the data I collected, however, why the students picked the details they did and whether or not it was the result of thinking specific to the investigation or a more generic approach to demonstrating understanding of a text. The improved think-aloud and discussion protocols described previously might allow me to collect data to make such an analysis.

Another change I would make to the research questions would not involve changing the research question, but rather my approach to answering the question. This applies to my third research question and sub-questions:

3. How do students coordinate evidence with claims when engaging in argumentation in science and history?

    A. What kind of reasoning do students use to coordinate evidence with claims?

    B. What are the similarities and differences between this activity in science and history?

The present study focused less on *how* students coordinated evidence with claims and more on *if* they included such a coordination in their arguments. This was largely due to the design of the activity and materials which provided few scaffolds to support students to include or state their reasoning. My desire to establish students' practices "at the ground floor," however, partly resulted in replicating previous findings that students frequently struggle to articulate or provide the reasoning element of arguments (De La Paz & Felton, 2010; Erduran et al., 2004; Jiménez-Aleixandre et al., 2000; Young & Leinhardt, 1998). Consequently, since these students did not include their reasoning I was not able to say very much about the strategies or kinds of thinking they do to tie evidence to claims. In a follow-up study I would prioritize making the students' thinking visible over observing their initial, unprepared practices in argumentation. I would do this by providing more supports in the materials and activities to help students provide the thinking behind the evidence they chose. For instance, in the procedure described above where students talk directly to the camera, I might include a prompt such as, "This piece of evidence proves my point because…" or other similar sentence stems.

Finally, I believe that I would separate my final research set of research questions into a separate study which would focus more on the differences between students' beliefs about the evidence, argumentation, and the disciplines of science and history, than on the connection between those beliefs and how they engage in argumentation. The original question and sub-questions were:

4. How do students understand the use of evidence in arguments in science and history, and the relationship between the two?
    A. What connections, if any, do they see between their work engaging in argumentation in science and history?

B. What connection, if any, do they see between their work engaging in argumentation and the work of scientists and historians?

C. Is there a connection between how students understand the work of scientists and historians and how they coordinate claims with evidence in each subject area?

The main reason for separating these research questions into a different study would be to refine the instruments I used to evaluate students' ideas in these domains. As indicated in Chapter 7, the data presented a number of limitations, which were partly due to how I asked students to complete the task. Some of these problems would be addressed by the methodological changes described above, which would support students to articulate their thinking more consistently. But I also believe the instrument could be refined to better capture what I am looking for. The wording of the question, "How do scientists/historians do their jobs?" for instance, might be adjusted to focus students' attention to the ways that practitioners in the disciplines construct knowledge, such as Barton's (1997) phrasing of, "How [do] people find out about the past?"

Another reason for taking the time to develop a better instrument would be to reincorporate the activities related to students' ideas about the disciplines in a later study, but for a different purpose. Given the findings presented here that students at this grade-level have nascent abilities in argumentation and seem ready to learn the practice in greater sophistication, the connection between their pre-existing ideas about the disciplines and their practice seems less interesting because they can clearly start to learn the practice with support. Instead, I believe it would be more important to understand whether engaging in the disciplinary practices changes their ideas about the disciplines or the relationship between "school" history and science and "real" history and science. My findings suggest that students continued to see a difference even after the investigations, but it is worth further investigation.

Instructional Implications

Although there were many aspects of high-quality argumentation that were absent from how the students completed these activities, my findings do suggest that it is worth the time to engage students of this age in this sophisticated work. Considering they did not have any guidance in making sense of sources or constructing an argument, the students did fairly good work. Overall, their final arguments were generally coherent and made use of multiple pieces of evidence. Furthermore, the small group discussions also showed that there were some students who already seemed able to engage in effective argumentation. Angela is a good example of this. She made her case to Solomon about the science investigation by referring him to the data and even managed to change his mind in response to her argument. This does not just show that Angela was able to use evidence appropriately and effectively, but that Solomon also saw that claims in an argument needed to be consistent with the available evidence and that one's ideas should change based on that evidence. Bearing this in mind, however, there are still a number of instructional practices and supports implicated by this study.

Students need Preparation for Discussions

As previous studies suggest (Ashman & Gillies, 1997; Gillies & Ashman, 1998), students need preparation, training, or scaffolds to engage in quality independent discussions. The students in this study were not given any such preparation, so it is perhaps unsurprising that their talk during small group work was often off-task or did not include a meaningful back and forth about their ideas. The practice of argumentation consists of more than just producing a written argument including a claim, evidence, and reasoning. It is a dialogic process which is often focused on consensus building among participants (Driver et al., 2000; Garcia-Mila et al., 2013). In many ways this is different than students' typical experiences in classroom discussion.

The students in this study would have benefitted from some kind of preparation in how to work in small groups and the kind of talk they needed to engage in. Primarily, they needed supports to help them move beyond simply taking turns sharing their ideas and instead respond directly to what other students said. One way to support this is through establishing discussion protocols. Protocols, such as Collaborative Reasoning (A.-M. Clark et al., 2003), structure the conversations that students are to have and guide them through the important stages of meaningful discussion. Relatedly, students may also be assigned roles to take on during a discussion and be provided with talk moves or questioning strategies to use in those roles (Herrenkohl, 2006; Palincsar & Brown, 1984). Assigning students roles in a discussion makes the different ways that people in a discussion interact with each other explicit, and ideally are a scaffold that can eventually be removed. Finally, teachers can simply model questioning strategies and establish them as the norm in classroom talk (Gillies & Khan, 2009). The findings from this study suggest that one, or some combination of these supports, could have enhanced the quality of students' argumentation by providing them with concrete moves or strategies to employ during their small group work.

Pushing Students beyond "Playing School"

In concert with other findings from this study, the students' lack of meaningful talk also suggests that students need to be pushed beyond "playing school" when engaging in a practice as sophisticated as argumentation. In other words, the way students behaved during and participated in the investigations may have been influenced by what they perceive as a teacher's expectations and the norms set during the typical school day.

As noted above, the discussions I was hoping the students would take part in were likely very different from what they typically experienced during discussions in their regular day. Although I did not have the opportunity to observe Ms. Fisher leading or initiating discussions, teachers may not typically engage students in the kind of talk that encourages an exchange of ideas. One example of such talk is the IRE talk pattern in which a teacher initiates (I) a question, the student responds (R), and the teacher evaluates the response (E), can make up a significant portion of teachers' interactions with students (Mehan, 1979).

Because talk patterns such as IRE focus on the teacher as the arbiter, students become unaccustomed to interacting with each other rather than the teacher. Furthermore, these talk patterns may also influence the ways students engage with each other. Students may mimic what teachers model and adopt the IRE talk pattern in their own discussions with one another (Prado-Olmos, 1994). Besides limiting students' experience with meaningful discussion, there is a second consequence of talk patterns such as IRE. Because the interaction ends with a teacher's evaluation of the statement, the implication is that students should be making contributions to the conversation that match the teacher's expectations and not necessarily those that express the student's ideas. The goal of classroom discussions, then, becomes to provide the "right" answer and not to explore one's thinking and engage with other's ideas. Using this lens of "playing school," the students' interactions during small group work may well have been influenced by their ideas of what was expected of them. They infrequently engaged substantively with one another's ideas and more often just took turns stating their ideas. And when they did challenge one another or provide feedback, it was often on details such as spelling and whether or not it was right.

Another way in which might play school is in their responses to texts. Students may not often do work that is disciplinary in history or science in schools, instead doing “school history” or “school science” that does not reflect the work of practitioners (Cuban, 1991). The “school” versions of the disciplines often rely on reporting facts rather than evaluating data, making inferences, or drawing conclusions. Consequently, when students read a text they may assume that they are expected to report details that demonstrate their comprehension of the text.

This provides a possible explanation for the dominant reading strategy among the groups. They often provided readings of the texts that focused on literal comprehension as opposed to interpretive or evaluative comprehension that went beyond details stated explicitly in the text. Many of their responses demonstrated the literal comprehension often required in reading instruction or the school versions of science and history. So rather than engaging in historical or scientific reading practices, they instead provided many readings that conformed to their understanding of the teacher’s expectations.

One way of getting students involved in disciplinary practices, then, is to guide them away from simply playing school and trying to meet the teacher’s expectation. This involves more than just calling students “scientists” or telling them, “today we’re going to be doing the work of historians.” Rather, it requires significant shifts in the nature of classroom dialogue and inquiry. Teachers not only need to remove themselves from the center of a discussion, as in IRE, and encourage more student to student interaction. They must also involve students in decisions about what is “right” or what counts as a good explanation. This may be an uncomfortable shift for some teachers because it means shifting the authority in the classroom away from the teacher, but it is essential to moving students away from playing school and towards engaging in more authentic disciplinary practices.

Supporting Students to Analyze for Argument

The results of this study also have implications for instruction in argumentation. The SenseMaker board, for instance, seemed to help support students to construct somewhat complete arguments, especially considering they had never done so before. The requirement that students graphically link their ideas to a source via an arrow seemed to reinforce the concept that their support for their claim had to come from the sources themselves. By gluing strips with the titles of sources onto the SenseMaker board, the students were literally constructing their argument by using the sources. Furthermore, the “AND” strips seemed to encourage them to use multiple pieces of evidence, something that can often prove difficult for students to do. Combined with the success of other students using SenseMaker boards (Bell & Linn, 2000; Herrenkohl & Cornelius, 2013), this suggests that well-designed graphic organizers can support students to make high-quality arguments without relying on essay writing as the only means of assessing their arguments.

One of the primary scaffolds built into the investigations was to separate the source evaluation from the argument construction elements of the task. This appeared to help the students in producing their arguments. In particular, it ensured that they looked at every piece of evidence before developing their argument and thus helped them include multiple pieces of evidence in their final argument. Examining their process of constructing their arguments, however, indicates that they may have needed the task to be broken down even further. Given the tendency of some of the students to fixate on singular pieces of evidence or to only look for confirmatory evidence, they would have likely benefitted from an intermediate step between *source evaluation* and *argument construction* which would have focused on *data analysis*. If an argument is a complex of data and reasoning in support of a claim, one should have a solid idea of what the claim is before developing the argument. The design of the activity in this study, however, compressed developing a claim and creating a supporting argument into one step. One consequence of this may have been to reinforce students' inclination to only look for evidence that confirmed their initial ideas about the investigation question. Separating the data analysis from argument construction may have made it clearer to those students that their answers needed to come out of the evidence, as opposed to their strategy of making the evidence fit their ideas. Furthermore, some students seemed to ignore the work they did on the source evaluation sheets and did not appear to use what they had written in their final analysis or argument. Making data analysis its own step, and one explicitly built on the first step of analyzing sources may have supported those students to see the link between their initial reading of the evidence and their final argument.

Careful materials design may be another way to support students in this kind of investigation. Particularly, the way that sources are presented to students may have an impact on how they use and interpret them. In this study, the two-part design of the source evaluation sheets along with the prompt appeared to guide some of the students to evaluate the evidence in appropriate ways and record their responses. One aspect of the sources that the students ignored or did not understand, however, was the type of source it was and where it came from. In particular, they treated primary sources in the same way that they treated secondary sources. One way to bring students' attention to the different kinds of sources they are using might be to include images of the original source along with the evaluation sheet, perhaps by reproducing the source on the opposite side of the worksheet. An image of a newspaper article from 1836 would likely have a greater impact in terms of students' understanding of it as a primary source than a simple attribution appearing below the modified text.

Conclusion

Given the importance being placed on disciplinary practices and argumentation in education, it is important to understand how students develop skills in analyzing evidence and constructing arguments, and how they engage in those practices across disciplinary lines. The findings from this study illustrate that while these practices can be challenging to younger students, they are nonetheless capable of beginning to engage in analyzing evidence to construct arguments. With almost no support or prior experience with developing claims from sources to answer an investigation question, the students were able to construct generally appropriate arguments. This was true even for the history investigation which was based on more challenging texts and was significantly removed from their everyday experience. Through their work and discussions, the students demonstrated that they were ready to learn to generate claims by analyzing evidence.

This dissertation contributes to the literature on argumentation in science and history by describing how students new to the practice begin to evaluate sources and develop and support claims. It also compared the students' activities across the two subject areas, science and history, to identify commonalities and differences in how they engage in the practice of argumentation. The findings add to our understanding of the kinds of challenges beginners face as they attempt the sophisticated work of developing and supporting claims with evidence. Furthermore, the findings shed light on students' comparative ideas about the disciplines of history and science, and the nature of evidence in each. This study has theoretical implications for the study of students' understanding of argumentation and suggests that as novices they have similar struggles across content areas. It suggests some of the building blocks that students may bring to the work of argumentation, even at a young age. The findings also have methodological implications for the field, particularly related to the challenges of working with novices and students of this age. Finally, the dissertation is relevant to teachers in the field working with students new to argumentation and in multiple content areas. Despite the students' struggles during the investigations there were numerous indications that even without direct instruction, they were able to engage in this difficult work and develop generally coherent arguments based on evidence.

WORKS CITED

Abd-El-Khalick, F., Waters, M., & Le, A.-P. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching*, *45*(7), 835–855. doi:10.1002/tea.20226

Afflerbach, P., & VanSledright, B. (2001). Hath! Doth! What? Middle graders reading innovative history text. *Journal of Adolescent & Adult Literacy*, *44*(8), 696–707. Retrieved from http://www.jstor.org.proxy.lib.umich.edu/stable/40018742

Ashman, A. F., & Gillies, R. M. (1997). Children's cooperative behavior and interactions in trained and untrained work groups in regular classrooms. *Journal of School Psychology*, *35*(3), 261–279. doi:10.1016/S0022-4405(97)00007-1

Au, W. (2009). Social studies, social justice: W(h)ither the social studies in high-stakes testing? *Teacher Education Quarterly*, 16.

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+* (p. 442). Chapel Hill, NC: Horizon Research, Inc.

Barbour, R. (2007). *Doing focus groups*. Thousand Oaks, CA: SAGE Publications.

Barton, K. C. (1997). “I just kinda know”: Elementary students' ideas about historical evidence. *Theory & Research in Social Education*, *25*(4), 407–430. doi:10.1080/00933104.1997.10505821

Barton, K. C. (2001). Primary children’s understanding of the role of historical evidence: Comparisons between the United States and Northern Ireland. *International Journal of Historical Learning, Teaching, and Research*, *1*(2), 21–30.

Barton, K. C., & Avery, P. G. (2016). Research on social studies education: Diverse students, setting, and methods. In D. Gitomer H. & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed., pp. 985–1038). Washington, DC: American Educational Research Association.

Basaraba, D., Yovanoff, P., Alonzo, J., & Tindal, G. (2013). Examining the structure of reading comprehension: Do literal, inferential, and evaluative comprehension truly exist? *Reading and Writing*, *26*(3), 349–379. doi:10.1007/s11145-012-9372-9

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*(8), 797–817. doi:10.1080/095006900412284

Berland, L. K., & McNeill, K. L. (2012). For whom is argument and explanation a necessary distinction? A response to Osborne and Patterson. *Science Education*, *96*(5), 808–813. doi:10.1002/sce.21000

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

Blank, R. K. (2013). Science instructional time is declining in elementary schools: What are the implications for student achievement and closing the gap? *Science Education*, *97*(6), 830–847. doi:10.1002/sce.21078

Bricker, L. A. (2008). *A sociocultural historical examination of youth argumentation across the settings of their lives: Implications for science education* (Ph.D.). University of Washington, United States -- Washington. Retrieved from https://search.proquest.com/docview/304447739/abstract/425B4F10F7D34C37PQ/1

Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education*, *92*(3), 473–498. doi:10.1002/sce.20278

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522. doi:10.1207/S1532690XCI2004_2

Carr, E. H. (1961). The historian and his facts. In *What is history?* (pp. 3–35). New York: Vintage Books.

Center on Education Policy. (2008). Instructional time in elementary schools: A closer look at changes for specific subjects. *Arts Education Policy Review*, *109*(6), 23–28. doi:10.3200/AEPR.109.6.23-28

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, *67*(2), 255–265. doi:10.1002/sce.3730670213

Chambliss, M. J., & Murphy, P. K. (2002). Fourth and fifth graders representing the argument structure in written texts. *Discourse Processes*, *34*(1), 91–115. doi:10.1207/S15326950DP3401_4

Clark, A.-M., Anderson, R. C., Kuo, L.-j., Kim, I.-H., Archodidou, A., & Nguyen-Jahiel, K. (2003). Collaborative reasoning: Expanding ways for children to talk and think in school. *Educational Psychology Review*, *15*(2), 181–198. doi:10.1023/A:1023429215151

Clark, D. B., & Sampson, V. D. (2007). Personally-seeded discussions to scaffold online argumentation. *International Journal of Science Education*, *29*(3), 253–277. doi:10.1080/09500690600560944

Clymer, T. (1968). What is 'reading'?: Some current concepts. In *Yearbook of the National Society for the Study of Education* (Vol. 67, pp. 7–29). Chicago: University of Chicago Press.

Collingwood, R. G. (1946). *The idea of history*. Oxford: Clarendon Press.

Cuban, L. (1991). History of teaching in social studies. In J. P. Shaver (Ed.), *Handbook of research on social studies teaching and learning*. New York: Macmillan.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, *97*(2), 139–156. doi:10.1037/0022-0663.97.2.139

De La Paz, S., & Felton, M. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*(3), 174–192. doi:10.1016/j.cedpsych.2010.03.001

De La Paz, S., Monte-Sano, C., Felton, M., Croninger, R., Jackson, C., & Piantedosi, K. W. (2017). A historical writing apprenticeship for adolescents: Integrating disciplinary learning with cognitive strategies. *Reading Research Quarterly*, *52*(1), 31–52. doi:10.1002/rrq.147

Delen, I., & Krajcik, J. (2015). What do students' explanations look like when they use second-hand data? *International Journal of Science Education*, *37*(12), 1953–1973. doi:10.1080/09500693.2015.1058989

Deng, F., Chen, D.-T., Tsai, C.-C., & Chai, C. S. (2011). Students' views of the nature of science: A critical review of research. *Science Education*, *95*(6), 961–999. doi:10.1002/sce.20460

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

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, *88*(6), 915–933. doi:10.1002/sce.20012

Fillpot, E. (2012). Historical thinking in the third grade. *The Social Studies*, *103*(5), 206–217. doi:10.1080/00377996.2011.622318

Fitchett, P. G., Heafner, T. L., & Lambert, R. G. (2014). Examining elementary social studies marginalization: A multilevel model. *Educational Policy*, *28*(1), 40–68. doi:10.1177/0895904812453998

Ford, D. J. (2006). Representations of science within children's trade books. *Journal of Research in Science Teaching*, *43*(2), 214–235. doi:10.1002/tea.20095

Ford, M. J., & Forman, E. A. (2006). Redefining disciplinary learning in classroom contexts. *Review of Research in Education*, *30*, 1–32.

Garcia-Mila, M., Gilabert, S., Erduran, S., & Felton, M. (2013). The effect of argumentative task goal on the quality of argumentative discourse. *Science Education*, *97*(4), 497–523. doi:10.1002/sce.21057

Gillies, R. M., & Ashman, A. F. (1998). Behavior and interactions of children in cooperative groups in lower and middle elementary grades. *Journal of Educational Psychology*, *90*(4), 746–757. doi:10.1037/0022-0663.90.4.746

Gillies, R. M., & Khan, A. (2009). Promoting reasoned argumentation, problem-solving and learning during small-group work. *Cambridge Journal of Education*, *39*(1), 7–27. doi:10.1080/03057640802701945

Goldman, S. R., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., … Shanahan, C. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist*, *51*(2), 219–246. doi:10.1080/00461520.2016.1168741

Greene, S. (1994). The problems of learning to think like a historian: Writing history in the culture of the classroom. *Educational Psychologist*, *29*(2), 89–96. doi:10.1207/s15326985ep2902_4

Greenleaf, C., Schoenbach, R., Cziko, C., & Mueller, F. (2001). Apprenticing adolescent readers to academic literacy. *Harvard Educational Review*, *71*(1), 79–130. doi:10.17763/haer.71.1.q811712577334038

Groarke, L. (2017). Informal logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017.). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/spr2017/entries/logic-informal/

Heafner, T. L., & Fitchett, P. G. (2012). Tipping the scales: National trends of declining social studies instructional time in elementary schools. *Journal of Social Studies Research*, *36*(2), 190–215. Retrieved from http://proxy.lib.umich.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=90622745&site=ehost-live&scope=site

Herrenkohl, L. R. (2006). Intellectual role taking: Supporting discussion in heterogeneous elementary science classes. *Theory Into Practice*, *45*(1), 47–54. doi:10.1207/s15430421tip4501_7

Herrenkohl, L. R., & Cornelius, L. (2013). Investigating elementary students' scientific and historical argumentation. *Journal of the Learning Sciences*, *22*(3), 413–461. doi:10.1080/10508406.2013.799475

Hexter, J. H. (1971). *The history primer*. New York: Basic Books.

Hitchcock, D. (2007). Informal logic and the concept of argument. In D. Jacquette (Ed.), *Philosophy of logic* (pp. 101–129). New York: North Holland.

Hogan, K. (1999). Sociocognitive roles in science group discourse. *International Journal of Science Education*, *21*(8), 855–882. doi:10.1080/095006999290336

Hudicourt-Barnes, J. (2003). The use of argumentation in Haitian creole science classrooms. *Harvard Educational Review*, *73*(1), 73–93. doi:10.17763/haer.73.1.hnq801u57400l877

Hug, B., & McNeill, K. L. (2008). Use of first-hand and second-hand data in science: Does data type influence classroom conversations? *International Journal of Science Education*, *30*(13), 1725–1751. doi:10.1080/09500690701506945

International Baccalaureate. (2009). The Primary Years Programme: A basis for practice. International Baccalaureate Organization.

Iordanou, K. (2010). Developing argument skills across scientific and social domains. *Journal of Cognition and Development*, *11*(3), 293–327. doi:10.1080/15248372.2010.485335

Jeong, H., Songer, N. B., & Lee, S.-Y. (2007). Evidentiary competence: Sixth graders' understanding for gathering and interpreting evidence in scientific investigations. *Research in Science Education*, *37*(1), 75–97. doi:10.1007/s11165-006-9014-9

Jiménez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In S. Erduran & M. Jiménez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3–27). Dordrecht: Springer Science + Business Media B.V.

Jiménez-Aleixandre, M. P., Bugallo Rodríguez, A., & Duschl, R. A. (2000). “Doing the lesson” or “doing science”: Argument in high school genetics. *Science Education, 84*(6), 757–792. doi:10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F

Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching*, *41*(7), 748–769. doi:10.1002/tea.20020

Kelly, G. J., & Takao, A. Y. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. *Science Education*, *86*(3), 314–342. doi:10.1002/sce.10024

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*(4), 664–687. doi:10.3102/00028312027004664

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, *101*(1), 136–160. doi:10.1037/a0013097

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*(2), 155–179. doi:10.17763/haer.62.2.9r424r0113t670l1

Lee, C. D., Goldman, S. R., Levine, S., & Magliano, J. (2016). Epistemic cognition in literary reasoning. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 165–183). New York, NY: Routledge. doi:10.4324/9781315795225.ch11

Lee, J., & Weiss, A. (2007). *The nation's report card: U.S. history 2006 (NCES 2007–474)*. Washington, D.C.: U.S. Government Printing Office: U.S. Department of Education, National Center for Education Statistics. Retrieved from https://nces.ed.gov/nationsreportcard/pdf/main2006/2007474.pdf

Lee, O. (2017). Common core state standards for ela/literacy and next generation science standards: Convergences and discrepancies using argument as an example. *Educational Researcher*, *46*(2), 90–102. doi:10.3102/0013189X17699172

Losh, S. C., Wilke, R., & Pop, M. (2008). Some Methodological Issues with "Draw a Scientist Tests" among Young Children. *International Journal of Science Education*, *30*(6), 773–792. doi:10.1080/09500690701250452

Maloney, J. (2007). Children's roles and use of evidence in science: An analysis of decision-making in small groups. *British Educational Research Journal*, *33*(3), 371–401. Retrieved from http://www.jstor.org.proxy.lib.umich.edu/stable/30032616

Martin, D., Wineburg, S., Rosenzweig, R., & Leon, S. (2008). Historicalthinkingmatters.org: Using the web to teach historical thinking. *Social Education*, *72*(3), 140. Retrieved from https://link.galegroup.com/apps/doc/A178673161/AONE?u=umuser&sid=AONE&xid=65a6c1ff

Masnick, A. M., & Klahr, D. (2003). Error matters: An initial exploration of elementary school children's understanding of experimental error. *Journal of Cognition and Development*, *4*(1), 67–98. doi:10.1080/15248372.2003.9669683

Mason, C. L., Kahle, J. B., & Gardner, A. L. (1991). Draw-A-Scientist Test: Future Implications. *School Science and Mathematics*, *91*(5), 193–198. doi:10.1111/j.1949-8594.1991.tb12078.x

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their 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., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, *45*(1), 53–78. doi:10.1002/tea.20201

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

Mehan, H. (1979). "What time is it, Denise?": Asking known information questions in classroom discourse. *Theory Into Practice*, *18*(4), 285–294. Retrieved from https://www.jstor.org/stable/1476655

Mink, L. O. (1987). *Historical understanding*. Ithaca: Cornell University Press.

Monte-Sano, C. (2008). Qualities of historical writing instruction: A comparative case study of two teachers' practices. *American Educational Research Journal*, *45*(4), 1045–1079. Retrieved from http://www.jstor.org.proxy.lib.umich.edu/stable/27667162

Monte-Sano, C. (2010). Disciplinary literacy in history: An exploration of the historical nature of adolescents' writing. *Journal of the Learning Sciences*, *19*(4), 539–568. doi:10.1080/10508406.2010.481014

Monte-Sano, C. (2011). Beyond reading comprehension and summary: Learning to read and write in history by focusing on evidence, perspective, and interpretation. *Curriculum Inquiry*, *41*(2), 212–249. doi:10.1111/j.1467-873X.2011.00547.x

Monte-Sano, C., & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research*, *44*(3), 273–299. doi:10.1177/1086296X12450445

National Council for the Social Studies (NCSS). (2013). *The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K-12 civics, economics, geography, and history*. Silver Spring, MD: NCSS.

National Governors Association Center for Best Practices (NGA Center), & Council of Chief State School Officers (CCSSO). (2010). *Common core state standards*. Washington, D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.

National Research Council (NRC). (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Achieve, Inc. on behalf of the twenty-six states and partners that collaborated on the NGSS.

Nokes, J. D. (2010). Observing literacy practices in history classrooms. *Theory & Research in Social Education, 38*(4), 515–544. doi:10.1080/00933104.2010.10473438

Nystrand, M., & Graff, N. (2001). Report in argument's clothing: An ecological perspective on writing instruction in a seventh-grade classroom. *The Elementary School Journal*, *101*(4), 479–493. doi:10.1086/499683

Oliveira, A. W., Akerson, V. L., Colak, H., Pongsanon, K., & Genel, A. (2012). The implicit communication of nature of science and epistemology during inquiry discussion. *Science Education*, *96*(4), 652–684. doi:10.1002/sce.21005

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2009). A qualitative framework for collecting and analyzing data in focus group research. *International Journal of Qualitative Methods*, *8*(3), 1–21. doi:10.1177/160940690900800301

Osborne, J. (2014). Teaching Scientific Practices: Meeting the Challenge of Change. *Journal of Science Teacher Education*, *25*(2), 177–196. doi:10.1007/s10972-014-9384-1

Osborne, J., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education, 95*(4), 627–638. doi:10.1002/sce.20438

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

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*(2), 117–175. Retrieved from http://www.jstor.org/stable/3233567

Palincsar, A. S., & Magnusson, S. J. (2001). The interplay of first-hand and second-hand investigations to model and support the development of scientific knowledge and reasoning. In S. M. Carver & D. Klahr (Eds.), *Cognition and Instruction: Twenty-five Years of Progress* (1st ed., pp. 151–193). Mahwah, NJ: Psychology Press.

Parkinson, J., & Adendorff, R. (2005). Science books for children as a preparation for textbook literacy. *Discourse Studies*, *7*(2), 213–236. doi:10.1177/1461445605050367

Paxton, R. J. (2002). The influence of author visibility on high school students solving a historical problem. *Cognition and Instruction*, *20*(2), 197–248. doi:10.1207/S1532690XCI2002_3

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Towards a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (1 edition., pp. 99–122). Mahwah, NJ: Lawrence Earlbaum Associates.

Pontecorvo, C., & Girardet, H. (1993). Arguing and reasoning in understanding historical topics. *Cognition and Instruction*, *11*(3/4), 365–395.

Prado-Olmos, P. L. (1994). Exploring structure and interaction in small groups: An ethnographic study of cooperative group life in a bilingual elementary classroom. *Bilingual Research Journal*, *18*(3-4), 135–159. doi:10.1080/15235882.1994.10162672

Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, *30*(1), 86–112. doi:10.1080/07370008.2011.634081

Rice, D. C. (2002). Using trade books in teaching elementary science: Facts and fallacies. *The Reading Teacher*, *55*(6), 552–565. Retrieved from www.jstor.org/stable/20205097

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Charlotte, NC: Information Age Publishing.

Rouet, J.-F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, *88*(3), 478–493. doi:10.1037/0022-0663.88.3.478

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education*, *96*(3), 488–526. doi:10.1002/sce.21006

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

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, *12*(1), 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

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics* (Vol. 2, pp. 142–175). New York, NY: Cambridge University Press.

Seixas, P. (1999). Beyond 'content' and 'pedagogy': In search of a way to talk about history education. *Journal of Curriculum Studies*, *31*(3), 317–337. doi:10.1080/002202799183151

Seixas, P. (2015). A model of historical thinking. *Educational Philosophy and Theory*, 1–13. doi:10.1080/00131857.2015.1101363

Simosi, M. (2003). Using Toulmin's framework for the analysis of everyday argumentation: Some methodological considerations. *Argumentation*, *17*(2), 185–202. doi:10.1023/A:1024059024337

Stevens, R., Wineburg, S., Herrenkohl, L. R., & Bell, P. (2005). Comparative understanding of school subjects: Past, present, and future. *Review of Educational Research*, *75*(2), 125–157.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge [Cambridgeshire]; New York: Cambridge University Press. Retrieved from http://hdl.handle.net/2027/

Toulmin, S. (1958). *The uses of argument* (Updated.). Cambridge, U.K.; New York: Cambridge University Press.

Toulmin, S., Rieke, R. D., & Janik, A. (1979). *An introduction to reasoning*. New York: Macmillan.

van Drie, J., & van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review*, *20*(2), 87–110. doi:10.1007/s10648-007-9056-1

van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge, UK: Cambridge University Press.

von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, *45*(1), 101–131. doi:10.1002/tea.20213

Voss, J. F., & Means, M. L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction*, *1*(4), 337–350. doi:10.1016/0959-4752(91)90013-X

Walton, D. N. (2001). Enthymemes, Common Knowledge, and Plausible Inference. *Philosophy and Rhetoric*, *34*(2), 93–112. doi:10.1353/par.2001.0010

Walton, D. N., & Krabbe, E. C. W. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning*. Albany: State University of New York Press.

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, *91*(2), 301–311. doi:10.1037/0022-0663.91.2.301

Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*(1), 73–87. doi:10.1037/0022-0663.83.1.73

Wineburg, S. (1999). Historical thinking and other unnatural acts. *The Phi Delta Kappan*, *80*(7), 488–499.

Wineburg, S., & Martin, D. (2009). Tampering with history: Adapting primary sources for struggling readers. *Social Education*, *73*(5), 212–216.

Young, K. M., & Leinhardt, G. (1998). Writing from primary documents a way of knowing in history. *Written Communication*, *15*(1), 25–68. doi:10.1177/0741088398015001002

APPENDICES

APPENDIX 1: Surveys

Pre-Survey

Name Group

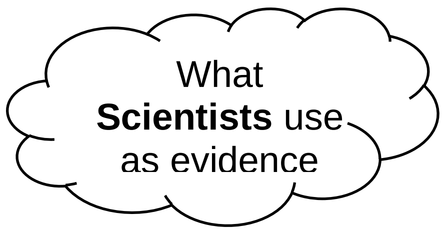
SCIENTISTS AND HISTORIANS

Think about how scientists do their job and how historians do their jobs.

How do **Scientists** do their job?	How are they **similar?**	How do **Historians** do their job?

SCIENTISTS

1. Scientists use evidence to do their work. What do you think **scientists** use as evidence? Make an idea web of different kinds of evidence **scientists** use:



2. Scientists use evidence to do their work. **Why** do you think **scientists** use evidence?

IB INQUIRY - SCIENCE CLASS

3. Do you ever use evidence in class during IB inquiry for **science**? ____
   _____

4. *If you do*, make an idea web of as many things as you can that you use as evidence in **science class**:

What **WE** use
as evidence in
Science Class

5. *If you do*, **why** do you think you use evidence in **science class**?

HISTORIANS

6. Historians use evidence to do their work. What do you think **historians** use as evidence? Make an idea web of different kinds of evidence **historians** use:

What **Historians** use as evidence

7. Historians use evidence to do their work. **Why** do you think

   Historians use evidence?

IB INQUIRY – SOCIAL STUDIES CLASS

8. Do you ever use evidence in class during IB inquiry for **social studies**? __________

9. *If you do*, make an idea web of as many things as you can that you use as evidence in **social studies class**:

What **WE** use as evidence in **Social Studies Class**

10. *If you do*, **why** do you think you use evidence in **social studies** class?

Post-Survey

Group_______________

*Think about the investigation you did in **Social Studies** how was it similar and different from what **historians** do?*

What **we** did	What we **both** do	What **historians** do

Think about the investigation you did in ***Science*** *how was it similar and different from what* ***scientists*** *do?*

What **we** did	What we **both** do	What **scientists** do

Think about the investigations you did in science and history. Were they the same in any ways? Were they different in any ways?
Ways our history and science investigations were the **SAME:**
Ways our history and science investigations were **DIFFERENT:**

APPENDIX 2: History Source Evaluation Worksheet

Source 1 – Results of the First Convention

<u>Headnote:</u> *This chart tells about the people who went to the first convention to vote on Michigan's statehood on* **September 28, 1836**. *Delegates from almost all of the counties in Michigan were there. The vote was very close. They decided not to become a state yet. People who did not want to let Ohio have Toledo voted against becoming a state.*

Item	Number
Number of counties present	20
Number of delegates present	49
Votes for accepting terms of statehood	21
Votes for rejecting terms of statehood	28

Source: Made from various sources by John-Carlos Marino, 2017.

Investigation Question:

Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?

What does this piece of evidence tell you about the investigation question?

___

___

___

___

___

___

___

___

Source 2 – Biography of Michigan's Governor	**Investigation Question:** Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?
Headnote: *This comes from a biography of the first Governor of Michigan, Stevens T. Mason. The author describes what happened after the first convention voted for Michigan not to become a state because of the fight over Toledo.*	What does this piece of evidence tell you about the investigation question?
After the first convention, it seemed like many people really wanted to have a new vote about becoming a state. Soon after, there were public meetings all through Michigan. These meetings were an example of democracy. People were meeting to decide what the State would do. In 1836 these citizens met in barns, schools, town halls, and village stores. They chose five men to get delegates for a new convention. They were successful. On December 14th another convention met, and seventy-one people were there.	
Source: *Stephens Thomas Mason: Misunderstood Patriot*, by Kent Sagendorph, pages 242-243, Published in 1947.	

Source 3 – Results of the Second Convention

<u>Headnote:</u> *This chart tells about the people who went to the second convention to vote on Michigan's statehood on* **December 14, 1836**. *There weren't as many counties at this convention as there were in the first one. Many in the Whig party did not think the vote should happen. So, some Whig counties did not attend. Almost all of the delegates at this convention were Democrats.*

Item	Number
Number of counties present	22
Number of delegates present	82
Votes for accepting terms of statehood	82
Votes for rejecting terms of statehood	0

Source: Made from various sources by John-Carlos Marino, 2017.

Investigation Question:

Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?

What does this piece of evidence tell you about the investigation question?

Source 4 – Report of the Democratic Meeting	**Investigation Question:** Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?
Headnote: This report comes from a newspaper article printed between the first and second conventions to vote on statehood. It describes a meeting the Democrats had. It tells what the Democrats decided to do after the first convention voted not to join the United States. It comes from a newspaper called "The Democratic Free Press." The newspaper was mostly read and written by Democrats. *We have decided*: The members of the Democratic party in Wayne county are worried about the dangers of voting against statehood for Michigan. We recommend meetings of our fellow citizens in every county. At the meetings people can show they want to call another convention and vote for Michigan to join the United States. **Source:** *Democratic Free Press* Nov. 16, 1836. "Report of the meeting of the Democratic Convention of Wayne County"	What does this piece of evidence tell you about the investigation question? ____

Source 5 – Book on Michigan History	Investigation Question:
Headnote: This comes from a history book about Michigan. The author teaches history at Michigan State University and ran a magazine called "Michigan History." It explains how Democrats called for a new election and that the Whigs were against that.	Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?
The Wayne County Democrats put out a notice that there would be another election. Each county would elect a new set of delegates. These delegates would meet to re-vote on Congress' offer for statehood. Not everyone agreed about having a second convention. Most Whigs were against the election and did not join in. Some counties did not send delegates. On December 14th eighty-two delegates met in Ann Arbor. The delegates voted for Michigan to become a state and agree that Toledo was in Ohio.	What does this piece of evidence tell you about the investigation question?
Source: *Michigan: Visions of our Past,* edited by Richard J. Hathaway, Published 1989	

Source 6 – Letter from the Governor

Headnote: *This comes from a letter written by Stevens T. Mason, the governor of Michigan at the time of the convention. He wrote to the leader of a meeting of a group of Democrats. He was also a Democrat. The letter was written after the first convention voted against statehood. In the letter, the Governor gave advice to the Democrats.*

The Congress says that the convention that decides about statehood should be a "convention of delegates elected by the people of the state." If the people really want to vote on statehood again then they should do it themselves. They should take the matter into their own hands and elect a new group of delegates from themselves. If their decision is from the people of Michigan, Congress will give Michigan statehood. If it is a majority of the people, Congress will give Michigan statehood.

Source: "Letter from Governor Stevens Thomas Mason to Ezekiel Pray, president of the meeting of the citizens of the county of Washtenaw." Printed in *Democratic Free Press* Nov. 16, 1836

Investigation Question:

Did the second vote for Michigan's statehood happen mostly because of popular opinion or mostly because of the work of the Democratic party?

What does this piece of evidence tell you about the investigation question?

APPENDIX 3: Science Source Evaluation Worksheets

Source 1 – Cart Weight Test

<u>Headnote</u>: *To get ready for building their push-cart each team did tests with different weights of push-carts. They got 4 cars with different weights and then pushed them with the 20 newtons of force. They timed how long it took each push-cart to travel 100 meters*

Push-Cart Weight	**Time**
80 kilograms	28 seconds
85 kilograms	29 seconds
90 kilograms	30 seconds
95 kilograms	31 seconds

Source: *4th Grade Engineering Journal*, by Karl, Edgar, Ebony, and Asia. 2017

Investigation Question:

Which car will win the race? The Blue Monster or the Green Bolt?

What does this piece of evidence tell you about the investigation question?

Source 2 – Engineer's Notes	**Investigation Question:** Which car will win the race? The Blue Monster or the Green Bolt?
Headnote: These are the notes of an engineer that makes push-carts as a hobby. She was testing her own push carts. She tried different combinations of force and weight. Today I tested two of my push-carts. Push-cart #1 weighs 71 kilograms. I pushed it with 19 newtons of force. Push-cart #2 weighs 98 kilograms. I pushed it with 24 newtons of force. I raced them against each other. Push-cart #1 beat push-cart #2 by almost 2 seconds. **Source**: *Personal Engineering Journal*. Andrea Hutchins. 2016	What does this piece of evidence tell you about the investigation question? ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

Source 3 – All About Air Resistance	**Investigation Question:** Which car will win the race? The Blue Monster or the Green Bolt?
Headnote: This comes from an encyclopedia article. It explains about air resistance.	What does this piece of evidence tell you about the investigation question?
When something like a car or plane is moving, air is flowing all around it. The moving air creates air resistance. Air resistance is a force that slows down an object. The shape of an object changes its air resistance. Something like a bus has a lot of air resistance. It has large flat parts that the air pushes against. Something like a race car has less air resistance. Its smooth curves let air go around it more easily.	
Source: *Encyclopedia of Physics*. 2010	

Source 4 – Edgar's Notes	**Investigation Question:** Which car will win the race? The Blue Monster or the Green Bolt?
Headnote: These are notes that Edgar made while he was building his push-cart. He is describing building the body of the cart. The body of the cart is the outside part that the driver sits in.	What does this piece of evidence tell you about the investigation question?
Designing this push-cart is tough! I am worried about the size of my push-cart. I don't know how to do woodworking very well. My cart has a flat front because I can't make curves. Asia's cart has smooth curves to it. I am worried that Asia's car will have less air resistance than mine because it is smoother and smaller.	
Source: *Push-Cart Design Journal.* By Edgar. 2017	

Source 5 – Push Force Test

Headnote: To get ready for building their push-carts the kids did tests with different forces on a push-cart. They got 1 push-cart that weighed 35 kilograms and pushed it four times with different amounts of force. Then they measured how fast each car traveled 100 meters

Push Force	Time
15 newtons	34 seconds
20 newtons	29 seconds
25 newtons	26 seconds
30 newtons	24 seconds

Source: *4th Grade Engineering Journal*, by Karl, Edgar, Ebony, and Asia. 2017

Investigation Question:

Which car will win the race? The Blue Monster or the Green Bolt?

What does this piece of evidence tell you about the investigation question?

Source 6 – Force, Weight, and Speed	Investigation Question: Which car will win the race? The Blue Monster or the Green Bolt?
Headnote: This comes from a book on cars. Cars can have different weights and make different amounts of force. It explains how weight and force can change speed.	What does this piece of evidence tell you about the investigation question?
Force and weight are two things that affect how fast cars go. Lighter cars often go faster than heavier cars. Cars pushed with more force often go faster than cars pushed with less force. If a heavy car is pushed with a lot of force it can go faster than a light car. If it is light enough, a car that is pushed with only a little force can go faster than a car pushed with more force.	
Source: *All About Cars*. 2000	

APPENDIX 4: Content Pre-Tests

Name:

History Investigation Pre-Test

1. Where is Toledo?

a. Michigan

b. Ohio

c. Indiana

d. Illinois

e. I don't know

2. What was the Toledo War?

a. A conflict between Ohio and Michigan over the city of Toledo

b. A war between the U.S. Army and the Toledo tribe

c. A war between the United States and the British

d. A conflict between the cities of Toledo and Ann Arbor

e. I don't know

3. Circle the one that became a state first:

Michigan

Ohio

I don't know

5. Circle all the names of political parties from the 1800s:

Democratic Party

Whig Party

Republican Party

Heritage Party

Name:

Science Investigation Pre-Test

1. What is a *newton*?

a. A tool for lifting things

b. A measurement of size

c. A measurement of force

d. A very small molecule

e. I don't know

2. What is *aerodynamics*?

a. Something that makes a machine stronger

b. How air moves around objects

c. A kind of scientist

d. Special kinds of planes

e. I don't know

3. Four friends ran a race. They all ran the same distance. They wrote their times in the chart below.

Number of batteries	**Time**
Angela	11 seconds
Cierra	10 seconds
Hannah	8 seconds
Gloria	9 seconds

Which friend ran the fastest?

a. Angela

b. Cierra

c. Hannah

d. Gloria

e. I don't know

4. Two cars are racing each other. Car #1 has one person in it. It weighs 2,500 pounds. Car #2 has 5 people in it. It weighs 3,100 pounds. Which one will probably win the race?

a. Car #1 will probably win

b. Car # 2 will probably win

c. It will probably be a tie

d. I don't know

5. A push or pull on an object is called …

a. inertia

b. mass

c. force

d. acceleration

e. I don't know

APPENDIX 5: Interview Protocols

Interviews 1 and 2 – Evaluating Evidence

1. Prior to the interview prepare the following:
   - Evidence sheets for each group
   - Blank evidence sheets
   - Evidence value sheets
   - A photograph of each student's SenseMaker Board

2. Introduction
   - *Hello, thank you again for all of your hard work yesterday. Today I'm going to be asking you all a few questions about the work that you did in the group yesterday. I really want to know what you were thinking about when you working in the group. There are no wrong answers to these questions, so please don't be afraid to say whatever it is you think about when I ask you these questions.*
   - *I also want you to really listen to each other's answers. We're going to have a conversation, so it's really important that you listen to each other's ideas so you can think about what other people say and think about how you agree or disagree.*
   - *I'm going to give you a photograph of your group's SenseMaker board to help you remember what you came up with yesterday.*

3. Review SenseMaker board
   - *First, I want you to look at your answers to the investigation question* Repeat question.
   - *Take a minute to look over it.*
4. Review Evidence
   - *Now I want you to look at each piece of evidence. As I reread each piece of evidence, think about how useful it was to answer the investigation question. Then circle your answer on the sheet.*

   Most important evidence
   - *First, we're going to talk about the evidence that you thought was really important. Go through and put a star next to the piece of evidence you thought was most important.*
   - *Who thought* PIECE OF EVIDENCE *was most important?*
   - *[Student 1], you said this evidence was the most important. Why did you think it was most important? How did you use it in your answer?*
   - *[Students 2, 3, and 4], you rated this piece of evidence as [rating]. What do you think about the way that [Student 1] used the evidence? Did you use [piece of*

evidence] to support your explanation? How important did you think it was? Why?

- Possible follow-up questions:
    i. *On the SenseMaker board, you said that this evidence connects to the claim [claim] because [reasoning]. Can you tell me more about that?*
- Repeat for each student.

5. Least important evidence
    - *Now I want you to think about a piece of evidence that you didn't think was important to your explanation. Put a star next to the one that was least important.*
    - *Who thought* PIECE OF EVIDENCE *was least important?*
    - *[Student 1], you said this evidence was the least important. Why did you think it was least important? How did you use it in your answer?*
    - *[Students 2, 3, and 4], you rated this piece of evidence as [rating]. What do you think about the way that [Student 1] used the evidence? Did you use [piece of evidence] to support your explanation? How important did you think it was? Why?*

6. Other evidence
    - *Finally, I want you to think about the evidence that you DIDN'T have. If you could have any piece of evidence to help you answer the question, what would it be?*
    - *[Student 1] what do you think?*
    - Repeat for each student.
    - *Now that you've heard each other's ideas what do you think about them? What do you think about [Student 1's] idea? What do you think about [Student 2's] idea?* Repeat for each student
    - Possible follow up questions and prompts if students are stuck
        i. *What would the evidence look like? Would it be the most like…* provide an example of each type of evidence.
        ii. *What would you want the evidence to tell you about?*
        iii. *How would that evidence help you?*

7. Conclusion
    - *That is all the questions I have for you. Thank you so much for helping me. You worked really hard on all of this, and you should be proud of the work that you did. Do you have any questions?* Answer questions.

Thank you!

Interview 3 – Connecting the disciplines and investigations

1. Prior to the interview prepare the following:
    - Collect all of the students' pre-assessments about what counts as evidence in science, history, and school.
    - Collect all of the students' response sheets about the connections between the disciplines and the investigations from the day 9 discussion.

2. Introduction
    - *Hello, thank you again for all of your hard work during these investigations! Today I'm going to be asking you all a few questions about the work that you did during both the science and history investigations. There are no wrong answers to these questions, so please don't be afraid to say whatever it is you think about when I ask you these questions.*
    - *I also want you to really listen to each other's answers. We're going to have a conversation, so it's really important that you listen to each other's ideas, so you can think about what other people say and think about how you agree or disagree.*

3. Comparing the investigations to history
    - *First, we're going to talk about the history investigation that we did. I'm going to pass you back these papers that you wrote on at the beginning of the investigations and yesterday. Look at the one called "Comparing our work to historians and scientists." Read to yourself what you wrote.*
    - *Now I'm going to read your answers out loud, so you can all talk about them. [Student 1], you wrote* [read response for comparisons to historians].
    - *[Students 2, 3, 4, 5] what do you think about that? Do you agree with what [Student 1] said? Do you disagree?*
    - Repeat for each student

4. Comparing the investigations to science
    - *Now, we're going to talk about the science investigation that we did*
    - *I'm going to read your answers out loud, so you can all talk about them. [Student 1], you wrote* [read response for comparisons to scientists].
    - *[Students 2, 3, 4, 5] what do you think about that? Do you agree with what [Student 1] said? Do you disagree?*
    - Repeat for each student

5. Comparing the investigations
    - *Now we're going to move on to what you thought about the science and history investigations. I really want you to think about what was the same and what was different about what you did during the two investigations. To remind you of what you said yesterday, here is your "Comparing the Investigations" sheet. Please read it to yourself to remember what you wrote.*
    - *I'm going to read your answers out loud, so you can all talk about them. [Student 1], you wrote* [read response for comparisons to scientists].

- *[Students 2, 3, 4, 5] what do you think about that? Do you agree with what [Student 1] said? Do you disagree?*
- Repeat for each student
- Possible follow up questions:
    i. *What about the way you used evidence?*
    ii. *What about the kind of evidence you used?*
    iii. *What about the thinking you had to do to connect the claim with the evidence?*

6. Conclusion
    - *Ok everyone, you've been really, really helpful! Thank you so much for all of your hard work and letting me talk to you during this job. Great work!*