ABSTRACT

Title of Dissertation:	USING DISCOURSE TO IMPROVE THE QUALITY OF STUDENT TALK AND HISTORICAL ARGUMENTATIVE WRITING
	Josue Otarola, Doctor of Philosophy, 2024
Dissertation directed by:	Dr. Susan De La Paz, Counseling, Higher Education, and Special Education

Frameworks that connect to the Common Core State Standards for English Language Arts and Literacy in Social Studies, such as the College, Career, and Civic Life (C3) Framework for Social Studies States Standards, highlight the need to engage in inquiry-based instruction (NCSS, 2013). Participation in such inquiry requires students to engage in disciplinary thinking and to articulate that thinking to others, both verbally and in writing. However, such disciplinary thinking does not come natural to students (Wineburg, 1991). Thus, students require instruction in disciplinary thinking to learn its complexities and nuances. Once students can engage in disciplinary thinking, they can communicate it and participate in valuable discourse. Therefore, the current dissertation was conducted to explore how students use discourse to engage in argumentation and historical thinking.

Chapter 2 of the dissertation is a research synthesis of studies that use discourse to improve learning outcomes in primary and secondary science and social studies classrooms. The purpose of the synthesis was to determine the impact of argumentative discourse on students'

learning outcomes and to understand the instructional components teachers use when holding discourse. Asterhan & Schwarz's (2016) Argumentation for Learning (AFL) framework guided the research synthesis and the subsequent multiple-case study. Results indicate that discourse can be improved by using multiple instructional groupings, incorporating explicit instruction, modeling, graphic organizers and technology, and engaging students in deliberation.

Chapter 3 offers findings from a multiple-case study that was designed to explore how argumentation inhibitors and enablers moderate dialogue characteristics and learning outcomes and to provide a rich description of discourse in ninth-grade US History classrooms with academically diverse students. More specifically, the study captured how students engaged in argumentative discourse and historical thinking using two different discourse structures. The study used a cross-case analysis (Yin, 2018) to compare the discourse across three cases. Each case included a teacher and four students. The first case occurred in a co-taught class, the second case included the same teacher in an honors class, and the third case included a different teacher in an honors class. The first and second case used a modified structured academic controversy (SAC), while the third case used Johnson and Johnson's (1988) approach to SAC.

The multiple-case study and the research synthesis informed the practitioner manuscript provided in Chapter 4. The manuscript details how teachers can use structure and supports to improve student participation and historical thinking in classroom discourse, especially for students with disabilities (SWD) and other struggling learners.

The current dissertation provides several important findings. First, my synthesis indicated that students achieve higher learning outcomes when teachers use multiple instructional groupings, students engage in deliberative discourse, and teachers provide students with explicit instruction, modeling, and graphic organizers. Second, the findings from the multiple-case study

offered insight into how students of differing academic abilities engage in argumentative discourse and historical thinking. Students of all academic abilities participated at high levels and engaged in deliberative argumentation, though there were differences in the quality of historical thinking skills. The instructional approach used in the multiple-case study is further expanded in the practitioner manuscript. Areas for future research are discussed in the dissertation.

USING DISCOURSE TO IMPROVE THE QUALITY OF STUDENT TALK AND HISTORICAL ARGUMENTATIVE WRITING

by

Josue Otarola

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Susan De La Paz, Chair
Professor Jade Wexler
Associate Professor Sarah McGrew
Associate Professor Andrew Brantlinger
Associate Professor Jeffery Nokes

© Copyright by
Josue Otarola
2024

Dedication

To my wife and rock, Daniela, and to my two children, Joaquín and Elena.

Acknowledgements

I first want to thank my dissertation chair, Susan De La Paz, for being a wonderful advisor. You chose me as your advisee and guided me every step of the way in pursuit of my PhD. You pushed me since I started in 2020 to become a better writer, researcher, and speaker, and you provided me with opportunities to present at many conferences, to teach at the university level, and to publish. You gave me the confidence and support when I needed it, and I will always be thankful.

I want to thank my dissertation committee members: Jade Wexler, Sarah McGrew, Andrew Brantlinger, and Jeffery Nokes. Thank you for challenging me and giving me confidence to engage in high-quality scholarship. I am very grateful to all of you for showing your interest in me and providing me the guidance I needed to finish my dissertation.

Thank you to my fellow doctoral students who supported me along the way and made the experience more enjoyable. Sara, Melissa, Cam, Pearl, and Max, we shared many laughs and a few tough times together, but we always persevered. I especially want to thank Sara for her moral support, sharp wit, and baby-sitting when Daniela and I needed it.

To my students and colleagues at Buck Lodge Middle School and MacFarland Middle School, thank you for reminding me every day of the importance of education and of educators. Mrs. Robin Maher, you were my first co-teacher and my mentor, and I learned so much from you. From professionalism to excellent instruction, I am thankful for all that you taught me.

I want to especially thank my parents, Lawrence and Xinia, who engrained the love of learning in me from infancy. I am grateful for your unconditional love and support. To my in-laws, Cesar and Lisette, you were always supportive and made sure that Joaquín and Elena were taken care of while I was busy finishing my dissertation.

To my loving wife and best friend, Daniela. You encouraged me to embark on this journey and supported me throughout. You took care of Joaquín and Elena when I was busy with schoolwork, even after spending all day working. Your love and sacrifice have made this possible, and I share this achievement with you. We somehow made this work, and now we can look back at some of the rough moments knowing it was worth it. I could not ask for a more loving partner and parent, and I will always be forever grateful.

Finally, I want to thank God for His enduring love and faithfulness to my family and me.

Table of Contents

Dedication ... ii
Acknowledgements ... iii
Table of Contents ... v
List of Tables ... viii
List of Figures ... ix
Chapter 1: Introduction ... 1
Introduction ... 1
Outline of the Dissertation ... 5
Definition of Key Terms ... 8
Chapter 2: The Influence of Discourse Goals, Instructional Tools, and Grouping Structures on Students' Learning and Reasoning in Science and Social Studies ... 11
Theoretical Framework ... 14
Current Study ... 15
Research Questions ... 17
Method ... 18
Search Procedure ... 18
Inclusion and Exclusion Criteria ... 19
Coding Procedures ... 19
Calculation of Effect Sizes ... 20
Results ... 21
RQ1. Does type of group formation (i.e., pairs, small groups, large group) affect student learning outcomes in science and social studies K-12 classrooms? ... 22
RQ2. What process supports have teachers utilized during discourse in science and social studies K-12 classrooms, and which show better results? ... 27
RQ3. How do dialogue characteristics (i.e., disputative or deliberation) affect student learning outcomes in science and social studies K-12 classrooms? ... 31
Discussion ... 32
Limitations ... 40
Implications for Practice ... 40
Future Research ... 41
Conclusion ... 42
Chapter 3: Effects of Structured Controversy with and without Guided Support on Academically and Linguistically Diverse Students' Historical Thinking ... 43
Disciplinary Discourse and Writing ... 45
Theoretical Framework ... 47
Current Study ... 49
Method ... 51
Research Design ... 51
Setting and Participants ... 53
Cresswell High School (CHS) ... 53
Teachers ... 54
Students ... 55
Co-Taught ... 55

Honors 1 56
Honors 2 56
Task Design 57
Structured Academic Controversy 58
Guided Structured Academic Controversy 59
Data Collection 61
Data Analysis 61
Results 64
RQ1: What is the quality of students' historical thinking during discussions, and do students engage in deliberative or disputative argumentation? 64
Similarities in Historical Thinking Between Co-taught and Honors 1 65
Differences in Historical Thinking Between Co-taught and Honors 1 66
Similarities in Argumentation Between Co-taught and Honors 1 68
Differences in Argumentation Between Co-taught and Honors 1 73
Similarities in Historical Thinking Between Honors 1 and Honors 2 73
Differences in Historical Thinking Between Honors 1 and Honors 2 75
Argumentation in Honors 1 and Honors 2 77
RQ2: What can we infer about students with different incoming academic and cognitive profiles? 80
Academic and Cognitive Differences Among Cases 81
Discussion 85
Limitations 90
Implications for Practice 90
Recommendations for Future Research 91
Chapter 4: Improving Students with Disabilities and Other Struggling Learners' Argumentative Discussions with Guided Support for Structured Academic Controversies 93
Structured Academic Controversy 96
Historical Investigation 97
Guided Structured Academic Controversy 98
Preparation 99
Warm Up 99
Step 1: Identify Perspectives 99
Step 2: Initial Hypothesis 100
Step 3: Find Evidence 100
Steps 4 and 5 Evaluate Each Document 101
Step 6: Whole Class Discussion 102
Planning to Write 102
Student Discussion in Action 103
Teacher Reflection 107
Conclusion 108
Chapter 5: Conclusion 109
Research Synthesis 110
Multiple-Case Study 111
Practical Implications 115
Future Directions 117

Conclusion 118
Appendix A 120
Appendix B 125
References 157

List of Tables

Table 2.1 *Description of Studies*
Table 2.2 *Demographics of Studies*
Table 2.3 *Effect Size by Treatment Condition*
Table 3.1 *Student Demographic Information*
Table 3.2 *Selected Student Characteristics*
Table 3.3 *Overview of the Six Historical Investigations*
Table 3.4 *Sourcing Historical Judgment Rubric*
Table 3.5 *Contextualization Historical Judgment Rubric*

List of Figures

Figure 2.1 *Prisma Diagram*
Figure 3.1 *The Three-node Argumentation for Learning*
Figure 3.2 *SAC+G Graphic Organizer*
Table 3.3 *Argumentative Utterances in Co-taught and Honors 1*
Table 3.4 *Argumentative Utterances in Honors 1 and Honors 2*
Table 3.5 *Historical Thinking Utterances in Co-taught and Honors 1*
Table 3.6 *Historical Thinking Utterances in Honors 1 and Honors 2*
Table 4.1 *SAC Graphic Organizer (De La Paz et al, 2014; 2017; Monte-Sano et al, 2017)*
Table 4.2 *Student Sample of SAC+G Graphic Organizer*

Chapter 1: Introduction

Introduction

Students engage in discourse every day, whether in conversation or for learning purposes in classrooms. Moreover, discourse has demonstrated to improve learning outcomes (Bennet et al., 2010; De La Paz & Wissinger, 2017; Erduran et al., 2024; Gotwals et al., 2022; Murphy et al., 2017). Discourse is referred to as discussion, talk, and dialogue, though we use the term discourse as it emphasizes the underlying thinking behind the act of speaking. In education, leading a discussion is considered a high leverage practice (HLP), and education programs instill the importance of education in teacher candidates (Brownell et al., 2019; Capobianco et al., 2018; TeachingWorks, 2023). This expectation has also been emphasized by several curriculum and policy initiatives linking discourse to student literacy. For example, a report from the Institute of Educational Sciences (IES) recommended the use of classroom discussion to improve adult literacy (Kamil et al., 2008).

Of course, not all discourse is the same. In school, teachers leverage the natural act of discussion and attempt to embed it in their instruction for multiple reasons. For example, in secondary subjects, teachers use discourse as opportunities for students to learn content, and how to think about academic disciplines (Larson, 2021). Therefore, while the way discourse is implemented in K12 classrooms varies (O’Connor & Snow, 2018) teachers frequently use discourse as opportunities for

students to consider differing ideas and learn core content, and they use multiple approaches to integrate it into their instruction.

One approach to argumentative discourse is the structured academic controversy (SAC), whereby students discuss controversial topics in small groups, encounter conflicting ideas and opinions, and collaboratively resolve the controversy (Johnson & Johnson, 1988). The SAC is a powerful means to incorporate discussion because it offers both a discourse structure and an instructional procedure (Parker, 2021). A Socratic seminar, for example, is a discourse structure but not necessarily an instructional procedure. Furthermore, one of the goals of SAC is to ensure all students participate by holding each individual in a group accountable for their learning (Parker, 2021).

Johnson and Johnson (1988) propose framing academic controversies as problems requiring solutions rather than win-lose scenarios. They created a routine to structure students' approach to deliberation (i.e., discussing, advocating, analyzing, evaluating, and rebutting information) to reach valid factual and judgmental conclusions. To be successful, such discussions require instructional support to build on students' prior knowledge, reading comprehension, and effective discourse; it also requires teachers' management of small group dynamics (Parker, 2021). A SAC also requires cooperative learning, which can be difficult to accomplish without the proper structure and support. Furthermore, there are also more practical impediments, such as taking the time to gather the necessary materials. Moreover, given diversity in students ability level and motivation, teachers may only hear from a vocal minority when holding class discussions.

All students should build the skills necessary to engage in productive discourse regardless of academic, linguistic, or cultural differences. This is challenging for teachers as research indicates that struggling learners, including students with disabilities (SWD) are less likely to engage in classroom discussions (Bueso, 2022; Louick & Wang, 2021). Given that nearly 75% of students with learning disabilities (LD) in the United States are educated within the general education classroom (U.S. Department of Education, 2022), general education teachers must learn strategies to effectively engage these students in discussion.

The College, Career, and Civic Life Framework for Social Studies State Standards (C3 Framework) recommends that teachers engage students in inquiry that promotes thoughtful discourse about debatable and multifaceted questions (NCSS, 2013). And the Next Generation Science Standards (NGSS) promote using discourse to grapple with complex problems (NRC, 2013). Both the C3 Framework and the NGSS are well-founded, considering the potential benefits of incorporating discourse. Moreover, these standards are achievable because studies show the efficacy of discourse for special education (McLeskey, 2017) and general education (Grossman et al., 2009) students.

Given that only 27% of 8th and 12th graders wrote at the proficient level on the most recent National Assessment of Education Performance (NAEP), there is a great need to improve writing (NCES, 2022). Prior research indicates that students can use argumentative discourse to help improve their argumentative disciplinary writing in science (Chen et al., 2016) and in the social studies (Wissinger & De La Paz, 2016). History, for example, requires students to engage in historical thinking.

Students use historical thinking to make arguments about a historical question, and they must use evidence to support their claims. One way to engage students and promote critical thinking is by using critical questions. Studies established that discourse that incorporates critical questions can improve written historical arguments (De La Paz & Wissinger, 2017; Wissinger & De La Paz, 2016) but more should be done to determine how best to do so.

The purpose of this dissertation was to explore how teachers can structure discussion to promote high-quality discourse between students, including academically diverse learners. I first conducted a research synthesis of science and social studies interventions that integrated discourse to affect learning outcomes, specifically disciplinary knowledge, disciplinary writing, argumentation structure, and holistic writing (Chapter 2). I adopted Asterhan and Schwarz's (2016) Argumentation for Learning (AFL) theoretical framework due to its helpful categorization of argumentation inhibitors and enablers, dialogue characteristics, and learning outcomes and explore the literature on discourse in two domains.

My overarching results indicated that interventions incorporating multiple instructional groupings (e.g., small groups, whole class) alongside a combination of explicit instruction, modeling, and graphic organizers using structured deliberative discourse, were most effective. These are examples of AFL's Task Design and Process support, respectively. However, researchers rarely provide a fine-grained analysis of what happens during discourse, and more information is needed to determine how social studies teachers should engage academically and linguistically diverse students in deliberative discourse.

Based on the findings in the research synthesis, I conducted a multiple-case study to explore how two purposefully selected ninth-grade US History teachers and their students engaged in discourse through a SAC and a revised, more supportive SAC (SAC+G) in co-taught and honors classrooms (Chapter 3). I adopted AFL as my theoretical framework for this study as the three-nodes (inhibitors and enablers of argumentation, dialogue characteristics, and learning outcomes) can be used to study the effects of the process supports that are embedded in both approaches to SAC. Based on the findings of the latter study, I wrote a practitioner manuscript offering detailed guidance on why and how teachers can use the SAC+G to improve participation and quality in class discussions (Chapter 4). Thus, the current dissertation makes important contributions by informing general and special educators of practical ways to facilitate robust disciplinary thinking about historical controversies, and to propose additional areas for future research.

Outline of the Dissertation

This dissertation is organized into five major sections: Introduction, Research Synthesis, Multiple-case Study, Practitioner Manuscript, and Conclusion. The current chapter (Chapter 1) introduces the problem and provides an overview of what this dissertation aims to address. Chapter 2 is a synthesis of studies that use discourse to improve learning outcomes in primary and secondary science and social studies. The research questions for the systematic review are:

1. Does type of group formation (i.e., pairs, small group, large group) affect student learning outcomes in science and social studies K-12 classrooms?

2. What process supports have teachers utilized during discourse in science and social studies K-12 classrooms, and which show better results?
3. How do dialogue characteristics (i.e., disputative or deliberative) impact student learning outcomes in science and social studies K-12 classrooms?

Chapter 3 is a multiple-case study that extends the findings of the research synthesis of discourse studies in Chapter 2. This study utilized cross-case analysis (Yin, 2018) to compare three cases. Each case contained one teacher and four students. One teacher taught the first case, which was a co-taught class (referred to as Mod), and the second case, which was an honors Mod. A second teacher taught the third case, a different honors Mod, and this case used Johnson & Johnson's (1988) SAC rather than the one that was used in the first and second case. While the instructional structures (i.e., SAC and SAC+G) were developed by others, my work provides unique and important information about the resulting discourse characteristics and students' historical thinking learning outcomes. The research questions for the multiple-case study are the following:

1. What is the quality of students' historical thinking during discussions?
    a. What types of dialogue characteristics are evident in small and large group discussions?
    b. How do students evaluate sources/make judgments?
2. What can we infer about students with different incoming academic and cognitive profiles?

We found that student participation was high across all cases, and teacher guidance was important for student engagement and understanding of the SAC

process. Moreover, students demonstrated high quality argumentation and historical thinking skills with two exceptions. First, students in the honors SAC minimally addressed sourcing, in comparison to both SAC+G cases. Moreover, students in the second case, the co-taught class who used the SAC+G, engaged in sourcing at a higher level. Conversely, students in the co-taught class had lower quality historical judgments than their peers in the honors class. While student academic and cognitive profiles appeared to impact historical thinking, the two SWD participated at similar or higher rates than their non-disabled peers. This was an unexpected yet welcomed finding, as prior research has shown that SWD participate at lower levels than their non-disabled peers (Bueso, 2022; Louick & Wang, 2021).

Chapter 4 provides a practitioner manuscript to inform teachers of the rationale for using a revised SAC (i.e., SAC+G) and to guide teachers through the steps needed to implement it with academically diverse students, including SWD. The manuscript details the benefits of using a SAC to explore historical controversies. Moreover, the manuscript outlines specific structures and supports for teachers to use with their students to improve student engagement and quality in classroom discussions. I include sample excerpts of student and teacher discourse, a SAC+G graphic organizer completed by a student, and a teacher's feedback of her experience using the SAC+G. Thus, the practitioner manuscript provides teachers with a clear rationale, recommended step-by-step directions, and an example they can implement to support student discourse in history classrooms.

I conclude with Chapter 5, which reviews the findings and implications of the research synthesis and the multiple-case study. In addition, the chapter synopsizes the

guidance presented in the practitioner manuscript. The chapter concludes with research questions that can be addressed in future studies extending the current dissertation. Lastly, the research synthesis, the multiple-case study, and the practitioner manuscript are presented as individual manuscripts. This is to reflect that each paper will be submitted for publication as standalone manuscripts, and I provide individual titles for each of these chapters.

Definition of Key Terms

Argumentation for Learning: A framework proposed by Asterhan & Schwarz (2016) that emphasizes the importance of understanding how different factors influence the quality and effectiveness of argumentative discourse to enhance educational practices and learning gains. It identifies three key components: inhibitors and enablers, dialogue characteristics, and learning outcomes.

Cognitive apprenticeship (CA): A cognitive apprenticeship is an approach to learning that reveals otherwise hidden cognitive processes to students and gradually promote independence (Collins, Brown, & Holum, 1991). De La Paz and her colleagues use CA to teach historical concepts and reading strategies and argumentative writing (2010, 2014, 2016, 2021). Teachers model historical reading and writing processes, explicitly teaching students to make credible judgments about authors and evidence. Scaffolding includes tools (called IREAD and How to Write) for students while reading, planning, and writing, and tools for teachers to monitor students' progress (Monte-Sano, De La Paz, & Felton, 2014a, 2014b).

Cross-case analysis: A data analysis approach involves systematically comparing multiple cases to identify patterns and differences, which enhances the depth and robustness of the study's findings (Yin, 2018). This method allows researchers to generate more comprehensive and generalized insights by examining how each case varies and aligns within the broader context of the research questions.

Deliberative argumentation: Argumentative dialogue in which the goal is to arrive at a viewpoint by comparing and evaluating alternatives collaboratively (Felton et al., 2009).

Disputative argumentation: Argumentative dialogue in which the goal is to defend a viewpoint and undermine alternatives (Felton et al., 2009).

Guided Structured Academic Controversy: A turn and talk routine that breaks down how students explore, comprehend and evaluate sources, with a graphic organizer that provides structure for writing ideas and quotes. It adds structure to the original approach for exploring academic controversies.

Historical thinking: The process of critically analyzing and interpreting past events and contexts involving a set of disciplinary skills (e.g., sourcing, contextualization, corroboration) that enables individuals to understand history (Nokes, 2017).

Multiple-case study: A methodology that provides a comprehensive framework for conducting in-depth investigations across multiple cases, allowing researchers to explore differences and similarities within and between cases (Yin, 2018). This approach enhances the robustness of findings by offering detailed contextual analyses and facilitating the development of theoretical insights.

Structured Academic Controversy (SAC): A process in which students can collaboratively make well-formed arguments about complex and contentious issues after examining at least two perspectives (Johnson & Johnson, 1988).

Chapter 2: The Influence of Discourse Goals, Instructional Tools, and Grouping Structures on Students' Learning and Reasoning in Science and Social Studies

Students engage in discourse every day, whether in conversation or for learning purposes in classrooms. Lecture, whole-class discussion, and small-group discussions are common ways to support learning (O'Connor and Snow, 2018). Educators use discourse for good reason. It is now recognized as a powerful teaching practice (Bae et al., 2021; Erduran et al., 2004; Gotwals et al., 2022; Murphy et al., 2009) that allows students to build collective knowledge (TeachingWorks, 2023). For example, it facilitates academic learning (Andriessen & Baker, 2014; Engle & Conant, 2002; Nystrand et al., 1997). Discourse is seen as a high-leverage practice (HLP) that teacher candidates practice (Brownell et al., 2019; Capobianco et al., 2018; O'Connor & Snow, 2018). Moreover, it has been found effective in general (Grossman et al., 2009) and special education classrooms (McLeskey, 2017).

With respect to academic learning, the Next Generation Science Standards (NGSS) and the College, Career, and Civic Life Framework (C3 Framework), recognize that classrooms are ideal settings for using discourse as a sense making tool (National Council for the Social Studies, 2013; National Research Council, 2013). For example, students have been observed to use causal inferencing when engaged in inquiry (Ford, 2008; Hynd, 1999). In science, students discuss ways to collect and analyze empirical data to support a hypothesis or determine causal mechanisms (Chen et al., 2016; Ford 2008). And in history, students read and discuss multiple perspectives to discern significance of events (Nokes, 2017; van Drie & van Boxtel

2018). In both disciplines, it is commonly expected that students should discuss their thinking to understand disciplinary concepts.

Though discourse takes place across a variety of academic subjects such as English Language Arts (Reynolds, 2018) and mathematics (Zolkower & Shreyar, 2007), we limit the scope of our study to science and social studies. Science and social studies are both well studied, so we will be able to see how disciplinary thinking and reasoning affects argumentation across disciplines. These particular disciplines require students to engage in inquiry, which demands discourse as part of the learning process (Hackling et al., 2010; Monte-Sano et al., 2021). Moreover, both subjects have a variety of disciplines adept for inquiry, such as biology and physics in science, and history and civics for social studies.

Experts engage in argumentative discourse in different ways within various disciplines (Hyland, 2014). For example, in science and history, what counts as evidence, how evidence is evaluated, and how it is used varies (Chen et al., 2016; Monte-Sano, 2014; Reisman, 2012) due to different epistemological foundations and discourse practices (Cervetti et al., 2014; Herrenkohl & Cornelius, 2013). Scientists justify how evidence supports a claim, using mechanistic reasoning when linking evidence and claim (Russ et al., 2008), for example, and historians use causal inferencing to weigh evidence connected to a claim (Nokes et al., 2007). Because students emulate experts, teachers must also learn these disciplinary skills.

In school, two types of discourse that have been widely used to promote reasoning in science and social studies are disputative and deliberative argumentation (Felton et al., 2009; Gronostay, 2016). Both involve presenting alternative viewpoints

by providing claims and supporting evidence then evaluating claims and evidence. Whereas dispute often leads students to combative discourse, deliberation allows individuals to compare and evaluate alternative viewpoints respectfully with the goal of arriving at a consensus. When asked to persuade fellow classmates, students become more competitive and their arguments suffer (Felton et al., 2009). This may be explained by considering the goal of disputative argumentation, which calls for convincing peers of "correct" findings and defending them. In contrast, deliberation calls for collaborating to arrive at a consensus where differences are presented, discussed, and resolved. Deliberation in Felton et al.'s (2009) study (and in more recent research; Asterhan & Schwarz, 2016; Garcia-Mila et al., 2013) led to improved student writing. Clearly, deliberation is a skill that is important for students to learn.

Expectations such as these create challenges for learners, especially those who are from academically, culturally, or linguistically diverse backgrounds. Therefore, educators must be equipped to impart these skills to all students, irrespective of their backgrounds. Students with disabilities (SWD) and multilingual learners (ML) make up a large proportion of the student population. According to the National Center for Education Statistics (NCES, 2024), the percentage of ML in the K-12 population grew from 9.2% in the fall of 2010 to 10.3% in the fall of 2020. That number grows as immigration continues to rise. Similarly, 15% of students receive special education services, with the most disability categories of struggling learners including those with specific learning disabilities (SLD), attentional challenges (ADHD), and speech or language impairments (SLI) that affect learning and social interactions in classrooms. These challenges may cause students to be less willing or struggle to

participate in discourse (Bueso, 2022; Lambert & Sugita, 2016), especially if they have not mastered English (Boyd & Rubin, 2002), leading to additional instructional challenges for teachers.

Theoretical Framework

Literature on argumentation for learning has shown that writing arguments can improve students' conceptual understanding and engagement in science (Akkus et al., 2007; Chen et al., 2016; Nam et al., 2008), and history (Schwarz, 2009). Previous literature also demonstrates that small group discussion allows students to engage in more purposeful dialogue and increase understanding (Bennett et al., 2010; Brooks, 2009). Education researchers find that carefully scaffolded discussion in both subjects functions as bridges between every day and disciplinary arguments, if students learn to critically evaluate evidence (Chen et al., 2016; Firetto et al., 2019; Reisman, 2012; Wissinger & De La Paz, 2016).

These and other reviews (Bae et al., 2021; Gotwals et al., 2022; Murphy et al., 2009) emphasize the importance of providing explicit instruction to hold effective classroom discussions. However, we lack detailed outcomes regarding the effects of discourse on academic content and disciplinary learning. Murphy et al.'s early (2009) meta-analysis indicated very few (of 42) discourse approaches were effective at increasing both (a) students' literal or inferential comprehension and (b) their critical thinking and reasoning about text. To our knowledge, no more recent synthesis has provided this information, thus we aim to provide more details in the current review.

Asterhan and Schwarz' (2016) three-node Argumentation for Learning (AFL) framework appears viable for us to conduct an in-depth study of the literature on

discourse. They explain that researchers should capture details about the purpose and intent of instructional decisions "task design" and the scaffolding that teachers develop "process support" when studying argumentative dialogue in order to identify various "inhibitors and enablers" of argumentation. This suggestion aligns well with other literature on the effects of cognitive supports such as planning tools which generally function as a bridge between oral and written argumentation (De La Paz et al., 2020; Matos, 2021; Shi et al., 2019). Individual student characteristics (such as having differences in cognitive knowledge and skills, epistemological beliefs, motivation, etc.) are also part of the AFL model.

Importantly, the AFL framework suggests how to study the characteristics of the resulting dialogue. Knowing who contributes to discussions and the quality of individual student contributions is important because merely engaging in discussion may not allow everyone in the group to participate due to a few students dominating conversation (De La Paz et al., 2021; Levin et al., 2021). Asterhan and Schwarz's framework also suggests how researchers may study the quality of students' scientific and historical discourse, and that doing so is also important in order to establish links between inhibitors and enablers and student learning.

Current Study

The connection between discourse and learning is complex and highly situated, which makes a direct connection imperfect (Asterhan & Schwarz, 2007; Murphy et al., 2009). The empirical research in this area continues to grow, and the field has identified aspects of discourse that can be identified and analyzed. We seek to use Asterhan and Schwarz's framework to study the argumentative discourse

literature in science and social studies. Accordingly, we explore the antecedents of argumentation and how they facilitate or inhibit argumentative discourse, the characteristics of the actual dialogue, and the learning outcomes of argumentation.

We seek to conduct a more fine-grained analysis of the process support (i.e., instructional components) that teachers use, as discourse instructions and teacher scaffolding shape how discourse in the classroom takes place. For example, what kinds of graphic organizers or scaffolds are used prior to, during, and after holding a discussion? Are teachers modeling how discourse should be held, and is there explicit instruction around discourse?

Equally important is how teachers structure discourse, be it in pairs, in small groups, as whole class discourse, or as a combination thereof. Pairs and small groups tend to promote higher participation rates among individual members (Kuhn & Crowell, 2011). In contrast, whole-class discussions can draw from a broader pool of ideas, including perspectives that may not arise in smaller settings (Herrenkohl, 2006). Consequently, it is essential for teachers to comprehend and evaluate the advantages and limitations of various instructional groupings.

In both science and social studies classrooms, teachers have successfully used inquiry to support a classroom community that facilitates students' abilities to develop knowledge and think collaboratively (Bae et al., 2021; Monte-Sano et al., 2021). Moreover, this is a natural consequence of learning in a social environment when students co-construct and critique knowledge (Mercier & Sperber, 2011). The current review considers disciplinary knowledge as crucial learning outcomes because it is difficult to master complex concepts if simpler concepts are not mastered

(Ford, 2008; Shanahan, 2014). Reading comprehension is also important but less commonly included as a dependent measure in the current discourse literature and was therefore not a variable in our work. Disciplinary writing is also vital for students because writing has been identified as a learning tool, and studies show that students can write to learn disciplinary knowledge (Asterhan & Schwarz, 2016; Bangert-Drowns et al., 2004; Newell, 2006). Finally, the current review adds to the literature by specifying how instructional groupings and type of discourse impact learning outcomes for diverse learners.

Research Questions

We used the AFL framework proposed by Asterhan and Schwarz (2016) to sharpen our focus on the role of specific antecedents (i.e., elements of task design and process support) that have been operationalized in the literature. Teachers' decisions related to group formation are considered part of task design, and process supports broadly include teacher scaffolding and discourse instructions. We also seek to study the role of deliberative or disputative argumentation discourse, or the role of dialogue characteristics more broadly, in the literature on scientific and historical discourse. The following research questions guided our review:

1. Does type of group formation (i.e., pairs, small group, large group) affect student learning outcomes in science and social studies K-12 classrooms?
2. What process supports have teachers utilized during discourse in science and social studies K-12 classrooms, and which show better results?
3. How do dialogue characteristics (i.e., disputative or deliberative) impact student learning outcomes in science and social studies K-12 classrooms?

Method

Search Procedure

We consulted systematic literature methodology papers and chapters (Alexander, 2020; Cooper and Hedges, 2009) and the updated Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (Page et al., 2021) to conduct a systematic search of the literature. We conducted the search using the following databases: Education Research Information Center (ERIC), Academic Search Ultimate, APA PsycArticles, APA PsycINFO, Education Source, EBSCOhost, Education Source, and E-Journals in the past 24 years (2000 to 2024). We identified terms related to discourse, science, and social studies, refining and expanding search terms as terminology differed in science and social studies. We utilized the following search terms: *discourse, discussion, dialogue, talk,* and *argument** (for argumentative or argumentation). We identified the following terms in science (*science education, science teaching, science learning*) and in social studies (*social studies*, *social studies education)* in primary and secondary U.S. classrooms.

We used Boolean search procedures to include discourse, argumentation, primary and secondary education, social studies, and science in our initial search. In total, the online database search produced 21,206 texts, including 13,029 studies, 6,783 reports, 2,024 magazines, 1,612 dissertations, 1,135 electronic resources, 871 reviews, 607 books, 295 conference materials, 137 news articles, 31 government documents, 24 trade publications, and four biographies. We further narrowed our search by requiring studies to be peer reviewed, quantitative or mixed methods, and between the year 2000 to May 2024. This resulted in 550 articles.

Inclusion and Exclusion Criteria

We screened titles and abstracts according to specific criteria. First, interventions had to include one discourse task. Second, we focused on science or social studies. Third, studies included students in primary and secondary grades (K-12). Fourth, foreign studies were acceptable if published in English. Fifth, we excluded studies that did not define a learning outcome. Of the 550 texts, 520 were excluded. The resulting 30 full texts were reviewed and included in the final sample. Reference lists were then reviewed but did not lead to additional articles. The total sample was 30 articles as shown in the Prisma diagram in Figure 2.1.

Coding Procedures

An early step in the research process involved organizing specific aspects about each study. We developed a coding matrix to categorize general study information, participant information, study setting, study design, independent variables, dependent variables, procedures, and intervention duration (dosage). We recorded whether the study involved disputative discourse or deliberative discourse (Asterhan & Schwarz, 2016). These elements were not always explicitly stated in the text, so we used the literature to guide operational definitions for the terms. Any type of discourse that had a goal of convincing or persuading or was competitive in nature was coded as disputative. Discourse that had a goal of arriving at conclusions through collaboration was coded as deliberation.

Studies were also coded for instructional grouping to learn whether these variables produced any effects. Researchers have studied ways to engage students in discourse through various instructional groupings, such as in pairs/dyads, in small

groups, whole-class, or a combination of these (Kuhn & Crowell, 2011; Reisman et al., 2019). Several process supports were identified and coded: explicit instruction (EI), teacher modeling (MODEL), use of graphic organizers (GO), and technology (TECH). We should note that EI and MODEL refer to their use without specifying how those supports were utilized. Learning outcomes, when students learned content specific to the discipline, were labeled disciplinary knowledge. Disciplinary writing was scored as a learning outcome if researchers scored historical or scientific thinking. Argumentation structure was scored as a learning outcome if researchers scored elements such as claim, evidence, explanation, counterclaim and/or rebuttal. Finally, holistic writing was scored as a learning outcome when researchers measured students' general writing skills (e.g., persuasiveness, spelling, grammar, etc.).

The first author coded all 30 studies, and one secondary coder who was blind to the research questions coded 11 studies (37%). The coder was trained by the first author on four randomly selected studies from the available pool and reached 89% agreement during training. Interrater reliability was calculated at 91%. Any differences were resolved through discussion.

Calculation of Effect Sizes

For studies that did not provide effect sizes (ES), a coder who was unfamiliar with the study independently calculated Hedges' g to investigate the efficacy of the interventions on learning outcomes. Hedges' g was calculated using the following formula: $Hedges'g = \frac{M_1 - M_2}{SD_{pooled}}$ where M1 is the treatment group posttest mean, M2 is the control or comparison group posttest mean, and SDpooled is the pool standard

deviation of both treatment and control group. To eliminate arithmetic errors, we used an online calculator (https://www.statology.org/hedges-g-calculator/). Four studies did not report ES or provide information needed to obtain them.

Results

The systematic review yielded 30 studies, including randomized (N=7), quasi-experimental (N=17), and mixed methods (N=6). Table 2.1 provides an overview. Table 2.2 provides information about the discipline, setting, instructional components, and student population (i.e., students with disabilities [SWD], multilingual learners [ML] and general education students [GEN]). Table 2.3 includes ES, with interpretation of ES (e.g., small, medium, large).

The student sample included primary (N=4), secondary (N=25), or both schooling levels (N=1). Studies were primarily conducted in the US (N=17), followed by Europe (N=10), Latin America (N=2), and the US and Europe (N=1). Only 13 studies identified SWD and eight identified ML students, (or non-native speaking students if in a foreign country). More than half of the studies (N = 17) did not report whether participants were SWD or GEN and almost three-quarters (N = 22) did not identify participants as ML or non-native speakers.

The types of discourse used in the studies were disputative (N=2), deliberation (N=23), or both disputative and deliberation (N=5). Discourse was held in pairs (N=6), small groups (N=6), whole class (N=3), a combination of all three (N=1), a combination of small groups and whole class (N=13), and one study did not specify. Half of our studies were conducted in social studies classrooms (N=15), with the remaining occurring in science classrooms (N=14), and both subjects (N=1). In all,

discourse was used by researchers and teachers in a variety of ways, with studies from three continents.

RQ1. Does type of group formation (i.e., pairs, small groups, large group) affect student learning outcomes in science and social studies K-12 classrooms?

We first report findings about the effects of group formation, as teachers' decisions regarding how to cluster students when they discuss differences in opinions, ideas and viewpoints may facilitate or negatively impact the resulting discourse. We then link these findings with four learning outcomes (disciplinary knowledge, disciplinary writing, argumentation structure, and holistic writing).

Pairs were used in six studies (de Vries et al., 2002; Felton et al., 2009, 2015; Garcia-Mila et al., 2013; Munneke et al., 2007; van Drie et al., 2005) with ES ranging from d = 0.74 to $\eta^2 = 0.41$, six studies used small group discussions (De La Paz et al., 2017; Casado-Ledesma et al., 2023; MacArthur et al., 2002; Murphy et al., 2018; Zhang et al., 2023), with ES ranging from g = 0.49 to $\eta^2 = 0.60$, three studies used whole class discussions (Gronostay, 2019; Nussbaum & Edwards, 2011; Reisman, 2012), with ES ranging from $d = 1.56$ to $\eta_p^2 = 0.03$.

In addition, 13 studies used both small group and whole class (Akbaş et al., 2019; Chen et al., 2016; De La Paz et al., 2021, 2022; Del Favero et al., 2007; Ferretti et al., 2001; Grooms et al., 2018; Kent et al., 2015; Larraín et al., 2018, 2021; Relyea et al., 2022; Wanzek et al., 2015; Wissinger & De La Paz, 2016), with ES ranging from g = 0.07 to g = 2.81, and one study used pairs, small groups, and whole class discussions (Iordanou et al., 2019) with ES ranging from g = 0.33 to g = 0.50.

These findings indicate that larger effect sizes in disciplinary knowledge and writing are associated with teachers using more than one instructional grouping.

Seventeen studies reported disciplinary knowledge gains, and with two exceptions, students who participated in some type of discourse group had better learning outcomes compared to students in a control group. Students in control groups differed in the process supports (e.g., Akbaş et al., 2019; De La Paz & Wissinger, 2017) and dialogue characteristics (Felton et al., 2009). For example, De La Paz and Wissinger (2016) found that students who engaged in disciplinary discussions (intervention group) as opposed to traditional discussions (control group) showed improved historical knowledge and historical thinking, including in far transfer (η_p^2 = 0.19). Furthermore, Larraín et al. (2018) found improved post-delayed disciplinary content knowledge (η^2 = 0.31) via lesson plans designed to foster dialogic and argumentative interactions (intervention group) as opposed to teacher-centered pedagogy (control group). Del Favero et al. (2007) however reported negative results and Wanzek et al. (2015), did not report significant differences in content knowledge after the intervention (ES of g = 0.07).

Of the 15 studies with positive outcomes between discourse and disciplinary knowledge, seven were in social studies (Akbaş et al., 2019; De La Paz & Wissinger, 2017; Ferretti et al., 2001; Kent et al., 2015; MacArthur et al., 2002; Reisman, 2012; Wissinger & De La Paz, 2016), eight in science (Casado-Ledesma et al., 2023; Felton et al., 2009, 2015; Grooms et al., 2018; Larraín et al., 2018, 2021; Murphy et al., 2018; Relyea et al., 2022), and one in both science and social studies (Iordanou et al., 2019). Six social studies interventions had moderate to large ES, ranging from η_p^2 =

0.03 to η^2 = 0.30. Seven science studies had positive moderate to large ES, ranging from *d* = 0.74 to η_p^2 = 0.44. Relyea et al. (2022) was an exception, with an ES of g = 0.43.

Murphy et al.'s (2018) science Quality Talk study was notable as findings indicated decreased teacher talk (and increased student talk) that led to improved conceptual understanding. Akbaş et al. (2019) found that using argumentation-based science learning (ABSL) allowed students to be more active in their learning as they comfortably expressed themselves, learned through inquiry, and held fun discussions compared to a more traditional teaching approach. These findings suggest that when students discuss content, they learn from each other. Furthermore, these results show no difference in findings across disciplines.

Improvements in disciplinary writing were found in 13 studies. Seven were in science (Casado-Ledesma et al., 2023; Chen et al., 2016; De La Paz et al., 2022; Grooms et al., 2018; Murphy et al., 2018; Relyea et al., 2022; Zhang et al., 2023) and six in social studies (De La Paz & Wissinger, 2017; De La Paz et al., 2014, 2017, 2021; Wanzek et al., 2015; Wissinger & De La Paz, 2016). Two studies, by Casado-Ledesma et al. (2023) and Chen et al. (2016) did not have the necessary information to calculate ES, so conclusions about these science studies are limited.

Of the remaining 11 studies, De La Paz et al. (2022) had positive moderate to large ES on near transfer (*d* = 1.08) and far transfer (*d* = 0.76), including for SWD (*d* = 0.73). Grooms et al. (2015), Murphy et al. (2018) and Relyea et al. (2022) had small to moderate ES, ranging from *d* = 0.40 to η^2 = 0.06. Students' writing in social studies also improved significantly in conditions involving discourse. One study by

De La Paz et al. (2021) had a smaller ES (d = 0.33) than related studies. The intervention was a single week, however, in contrast to other studies with longer interventions, which had moderate to large ES, ranging from d = 0.56 to η_p^2 = 0.45. Wanzek et al. (2015) also had small ES (g = 0.31 on the written content and g = 0.16 on support), though possibly due to low teacher fidelity.

The writing ranged from relatively simple to more complex products. For example, in Grooms et al.'s (2015) study, students engaged in an argumentation session after analyzing data, followed by writing scientific investigation reports that were revised after undergoing a double-blind group peer review (e.g., Grooms et al., 2015). Students in De La Paz et al. (2017) wrote five-paragraph argumentative essays about a historical controversy after reading, annotating, and discussing two documents. In sum, discourse can improve disciplinary writing in social studies and science classrooms regardless of the complexity of the writing task.

Thirteen studies demonstrated improvements in oral or written argumentation structure. Eight took place in science classrooms (Chen et al., 2016; De La Paz et al., 2022; de Vries et al., 2002; Felton et al., 2009, 2015; Garcia-Mila et al., 2013; Nussbaum & Edwards, 2011; Zhang et al., 2023), four in social studies classrooms (Ferretti et al., 2001; MacArthur et al., 2002; van Drie et al., 2005; Wissinger & De La Paz, 2016), and one took place in both science and social studies (Iordanou et al., 2019).

Four studies focused on oral deliberation (Felton et al., 2009, 2015; Garcia-Mila et al., 2013; MacArthur et al., 2002). These studies had small to large ES, ranging from r = 0.33 to η_p^2 = 0.17. For example, Garcia-Mila et al. (2013) found that

process support and dialogue characteristics affected the way students formed oral arguments. Students who engaged in deliberation crafted better claims, supported their evidence more effectively, and composed better rebuttals compared to students who engaged in disputative argumentation. In addition, students who engaged in discourse stated claims supported by relevant evidence and analyzed the evidence more effectively.

Felton et al. (2009, 2015) and Garcia-Mila et al. (2013) collectively show that students in deliberation conditions constructed higher quality arguments than those in the disputative condition. MacArthur et al. (2002) demonstrated that general education and special education students can construct oral arguments, though they did not explicitly warrant claims. Nevertheless, students with and without disabilities had high levels of participation in the debate.

Of the 13 studies that showed improved oral or written argumentation, 10 studies demonstrated a causal connection between discourse and improved written argumentation structure (Chen et al., 2016; De La Paz et al., 2022; de Vries et al., 2002; Iordanou et al., 2019; Munneke et al., 2007; Murphy et al., 2018; Nussbaum & Edwards, 2011; van Drie et al., 2005; Wissinger & De La Paz, 2016; Zhang et al., 2023). Chen et al. (2016) and de Vries et al. (2002) did not report nor provide necessary information to calculate ES. The science studies' ES were moderate to large, ranging from d = 0.73 to η^2 = 0.31. The social studies interventions had larger ES, ranging from g = 1.12 to η_p^2 = 0.41. Overall, discourse helped improve students' oral and written argument structure.

Three studies measured holistic writing (De La Paz et al., 2014, 2017; Wissinger & De La Paz, 2016). The control group in De La Paz et al (2014) had higher holistic quality (ES of $d = 0.10$) than the treatment group. After modifying their 2014 intervention to address lack of improvement in holistic quality, De La Paz et al. (2017) showed large ES in holistic quality ($\eta_p^2 = 0.32$). Wissinger and De La Paz (2016) did not show a difference in holistic writing because the study design focused on different types of process support rather than controlling for presence or absence of discussion. With such studies including holistic quality, more studies are needed to determine whether discourse functions as a mediator, when other instructional variables are not controlled.

RQ2. What process supports have teachers utilized during discourse in science and social studies K-12 classrooms, and which show better results?

Seventeen studies utilized EI as an instructional component (Casado-Ledesma et al., 2023; Chen et al., 2016; De La Paz & Wissinger, 2017; De La Paz et al., 2014, 2017, 2021, 2022; Ferretti et al., 2001; Grooms et al., 2018; Kent et al., 2015; Munneke et al., 2007; Murphy et al., 2018; Reisman, 2012; Relyea et al., 2022; Wanzek et al., 2015; Wissinger & De La Paz, 2016; Zhang et al., 2023). Teachers used explicit instruction to guide student learning in both social studies and science, and vocabulary instruction for the disciplines was common. Studies using EI provided small to large ES, ranging from $\eta_p^2 = 0.03$ to $\eta_p^2 = 0.45$, though two studies did not have the necessary data to calculate ES (Casado-Ledesma et al., 2023; Chen et al., 2016). These results suggest that EI improves learning, which is not surprising as this

is an evidence-based approach for teaching vocabulary, disciplinary knowledge, disciplinary writing, and argumentation.

Fifteen studies used MODEL in their instruction (Casado-Ledesma et al., 2023; De La Paz & Wissinger, 2017; De La Paz et al., 2014, 2017, 2021, 2022; Kent et al., 2015; MacArthur et al., 2002; Murphy et al., 2018; Nussbaum & Edwards, 2011; Reisman, 2012; Relyea et al., 2022; Wanzek et al., 2015; Wissinger & De La Paz, 2016; Zhang et al., 2023). Unique to this group of studies, Casado-Ledesma and colleagues used video modeling, which significantly predicted learning gains in argumentation. However, it was not possible to determine an ES from the manuscript. All studies using MODEL had small to large ES, ranging from $\eta_p^2 = 0.03$ to $\eta_p^2 = 0.45$. Of the social studies interventions that used MODEL, 11 studies had moderate to large ES. Only Reisman (2012) and Wanzek et al. (2015) had small ES ($\eta_p^2 = 0.03$ and g = 0.31, respectively), and that was in factual and content knowledge. However, Reisman's findings are not surprising given that teacher fidelity in her study was low. Wanzek et al. (2015) indicated similar challenges, teachers may have needed more time to implement the Team-Based Learning (TBL) due to its complicated nature. With these exceptions noted, it appears that studies using MODEL had moderate to large ES with two exceptions, suggesting that MODEL is an effective strategy.

Twenty-three studies utilized GO in their interventions (Akbaş et al., 2019; Casado-Ledesma et al., 2023; Chen et al., 2016; De La Paz & Wissinger, 2017; De La Paz et al., 2014, 2017, 2021, 2022; de Vries et al., 2002; Ferretti et al., 2001; Grooms et al., 2018; Iordanou et al., 2019; Kent et al., 2015; MacArthur et al., 2002; Munneke et al., 2007; Murphy et al., 2018; Nussbaum & Edwards, 2011; Reisman, 2012;

Relyea et al., 2022; van Drie et al., 2005; Wanzek et al., 2015; Wissinger & De La Paz, 2016; Zhang et al., 2023). These studies reported small to large ES, ranging from $\eta_p^2 = 0.03$ to $\eta_p^2 = 0.44$. Nineteen had moderate to large ES. For example, Nussbaum and Edwards (2011) found that including critical questions in argumentation vee diagrams (AVD) improved argumentation (e.g., more integrated arguments). In sum, helping students organize their arguments in a visual manner can improve learning outcomes.

The small ES in Reisman (2012) and De La Paz et al. (2021) are not surprising given teachers' low fidelity in the former study and the brevity of the intervention in the latter. The low ES sizes in Wanzek et al. (2015) are also unsurprising given teachers needed more time to effectively implement the intervention. Finally, Casado-Ledesma et al. (2023) showed no benefit to adding a GO to an experimental group that integrated EI and deliberative discourse.

These results collectively highlight the importance of teachers providing students with effective process supports such as graphic organizers. In this literature, teachers used GO to facilitate the organization and integration of new knowledge. GO provide a structured framework that allows students to focus on understanding and engage in higher-order thinking skills.

Seven studies incorporated TECH in their intervention. Students in four studies engaged in discourse via a computer chat (de Vries et al., 2002; Iordanou et al., 2019; Munneke et al., 2007; van Drie et al., 2005). One study included video modeling (Casado-Ledesma et al., 2023), one study used a computer-enhanced curriculum (Larraín et al., 2021), and one study used deliberation using technology

(Ferretti et al., 2001). Casado-Ledesma and colleagues recruited students of the same age as participants to video model the discussion task. They provided these students with a script of the EI. Larraín et al. (2021) used an app to support argumentation pedagogy through the integration of purposely designed whole-class and group-work interactions. The app allowed teachers to control the flow and interactions of preloaded lessons on one tablet, and student groups would receive a separate tablet with the student view of the app through which they received task instructions, visual resources, etc.

The studies using chats showed varied ES, ranging from g = 0.33 to $\eta_p^2 = 0.41$, except in de Vries et al.'s (2002) study, where ES could not be determined due to inadequate data. Only Iordanou et al. (2019) had a small ES (g = 0.33). It appears likely that using virtual chats also aided students to learn when engaged in discourse, though more research is needed.

It is worth mentioning that Casado-Ledesma et al. (2023) evaluated each instructional element. They used four groups to evaluate EI with MODEL via video modeling (TECH), GO, and deliberative discourse. They found that students in all conditions improved in their socio-scientific knowledge, but EI combined with deliberative discourse demonstrated the greatest gains. Nine studies with three instructional components took place in social studies classrooms (De La Paz & Wissinger, 2017; De La Paz et al., 2014, 2017, 2021; Ferretti et al., 2001; Munneke et al., 2007; Reisman, 2012; Wissinger & De La Paz, 2016) and one was conducted in science classrooms (De La Paz et al., 2022). Small to large ES were reported, ranging

from d = 0.33 to g = 2.81. Most studies had large ES, indicating that three instructional components may be sufficient to promote learning.

RQ3. How do dialogue characteristics (i.e., disputative or deliberation) affect student learning outcomes in science and social studies K-12 classrooms?

Of the 17 studies that reported disciplinary learning, all but two (Del Favero et al., 2007; Wanzek et al. 2015) reported increased learning through discourse. Of the 15 studies with positive findings, 13 used deliberation (Casado-Ledesma et al., 2023; De La Paz & Wissinger, 2017; Ferretti et al., 2001; Grooms et al., 2018; Kent et al., 2015; Larraín et al., 2018, 2021; Munneke et al., 2007; Murphy et al., 2018; Reisman, 2012; Relyea et al., 2022; Wanzek et al., 2015; Wissinger & De La Paz, 2016) with small to large ES, ranging from η_p^2 = 0.03 to η^2 = 0.44. Two studies used both disputative and deliberation (Felton et al., 2009; Iordanou et al., 2019) with small to moderate ES, ranging from g = 0.33 to η^2 = 0.11. Felton et al. (2009) compared both and found deliberation to be more effective.

One study used disputative discourse (MacArthur et al., 2002) with moderate ES (g = 0.49). MacArthur et al. (2002) reported that student engagement was high during the debate, and that students demonstrated gains in knowledge. However, the eight-week project-based investigation contained more than just the debate, so it is difficult to conclude that the debate led to the increased knowledge rather than the other parts of the curriculum (e.g., small-scale investigation, “migration and conflict” schema). Moreover, students used discourse that resembled everyday argumentation rather than academic argumentation. The studies using deliberation reported large ES,. This evidence suggests that when students engage in consensus building to arrive

at an integrative position, they are more likely to comprehend and retain information. Findings from studies using critical questions (e.g., De La Paz & Wissinger, 2017; Wissinger & De La Paz, 2016) suggest that students who participated in deliberative argumentation engaged in critical thinking.

Use of deliberation also led to better (i.e., more sophisticated reasoning or higher quality) written texts (Casado-Ledesma et al., 2023; Chen et al., 2016; De La Paz & Wissinger, 2017; De La Paz et al., 2014, 2017, 2021, 2022; de Vries et al., 2002; Grooms et al., 2018; Murphy et al., 2018; Nussbaum & Edwards, 2011; Relyea et al., 2022; van Drie et al., 2005; Wanzek et al., 2015; Wissinger & De La Paz, 2016) with small to large ES, ranging from d = 0.33 to η_p^2 = 0.45. One study that had both deliberation and disputative (Iordanou et al, 2019) reported small to moderate ES, ranging from g = 0.33 to g= 0.50. Larger ES were present when students deliberated, suggesting that deliberation led to higher quality written products. Moreover, students applied disciplinary thinking at higher levels in their writing.

Discussion

This systematic literature review was designed to examine the impact of discourse on students' learning in science and social studies classrooms. Following Asterhan and Schwarz's (2016) AFL framework, we wanted to know how argumentation inhibitors and enablers (e.g., process support, task design) plus dialogue characteristics (e.g., deliberative, disputative) affected learning outcomes. More specifically, we wanted to know whether and how instructional groupings, instructional components, and/or discourse type had an impact on students' disciplinary knowledge and writing in science and social studies classrooms. The

systematic review yielded 30 studies (seven randomized, 17 quasi-experimental, and six mixed methods). These studies came from Europe, South America, and North America. We also provided the ES of studies based on the independent variable and the learning outcome when possible.

How tasks are designed, including group formation, play an important role as an inhibitor or enabler of argumentation (Asterhan & Schwarz, 2016). In answering our first research question regarding instructional groupings, we found that students benefited from discourse when they were in more than one instructional grouping. Integrating instructional groupings of various sizes and purposes enables students to engage in extensive and meaningful discussions with multiple peers, along with receiving teacher feedback. Consistent with prior research, this approach leads to improvements in disciplinary knowledge (Asterhan & Schwarz, 2007, 2016; Colley & Windschitl, 2016; De La Paz & Wissinger, 2017; Murphy et al., 2017), highlighting why discourse is considered a high-leverage practice (Brownell et al., 2019; Capobianco et al., 2018; O'Connor & Snow, 2018). To achieve this, it is essential for teachers to establish norms and routines that foster a collaborative classroom environment (Monte-Sano et al., 2021).

Different instructional groupings are often used to build on each other (e.g., De La Paz et al., 2022; Hemberger et al., 2017) and for different purposes. During whole class discussions, providing students with a set of critical questions will facilitate analysis and critique of ideas (Felton et al., 2022; Nussbaum et al., 2023). In addition, assigning students to work in pairs allows students more opportunity to engage in discourse, especially students who otherwise would not participate in either

small group or whole class discussion, such as ML and SWD (Boyd & Rubin, 2002; Kelly, 2007; Walqui & Heritage, 2018).

Educators can expand upon pair discussions by organizing students into small groups (e.g., Iordanou et al., 2019). These small groups provide students with the opportunity to engage with multiple peers, who may present distinct or similar viewpoints. Effectively structured small group interactions can foster meaningful deliberative discourse, regardless of initial pair placements. Students can test the ideas constructed in pairs with a larger audience, thus gaining more confidence and fostering argumentation skills (Iordanou et al., 2019). Furthermore, small group discussions serve as a platform for further knowledge construction, argumentation, or preparation for subsequent whole-class discussions (e.g., Akbaş et al., 2019; Ferretti et al., 2001).

Whole class discourse may then be used for broader evaluation of ideas after students discuss them in small groups (e.g., Felton et al., 2022; Grooms et al., 2018). Whole class discussion allows students to engage with all their peers in the classroom and their teachers to address varying points of view (Chen et al., 2016; Gronostay, 2019; Reisman, 2012). Teachers can play a substantive role in facilitating and supporting meaningful student engagement in disciplinary interactions (Felton et al., 2022). Such discussions can be advantageous as students evaluate the strengths and weaknesses of various arguments, ultimately identifying the most compelling one with appropriate guidance. Through discourse, students transform into active participants, engaging with peers and contributing to their own learning process,

rather than remaining passive recipients of teacher-led instruction (Erduran et al., 2004; O'Connor & Snow, 2018).

Teachers can move from one instructional grouping to another based on the quality or engagement of discourse (e.g., Iordanou et al., 2019; Wissinger & De La Paz, 2016). Unfortunately, curricula that focus on breadth of content rather than reasoning will impede use of more than one instructional grouping. Our findings suggest that teachers make time to structure more than one instructional grouping. Doing so will allow them to effectively integrate learning from peers that may even promote learning transfer (Felton et al., 2022). Teachers and curriculum designers should work together to integrate more discourse. As Monte-Sano et al. (2021) noted, "…we see the curriculum and teachers as partners in cultivating discourse" (p. 176).

In answering our second research question, which focused on the effectiveness of four instructional components, we analyzed how teachers used EI, MODEL, GO, and TECH when engaging in discourse. Students benefitted from receiving EI and MODEL, as they can see and hear what discourse focuses on when teachers verbalize their thinking (e.g., Berland & Lee, 2012; Casado-Ledesma et al., 2023). In line with previous findings from Murphy et al. (2009) and Bennett et al. (2010), students need to visualize how to engage in discourse prior to engaging in discourse. Most studies did not clarify whether teachers actually modeled how students should engage in discourse. This is unfortunate because students need opportunities to observe argumentative discourse (Asterhan & Schwartz, 2016); in fact, EI and MODEL are likely beneficial when used to help students practice discussions as they are both evidence-based practices (EBP) and HLP in themselves.

Perhaps seeing similar-aged students enact a task they will soon participate in encourages students to engage. Likewise, video modeling, especially when teachers embed "chunking," is likely to be effective as doing so has the advantage of replaying parts of EI that students may have missed. Moreover, students can see the social negotiation in using evidence-based arguments to exchange, evaluate, challenge, and defend ideas within their classroom community to arrive at a consensus (Berland, 2011). For example, Casado-Ledesma et al. (2023) used students to simulate an expert discussion to reach collaborative solutions. They demonstrated through the video the processes involved in selecting arguments from sources, comparing arguments from differing positions, formulating an integrative conclusion, and writing an argumentative synthesis that incorporates this conclusion and its justification. Although future research is needed to determine effects, video modeling should be considered when providing EI and MODEL in future studies.

Additionally, providing students with GO allows them to organize their thoughts as they prepare for discourse (e.g., Nussbaum & Edwards, 2011) or as they shift from discourse to writing (e.g., De La Paz et al., 2017; Grooms et al., 2018). Moreover, findings suggest that graphic organizers promote active learning and critical thinking, as students engage in organizing and mapping out information, which fosters deeper cognitive engagement and facilitates the development of higher-order thinking skills. Diagrams help students categorize information for arguments (e.g., arguments, counterarguments), such as Nussbaum and Edwards' (2011) argumentation vee diagram (AVD). These and other organizational tools are

particularly beneficial for diverse learners, including those with learning disabilities, as they can be tailored to meet individual learning needs.

In Asterhan and Schwarz's (2016) AFL framework, computer-mediated discourse and software design are considered argumentation inhibitors and enablers. Certainly, TECH may be helpful in facilitating discourse through digital GO, online discussion forums, and videos -- however future studies are needed to determine how teachers might incorporate technology in classroom discussions. In the current age of mis/disinformation, it is crucial to teach students how to strategically evaluate the credibility of online sources (Wineburg & McGrew, 2019) including those seen or viewed through applications like Instagram or TikTok. It is also important that teachers leverage TECH especially as many schools began to offer students electronic devices with internet connection during the Covid-19 pandemic. In academic settings, discourse through digital means affords students who may not otherwise participate in oral discourse, such as ML and SWD, the opportunity to participate. It also provides students more time to think about what they will say and the opportunity to revise what they will communicate (de Vries et al., 2002).

In answering our third research question, on the effectiveness of discourse intent, our findings corroborate Felton et al. (2009) and others who reported that deliberation is more effective than disputative discourse at helping students learn disciplinary knowledge (Zhang et al., 2023). This is true in both science (Casado-Ledesma et al., 2023; Larraín et al., 2018, 2021) and in social studies (De La Paz et al., 2017; Wissinger & De La Paz, 2016), presumably because in both content domains, as students co-construct knowledge, they can more easily internalize it.

Another benefit of deliberation is that it can facilitate written argumentation by providing a "missing interlocutor" for students (Graff, 2003). Thus, when students are asked to collaborate to arrive at a consensus, they can resolve differences for the good of their argument. Moreover, even if students do not arrive at a consensus, deliberation provides them an opportunity to improve their argument without the negativity that can arise in disputative argumentation. Furthermore, argumentation may be hindered by perceptions of interpersonal competition that can diminish cognitive flexibility and receptiveness to different viewpoints, which can further lead students to concede prematurely (Asterhan & Schwarz, 2016).

This review has shown that SWD, particularly learning disabilities, and ML can use deliberative discourse to improve their learning, both in terms of content knowledge (MacArthur et al., 2002; Reisman, 2012) and in their writing (c.f., De La Paz et al., 2017). Unfortunately, 57% of the studies did not report on the disability status of participants and 73% did not report on whether students were native speakers. Struggling learners, especially ML, are less likely to participate in classroom discourse (Boyd & Rubin, 2002; Walqui & Heritage, 2018). Fortunately, when teachers provide appropriate scaffolds for SWD and ML learners, struggling learners have been shown to make comparable learning gains (De La Paz et al., 2023; Grooms et al., 2018; Relyea et al., 2022).

At a minimum, researchers should report findings regarding vulnerable populations. In addition, because EI, GO, and MODEL are all EBPs for teaching students with disabilities, studies that report findings for these student populations

could help researchers understand what, for whom, and under what conditions interventions work (Langley et al., 2009).

We found that discourse was effectively used to promote student outcomes across each of the countries in this review. The recent emphasis in American schools to incorporate discourse (NRC, 2013), and the minimal implementation of discourse in classrooms (O'Connor & Snow, 2018), suggests that there is much room for improvement in the United States. It would be interesting to explore whether other countries have similar struggles in integrating discourse in their classrooms. Likewise, with the increase in global migration, it would be interesting to compare results with non-native speakers in other countries to ML in the United States.

NGSS and the C3 framework call for students to hold discussions in their respective disciplines. Consistent with prior research (Asterhan & Schwarz, 2007, 2016; Nystrand et al., 1997), this review showed that students who engaged in discourse showed improved disciplinary knowledge in both science (e.g., Felton et al., 2009; Larraín et al., 2018, 2021) and social studies (e.g., MacArthur et al., 2002; Reisman, 2012). When students discuss their ideas, they can construct knowledge and learn disciplinary knowledge as they exchange and critique each other's ideas.

Finally, this study shows that when learning was measured through writing, students improved their argumentation structure and overall ability to communicate. This finding is consistent with prior research that shows that discourse can be a bridge to writing (Shi et al., 2019). The majority of studies in this review contained argumentation examples of students engaged in deliberative discourse stating a claim, supporting it with evidence, evaluating the evidence, and making counterclaims and

rebuttals. Students are then asked to either defend their claim based on the selected evidence or to build on their claim by collaborating with peers to settle disputes. We note, however, that holistic writing was only addressed in three studies. This shows the main focus of most studies was on disciplinary or argumentative writing. However, there is some evidence that students can improve their general writing through disciplinary and argumentative writing (e.g., De La Paz et al., 2017).

Limitations

This review has a few limitations. Findings from Casado-Ledesma et al. (2023), Chen et al. (2019), Gronostay (2019), and de Vries et al. (2002), should be viewed as tentative as it was not possible to determine ES from these investigations. Second, nine studies do not clarify whether the explicit instruction and/or modeling relate to how to hold a discussion or if the supports were used for different purposes. Finally, while our search methodology was as robust as possible, we may have unintentionally missed a study that could have been considered in our sample.

Implications for Practice

Curriculum writers should ensure lessons allocate appropriate time for discourse and focus more on direct instruction (including videos and metacognitive modeling) and teachers will need to consider the time allotted to each instructional grouping (Monte-Sano et al., 2021). Thus, when teachers combine discourse in pairs with small groups and then whole class discussions, students engage in substantive disciplinary discussions and construct a higher understanding of the disciplinary knowledge and reasoning. Moreover, deliberation allows students to reason with their peers collaboratively, whereas disputative argumentation can lead to a combative

environment in which struggling learners may not feel comfortable participating. All students benefit from EI, MODEL, and GO in discourse. Students also are likely to appreciate having the option to engage in electronic discourse, and using tools such as Google docs allow teachers to see in real time what students are thinking if they type their discourse in a digital GO. Lastly, teachers can also use video modeling for students to replay. Video models have been identified as EBP for SWD and have even been found to be productive for college student writers (Alexander et al., 2023).

Future Research

Future studies should explore how task design (i.e., instructional groupings) and process support (i.e., scaffolding) impact how students engage in discourse and disciplinary thinking. Furthermore, studies should expand on discourse instructions and teacher scaffolding given the promising results provided here and students' poor performance on national writing assessments. Research should also investigate the best ways to use technology in discourse (e.g., blogs, vlogs, etc.), especially as technology becomes more prevalent in schools. Moreover, the unprecedented rise of artificial intelligence (AI) poses significant challenges and benefits for educators. AI chatbots could help students, especially struggling learners, participate in disciplinary discussion. Productive suggestions on how to use AI for learning (e.g., from Edutopia, the International Society for Technology in Education, and/or CommonSense Media) provide suggestions for researchers to explore in discourse interventions.

Conclusion

This systematic review illustrates how discourse may be effectively used in science and social studies classrooms. Discourse may be supported in more than one way, and still be effective. It is especially effective when students engage in deliberation. Future researchers should validate new structures and processes that integrate discourse in interventions and curricula using to push instruction, for example when using video modeling and technology, and in new content areas, such as civics education. We believe that doing so will also help students to be equipped for post-secondary contexts when they consider complex or controversial social issues that also require a reasoned approach to discussion.

Chapter 3: Effects of Structured Controversy with and without Guided Support on Academically and Linguistically Diverse Students' Historical Thinking

Educators are increasingly encouraged to use discourse to help students make sense of information they grapple with in history and social studies classrooms (De La Paz et al., 2014, 2017; Monte-Sano, 2008; Nokes, 2017; Reisman, 2012, 2015; Reisman et al., 2018). Monte-Sano et al. (2021) note that in history classes, discourse has the potential to make thinking visible. Moreover, they posit that it is grounded in sources that explain meaning and creates knowledge that everyone can utilize. Both the teacher and students draw on points regarding textual evidence from multiple sources, laying the groundwork for analysis, evaluation, and synthesis. However, as emphasized by O'Connor and Snow (2018), facilitating discourse requires adequate student participation and motivation to guarantee equitable access to knowledge, ensuring the integrity and coherence of the knowledge being shaped, and managing classroom constraints, including time.

In framing academic controversies through collaborative learning, Johnson and Johnson (1979, 1985, 1988) state educators should define controversies as problems in need of solutions rather than win-lose situations. Moreover, in discussions about academic controversies, students have conflicting ideas and opinions that they must work through until they reach an agreement. These authors developed a routine that leads students to engage in deliberation, known as structured academic controversy (SAC), so that students discuss information and advocate for a position, analyze, evaluate, and rebut information, and arrive at factual and judgmental conclusions. In a SAC, students acquire knowledge as they learn to

employ language and information to interact with others regarding controversies over facts, definitions, and values (Parker, 2021). The academic content is demanding, thus requiring instructional supports to construct students' lacking prior knowledge and strategies to leverage their funds of knowledge. In addition, instruction also presents challenges as it requires adept management of small groups alongside facilitation of reading comprehension and effective interpersonal discourse.

van Drie & van de Ven (2017) describe how students used whole class discussion to discuss a historical question prior to writing individual texts. In their exploratory study, they show that students integrated arguments from documents and whole class discussion. They demonstrate how ideas can evolve "from interpersonal classroom interactions to intrapersonal individual writing" (p. 537). The primary role of the teacher is to facilitate the discussion by allocating turns, posing questions, requesting evidence for claims, and challenging the students' assertions. Reisman et al. (2019) found that teacher candidates initially engage students, with high participation rates, in responding to open-ended questions after the teacher candidates receive explicit instruction in social studies education coursework. In addition, discourse can help students make sense of documents and plan before writing (Monte-Sano et al., 2014; De La Paz et al., 2021). However, students continue to struggle when writing argumentative essays.

According to the latest National Assessment of Educational Progress (NAEP) writing results from 2011, only 27% of 12th and 8th grade students are at or above proficient levels of writing performance, and only 5% of students with disabilities (SWD) are at or above proficient in writing (NCES, 2022). It is noteworthy that

because the proficiency levels between both grade levels for all students and SWD are identical, the data shows an absence in growth from middle to high school for both groups of students.

Clearly, instruction needs to improve to help students make progress in their writing. There is extant literature on the impact that discourse, specifically disciplinary discourse, can have on student writing. Interventions that focus on or utilize argumentative discourse prove effective at improving argumentative writing in science (Chen et al., 2016) and in history (De La Paz et al., 2014, 2017, 2021; Reisman, 2012). Such interventions have also proved effective at improving argumentative writing for SWD in upper elementary and middle school (De La Paz & Wissinger, 2017; Wissinger & De La Paz, 2016; Wissinger et al., 2021).

According to the U.S. Department of Education, 64.8% of SWD received special education services in general education classrooms in 2019 (NCES, 2019). Students with high incidence disabilities account for 77% of students receiving special education services (Gargiulo & Metcalf, 2023). Moreover, research shows that SWD are less likely to engage in discussion (Bueso, 2022; Louick & Wang, 2021) and they struggle more to write compared to their neurotypical peers (Dockrell & Connelly, 2016; Saddler et al., 2017). Our goal in this study is to explore how to help all students to engage more meaningfully in discussion, as well as how to transfer their ideas from discourse to writing.

Disciplinary Discourse and Writing

Though students may be familiar with making arguments, they must differentiate everyday arguments from disciplinary argumentation to read and write in

secondary contexts (Moje, 2008). National initiatives such as the Common Core Standards and the C3 Framework for Social Studies have created standards to map ideas about the development of argumentative milestones, including to prepare students for civic life (Waring, 2024). The C3 Framework for Social Studies calls on students to develop disciplinary practices and thinking by engaging in inquiry (NCSS, 2013). In social studies classrooms, particularly history and civics classrooms, inquiry opens a forum where students can engage in meaningful discussion as they investigate debatable and complex questions. And, in history, students consider historical controversies by analyzing diverse sources with conflicting evidence and ultimately develop conclusions (De La Paz et al., 2014, 2017, 2021; Grant et al., 2017; Monte-Sano et al., 2014; Nokes, 2017; Reisman, 2012).

To adequately participate in such controversies, students must acquire the skills of historical thinking and the dispositions that historians possess. This means understanding concepts such as contextualization, corroboration, and sourcing (Wineburg, 1991), historical significance and cause and effect (Seixas, 2006) and discipline-specific skills to think, read, and write about history (Nokes, 2017). Historians seek to construct new knowledge by addressing historical questions that have not been attended to previously or have been misinterpreted in their view (Nokes & De La Paz, 2018). Understanding why people do these things is likewise an important disposition.

In contrast, students typically view history as merely fixed events that happened in the past (Lee, 2005; MacArthur et al., 2002; Nokes & De La Paz, 2018; Wineburg, 2001). For example, students often have difficulty or are not able to place

events in appropriate historical context but rather view past events through present-day standards, values, or worldviews (Reisman, 2015). Students' aptitudes in engaging in historical thinking can then be defined by their ability to negotiate productive solutions through specific skills. Though these skills are challenging, students can achieve them with proper instruction and supports in reading, thinking, and writing about history (De La Paz et al., 2014, 2017, 2021; Monte-Sano, 2008, 2014, 2017; Wissinger & De La Paz, 2016).

Deliberative argumentation can allow students to develop written argumentation by permitting students to communicate their ideas to the "missing interlocutor" (Graff, 2003). Chapter 2 presents a synthesis of disciplinary discourse in science and social studies to improve learning outcomes, including disciplinary knowledge and writing. We found that deliberation was the most common and most powerful (i.e., medium to large ES) type of discourse (Casado-Ledesma et al., 2021; Chen et al., 2016; De La Paz & Wissinger, 2017; De La Paz et al., 2017; De La Paz et al., 2021; de Vries et al., 2002; Nussbaum & Edwards, 2011; van Drie et al., 2005; Wissinger & De La Paz, 2016).

Theoretical Framework

While originally designed as a means for systematically reviewing empirical studies, Asterhan and Schwarz (2016) developed an argumentation for learning (AFL) framework that considers inhibitors and enablers of argumentation such as task design and process support that help operationalize how argumentation is conceptualized and studied. The current study uses a well-developed task design (i.e., historical controversies, as described below) with all student participants, but it adds

to the literature by also comparing two types of process support for classroom discourse and student learning outcomes.

Asterhan & Schwarz (2016) recognize the importance of individual learners' characteristics as additional enablers or inhibitors of argumentation. Characteristics such as students' cognitive levels, motivation, epistemological beliefs, and gender have all been shown to matter in studies in which these are factored into the data analysis. We anticipate that these characteristics and others (such as performance on high stakes social studies tests) will also make a difference in the current study. However, our analysis did not encompass students' motivation or epistemological beliefs, as these variables were neither measured nor available in our dataset. Additionally, while we report students' gender, it was not a focal point of our investigation. This approach aligns with Asterhan and Schwarz's (2016) recommendation to include gender data for the purpose of facilitating cross-study comparisons, even when gender is not a primary research objective.

With respect to the characteristics of dialogue, Asterhan and Schwarz, (2016) suggest that productive dialogue includes deliberative as opposed to disputative argumentation or consensual co-construction, in which speakers expand or explain others' ideas. Therefore, the current study explores diverse students' actual dialogue characteristics, to determine whether the different process supports influence different types of student learners.

Finally, several learning outcomes are suggested in the accompanying framework. For the purpose of our study, we focus on students' argumentation and historical thinking skills during discussion. Our interest lies in how students construct

arguments and the quality of their historical judgments during discussion. We plan to look at other learning outcomes by analyzing students' written arguments in a subsequent study to seek evidence of additional types of learning. Finally, as noted by Asterhan & Schwarz (2016), the limited evidence of a causal link between argumentation and domain specific knowledge gains remains problematic. In analyzing data collected for this study, we agree that inhibitors and enablers, dialogue characteristics, and learning outcomes should be researched together.

With respect to task design, each Historical Investigation (HI) provides a coordinated way to introduce historical thinking, reading, and writing skills. We employ a cognitive apprenticeship (CA) instructional approach, which is useful when teaching complex tasks (Collins et al., 1991). Collins et al. (1991) promote CA as characterized by distinct phases aimed at promoting student mastery through a gradual release of responsibility, tailored to students' developing understanding. Clearly, this instructional approach should help researchers to create developmentally appropriate historical reasoning tasks for adolescents, and the fact that our approach has led to positive learning outcomes for a wide range of learners in the past demonstrates that our task design functions as argumentation enablers.

Current Study

The purpose of this study is to explore how argumentation inhibitors and enablers moderate dialogue characteristics and learning outcomes, and to describe the quality of the resulting discourse in ninth-grade US History academically diverse students. Moreover, the current study explores whether two forms of process support make a difference with diverse learner groups. We compare two types of structured

academic controversies with students who receive instruction in co-taught vs. honors classrooms. In addition, while most instruction was conducted face-to-face, students watched a video in which adults modeled and gave specific instructions for how discussion should progress (from hypothesis, to finding and evaluating evidence, to deliberation).

Prior research has established the viability of a cognitive apprenticeship approach to teaching students to write historical argumentative essays (De La Paz et al., 2014, 2017, 2021; Monte-Sano et al., 2014; Wissinger & De La Paz, 2016). However, work is still needed to determine the role of disciplinary discourse as a means for students to make sense of documents for themselves. Moreover, the literature points to a need for structure in discourse, both in group formation and in process support, and deliberative argumentation to achieve positive learning outcomes (see Chapter 2). The current study aims to provide a rich description of how two distinct SAC that integrate critical questions during the discussion phase of a CA approach to instruction work. This study aims to address the following primary research questions:

1. What is the quality of students' historical thinking during discussions?
    a. What types of dialogue characteristics are evident in small and large group discussions? Do students engage in deliberative or disputative argumentation?
    b. How do students evaluate sources/make judgments?
2. What can we infer about students with different incoming academic and cognitive profiles?

Method

Research Design

This study is a descriptive case study using a multiple-case study design (Yin, 2018) that involves collecting and analyzing data from two teachers in three class periods (called Mods in the participating district). Case studies permit in-depth analysis of a case, and having multiple cases allows for cross-case analysis to detect patterns across cases (Yin, 2018). The case study design was selected because it offers rich, in-depth, and insightful descriptions of the relationship between student discourse and discourse structure, as well as the quality of the discourse, by identifying similarities and contrasts across cases (Merriam, 1998). According to Miles & Huberman (1994), including multiple cases in a study can bolster "the precision, the validity, and the stability of the findings" (p. 29).

In this study, we explored three cases, which are detailed below. The first case (henceforth, "Co-taught") examined how four students used an enhanced SAC (i.e., one that provided an explicit, guided routine for deliberation) to support discussion in one of the teachers' co-taught classrooms. Students in co-taught classrooms generally had more learning challenges than students in other settings regardless of whether they had disabilities or not. The second case involved a different, four-student group with the same teacher (henceforth, "Honors 1" who used the enhanced SAC (henceforth, "SAC+G") in her honors Mod. Doing so allowed us to compare the same process supports with students with differing incoming cognitive characteristics. Lastly, the third case (henceforth "Honors 2") involved four students in a different

honors Mod, with a teacher who used Johnson and Johnson's (1988) original approach to holding a SAC to drive discussion. Students in honors classes generally exceeded academically compared to peers in other settings. The purpose of the contrast between Honors 1 and Honors 2 classes was to explore how students with similar academic and cognitive profiles used different discourse structures.In the larger project, the second author had an MOU with the school district that specifies provision of six historical investigations that address topics specified by the district, along with teacher professional development (PD). Much of this has been described elsewhere, and results from the first two years of our work has been published or presented at national educational research conferences. Finally, there is ample published empirical evidence on the use of the overall approach (i.e., a CA approach to reading, thinking and writing historical essays) in terms of improving culturally, linguistically, and academically diverse middle and high school students' written argumentative essays.

In our approach, students spend one day annotating, one day discussing and one day writing about historical controversies. To extend previous work, we focused on two opportunities: (a) teachers had been observed to struggle eliciting students' historical thinking (i.e., sourcing, corroboration, and contextualization) during small and whole class discussions (especially so with students who have been identified as having LD or who are ML); and (b) we have yet to uncover the degree to which critical questions or a structured approach to deliberation lead to improved disciplinary discussions.

While all U.S. history teachers in the district had access to the instructional materials (in digital and traditional formats, with differentiated reading and writing scaffolds and translated primary source documents) and personalized PD, in the current academic year we worked with two purposefully selected high schools, with very different student demographics, and eight ninth grade US history teachers.

Setting and Participants

The study took place at one high school. This was necessary as the lead social studies teacher at the second high school opted for a lesson on a different historical topic for the fifth HI. We collected teacher fidelity data at each school and student work (written annotations and final essays) from students whose parents consented to provide us with demographic information (e.g., social studies assessment scores, disability status, primary language, etc.). Student participants were purposefully selected using additional criteria: (a) attendance in a participating teacher's class period for both semesters; and (b) teacher nomination and student confirmation that the student was comfortable speaking with peers in small group discussions.

Cresswell High School (CHS)

Data were collected in three, ninth-grade US history Mods in a middle to upper income suburban, public high school in the mid-Atlantic region of the United States. According to the latest available data from the National Center for Education Statistics (NCES) for the 2022-23 school year, the school has a low poverty rate as measured by FARM data, with 13.8% of students qualifying. The racial demographics at CHS are as follows: 0.4% American Indian/Alaska Native, 7.6% Asian, 20.2% Black, 12% Hispanic, 0.2% Native Hawaiian/Pacific Islander, 54.2%

White, and 5.2% two or more races. CHS used an A/B-day block schedule, whereby students attend the class every other day, and each period consist of 71 minutes.

Teachers

Ms. Adams is a White, female teacher in her first year as a teacher. The previous school year, she interned at the same school with a ninth-grade US history teacher who had also participated in the second author's research. Later that year, she became a substitute teacher for another participating ninth grade US history teacher who took a leave of absence. Ms. Adams routinely provided instruction that was responsive to the needs of her students. She created additional scaffolds to help students remember content in previous lessons. She also showed great rapport with her co-teacher and discussed the instructional goals and purpose of the stages of the cognitive apprenticeship during planning meetings. Her high level of energy was an asset to instruction, and her pacing of lesson content was appropriate.

Mrs. Fern is a White, veteran female teacher. She was the department chair (DC) at CHS, and she actively engaged with the researchers and her fellow teachers in the larger study. She facilitated collection of consent forms and communicated intended adaptations of researcher lesson plans for feedback before use. She also led most US history meetings, though she encouraged the rotation of leading the meetings among her colleagues. Due to her role as DC, Mrs. Fern only taught one honors US history Mod, though she also taught a Mod of AP psychology. Mrs. Fern won a prestigious statewide social studies award, the Civic Education Coalition Civic Leader of the Year award, during the year our study took place. She also actively participated with the researchers in prior years of the research practice partnership

(RPP). Finally, she provided high fidelity to teaching the HI using a CA approach to instruction for this study.

Students

Seventy-two ninth-grade students participated in the current study. Table 3.1 shows the demographic information for all students in the six Mods. All participating students were familiar with the first author, who was present to conduct the audio recording. We selected one Mod from Ms. Adams to represent the teacher's overall population (i.e., all Mods), taking care to reflect the same demographics. Since Mrs. Fern only taught one US History Mod, this was not possible.

We then purposefully selected four students from each of the selected Mods considering (a) demographics (i.e., gender, IEP, 504), (b) performance on a district history assessment at the beginning of the school year, and (c) adequate discourse data for each student based on transcripts. In addition, students had to be in the same small group. Twelve students comprised the final sample (see Table 3.2). All student names provided are pseudonyms.

Co-Taught

Two pairs of students (Ford and Kristen, and Bryce and Julius) were taught by Ms. Adams, in her third period co-taught Mod. The co-taught Mods included SWD and students without disabilities. Two of the students were identified as SWD; Ford had a learning disability (LD), and Bryce had a diagnosis of other health impairment (OHI). The co-taught Mods included a special educator and a paraprofessional, who both were present during instruction. Ms. Adams used our SAC+G to guide the classroom discussion. The co-teacher pulled out students to the library who did not

have consent to participate. The 12 target students were present in the classroom. Students sat in pairs, and an audio recorder was placed on the table between the students to capture their discussions.

Honors 1

The second case, Honors 1, consisted of two pairs (Heather and Rex, and Jake and Taylor) in Ms. Adam's fifth Mod. There were 31 students in the classroom, with no pull-out groups. The seating arrangement was similar to the co-taught Mod, and the audio recorder was similarly placed to capture the student pair's discussion. While Ms. Adams circulated the classroom during discussion activities, the large number of students precluded her from interacting with some of the students in this case. Ms. Adams used the SAC+G with this Mod as well. The school held mandated assessments on the originally scheduled Day 2 of the HI. Thus, students returned to the task four days after Day 1 due to the A/B block schedule.

Honors 2

Mrs. Fern led the third case (Honors 2), which consisted of four students, Paisley, Nel, Yogi, and Isle. Mrs. Fern used an SAC that students used in similar discussion activities in previous HIs. Her setup for the day comprised small groups of four students. Students sat in groups of four, and the audio recorder was placed on one desk within recording distance of all group members. The structure of the discussion included four steps, and each step consisted of a small group discussion followed by a whole class debrief. There were 30 students present in the classroom at the time of the study. Students in this Mod, just as in Ms. Adams' Honors 1, also missed the original Day 2 due to the mandated assessment.

Task Design

We created a series of HIs for the cooperating school district in prior years of the project. This study focuses on HI 5, the final investigation in which students discussed contrasting sources together before writing (see Table 3.3 for an overview of the six historical investigations). Students explored one central historical question (CHQ) in each investigation. Our original goal was that students learn to develop a particular stance that is supported by evidence and makes sourcing, contextualization, and corroboration judgments (determining the influence of and reliability of the source, the context, and the quality of the author's facts and examples in relationship with the overall CHQ). Specific scaffolds, or tools to help teachers and students engage in this type of thinking have been described as IREAD and How to Write, or H2W, have been well described in published reports with culturally, linguistically, and academically diverse students (see De La Paz, Otarola, and Butler, 2023 for a report on the use of virtual professional development using this approach with high school social studies teachers during the pandemic).

The CHQ for HI5 asked students to consider whether the construction of the inner-city harbor, "Harborplace", in the late 1970s in Baltimore would improve life in the city and based on their reasoning to decide whether they would have voted to approve the construction of the inner harbor at the time in which it was debated. Students are provided with background information and primary sources related to the CHQ. Instruction during the three days is modeled after prior HI instruction (De La Paz et al., 2022). The focus of this study is the discussion, which occurs during the second day of instruction. Students in the co-taught Mod read and analyzed two

documents, while students in the honors Mods analyzed four documents (see Appendix A). The documents detailed the various potential impacts of the construction of Harborplace. Two documents had a more positive perspective upon the construction, such as the potential for increased economic activity. The other two had a less favorable perspective upon its construction, such as the displacement that occurred in poor, predominantly African American neighborhoods.

Structured Academic Controversy

The SAC is designed to engage students in structured, cooperative debates on controversial issues, fostering critical thinking, perspective-taking, and collaborative problem-solving skills (Johnson & Johnson, 1979, 1985, 1988). Students in Honors 2 used a SAC and were provided with critical questions to guide students' historical thinking. Using the SAC promotes higher-order thinking skills, such as analysis, evaluation, and synthesis (Nathan & Lee, 2004). Johnson and Johnson (1988) outline that the teacher's responsibilities include defining the objectives, preparing for the activity, explaining and structuring the task and controversy, supervising student performance and intervening when necessary, and evaluating and reflecting on the lesson. Initially, we provided teachers with a SAC designed for use in student small groups, allowing teachers to determine the group size at their discretion.

The original SAC graphic organizer begins with students discussing and documenting their group's initial ideas on the historical investigation. Teachers offer students at least two options for the CHQ. Students then decide whether they support one perspective of the CHQ and provide evidence in the form of quotations from one of two texts that align with that perspective, promoting corroboration. Subsequently,

students repeat this process for the opposing perspective. Finally, students discuss two questions: one regarding the author's perspective and the other concerning their final decision. The question about the author's perspective is designed to prompt students to evaluate the author's reliability or credibility critically. The teacher then reconvenes the class for a final discussion on students' decisions regarding the CHQ and their evaluations of the authors' perspectives. Following this discussion, students record their final stance on the CHQ, which they use to prepare for writing a final essay. In sum, the SAC method is grounded in the principles of cooperative learning and encompasses active engagement and critical thinking to understand and respect differing viewpoints.

Guided Structured Academic Controversy

In an observation of a previous HI two months prior to the HI in the current study, we noted a lack of participation from most students in the discussion. Moreover, students appeared confused by the CHQ, and Ms. Adams had to clarify what the CHQ was asking. The field notes indicated that more students became engaged in the discussion as the teacher circulated but overall had more off-topic conversations than on-topic. This was consistent with prior observations on Day 2 of HI, so we decided to revise how students could engage in discussion. Students in Ms. Adams' co-taught and honors Mods used a more supportive and guided SAC. The SAC+G (SAC plus guided support), which we used as process support, was designed to be responsive to SWD's needs. It is based on the general premise that providing explicit instructions to students, in the use of a turn and talk routine (Swanson et al.,

2019), and in the use of modeling and using a graphic organizer leads to increased discourse engagement and quality.

Students were prompted to engage in domain general and disciplinary reading utilizing the IREAD tool and critical questions for evaluating sources and evidence. Next, students began discussion in pairs, then each pair merged with another pair to create a small group of four students who engaged in further discussion before a teacher-led whole class discussion. Before its use in class, teachers showed students a video where members of the research team modeled its use. This allowed students to have explicit instruction in disciplinary discussion, which was intended to help them to visualize themselves engaging in discourse, potentially leading to more students participating during instruction, and ideally, higher quality discourse. We modeled this after Casado-Ledesma et al. (2023), though they used students as participants in the video modeling.

Teachers guided students in small group and large group discussion, incorporating discussion facilitation suggested by Monte-Sano et al. (2017) and Swanson et al. (2019). As students engaged in discussion, they completed a graphic organizer (GO) to record their ideas (see Figure 3.2). Students subsequently selected ideas from the discussions independently, utilizing the completed GO as a planning tool to organize textual evidence they deemed most credible and the judgments that supported their beliefs. Completing the graphic organizer resulted in key information that could be effectively used when composing their historical arguments.

Data Collection

Data included teacher and student audio recordings, permanent products (written graphic organizers, digital work), semi-structured interviews with teachers after data collection, and field notes. Target student groups were audio recorded. During whole class discussion, one audio recorder captured discussion between the teacher and her students. The semi-structured interviews took place one week after the instruction via a virtual meeting. This gave teachers an opportunity to reflect on how the discussion unfolded, suggestions for future modifications (i.e., informal social validity check), and to answer the first author's questions. We examined five student transcripts of audio recordings ranging from 28 to 35 minutes. These transcripts come from audio that was edited to only include the relevant discussion (e.g., video modeling was not included, periods of silence). All audio recordings were transcribed using online software (Rev). Finally, we used observations, field notes, and interview data to compare with audio recordings for triangulation.

Data Analysis

Our approach to conducting a cross-case analysis (Yin, 2018) of classroom discourse compared classes using the SAC+G (comparing students in the co-taught class with Honors 1) and compared outcomes in Honors 1 and Honors 2 (with the original SAC and SAC+G). We did not compare the co-taught class to Honors 2 because we anticipated that SWD were at a disadvantage without significant support in the highly structured routine.

The discourse analysis in the current study requires an in-depth analysis of how students' discussion was developed to unpack the relationships between small group and whole class discourse and the development of historical thinking. The unit of analysis was an utterance contributing to the development of knowledge. We analyzed each utterance, which we define as one turn at talk, for codes. Chen et al. (2016) suggested researchers study separate stages in discourse, and we adapted their approach in science to our work. The three stages are: (1) identifying distinct activities within a class, (2) identifying events within an activity, and (3) analyzing data related to discussion by using a cross-case analysis method.

An activity constituted a step of the SAC+G graphic organizer, which could include more than one related question, such as Q1 and Q2 (see Figure 3.2). This was similar for Honors 2, though each question of the SAC+G constituted a step and an activity for this case. An event would constitute the responses uttered by the students and teachers related to an activity. Where applicable, we used multiple codes for the same utterance.

Sixteen codes emerged: initial answer, final answer, reasons (for and against), evidence, sourcing, contextualizing, causal inferencing, rebuttal, quality of facts, corroboration, reaffirm, elaborate, add idea, oppose idea, and prompt or question. Utterances not directly relevant to the task were coded as other. Next, the codes were organized into historical thinking based on the literature: evidence, sourcing, contextualization, causal inferencing, quality of facts, and corroboration. All transcripts were double coded by the first author and by a researcher familiar with the

larger project to determine interrater reliability. Interrater reliability was calculated at 87% exact agreement. Any disagreements were resolved through discussion.

The semi-structured interviews with the teachers were systematically analyzed and coded. The following themes emerged: (a) structure increased struggling learners' engagement in discussion; (b) the quality of discussion was overall higher quality compared to other lessons; (c) writing and discussing was overwhelming for some students prior to SAC+G; and (d) honors students in Ms. Adams' Mod participated less than other peers and required extrinsic motivation to engage in discussion. In addition, we analyzed the students' graphic organizers and field notes from a prior HI.

To measure the quality of historical thinking discourse, we adapted historical thinking writing rubrics for sourcing (Nokes, 2017) and contextualization (Monte-Sano & De La Paz, 2012). In line with Nokes' (2017) scoring for sourcing, we used a spectrum and focused on whether students mentioned the source, whether students used ahistorical thinking in their attempts to source, and whether students used sourcing information to critically evaluate sources (see Table 3.4 for examples). In scoring for contextualization, we focused on the degree to which students accurately identified and situated their arguments within the appropriate temporal, spatial, and situational frameworks, thereby linking related events (see Table 3.5 for examples). We scored the transcripts similarly to how a written product would be scored. All transcripts were scored by the first author and by a researcher familiar with the larger project to determine interrater reliability. Interrater reliability was calculated at 89% exact agreement. Any disagreements were resolved through discussion.

Results

We report the results of this cross-case analysis by research question, comparing the co-taught Mod with Honors 1, and comparing Honors 1 with Honors 2. We captured utterances as argumentative and historical thinking. Argumentative encompassed utterances that reaffirmed ideas, elaborated ideas, added ideas, opposed ideas, rebutted ideas, and prompted or questioned ideas further. Historical thinking included utterances related to textual evidence, sourcing, contextualization, causal inferencing, quality of facts, and corroboration. We also corroborate students' transcripts with their written graphic organizers. Finally, we include the teachers' semi-structured interviews to support the research questions' findings.

RQ1: What is the quality of students' historical thinking during discussions, and do students engage in deliberative or disputative argumentation?

To answer the second research question, we first examined the utterances that involved historical thinking. Historical thinking was comprised of utterances related to using evidence, sourcing, contextualizing, causal inferencing, quality of facts, and corroboration. The AFL framework facilitated the coding of these utterances through process support, specifically discourse instructions. Our instructions prompted students to think like historians, so we asked questions about the source and the evaluation of documents using historical judgments (e.g., contextualization). Next, we examined the quality of students' use of sourcing, which we modified from Nokes (2017) and is found in Table 3.4, and the quality of contextualization, which we modified from Monte-Sano and De La Paz (2012) and is found in Table 3.5. Finally,

we looked at the overall arguments students made to determine whether they engaged in deliberative or disputative argumentation. We utilized the AFL framework during coding as we looked at the process support, specifically discourse instructions and teacher scaffolding. Our instructions were designed to engage students in deliberative argumentation, and the teacher scaffolding supported such discussions.

Similarities in Historical Thinking Between Co-taught and Honors 1

The total numbers of historical thinking utterances in Co-taught and Honors 1 were similar (see Figure 3.5). This demonstrates that students in Co-taught engaged in historical thinking to the same degree. For example, Ford used causal inferencing in the following statement, "Basically I say yes because due to the amount of how much entertainment and how much shop opportunities there will be and whatnot in that area, it will cause the people to probably go there more often and have jobs there for some people that cannot get a job normally." Ford inferred a cause-and-effect relationship between the construction of Harborplace and the attraction of people to the area, which led to more jobs. Similarly, Heather made the following observation, "Yeah, it's all about the displacement of, I don't know, lower class citizens of Baltimore. This particularly affected Black Americans who had moved in and are now being displaced, meaning more crowded poor areas and homelessness." Heather implied a cause-and-effect relationship with displacement leading to crowding and homelessness.

Students in Co-taught and Honors 1 used corroboration, and Ms. Adams prompted students in both classes. During step 6, she asked each Mod if the

documents agreed, to which most students responded no. Ford explained how the documents did not agree, “I would say that they do not quite agree because Document A talks about how all these things are going to possibly be happening to the city and whatnot and people are not quite reassured of it and plus they're kicking people out of their homes just to build things, which is a bad thing for a popularity rate. But with Document B, it's talking about how the opportunities of jobs that could possibly happen if that harbor thing is built.” In Honors 1, Rex concurred that the documents did not agree and provided a simpler explanation, “Because they have different opinions.” Ford articulated his corroboration judgment at a higher level than Rex, as he provided solid reasoning compared to Rex’s basic explanation. Furthermore, when asked about the quality of the discussion, Ms. Adams responded: “The discussion was really good. I would say… compared to like regular lesson discussions.”

Differences in Historical Thinking Between Co-taught and Honors 1

The quality of the historical judgments differed, as students’ discourse varied based on their ability to critically analyze sources using multiple historical thinking skills. For example, Bryce attempted contextualization (italicized), “All right, so I had document B… so it's pro African American *in this time*, which is huge, so they just want more job opportunities for the less fortunate.” Bryce showed that context is important by emphasizing it as “huge”, but he did not go further with his judgment, giving it a score of a Level 2. By contrast, Taylor’s contextualization was more sophisticated, “Our author has lived in Baltimore for a while, so he knows the area and knows a lot of people in the area probably. So, he would understand the problems

that are going in all Baltimore." Taylor implied that the author is a long-term resident of Baltimore, so he was more likely to be informed about the city. This is an example of a Level 4 contextualization judgment.

The quality of sourcing judgments also differed between students in Co-taught and Honors 1. Julius made the following judgment, "On the document perspective, it's from Ron Howell, a reporter for the Baltimore Sun, and his answer is no to whether or not he supports the building of the Harborplace in Baltimore because this quote from the source he said that 'this project will have an effect on the future racial socioeconomic makeup of several Baltimore neighborhoods'." Julius properly attributed the source of the document but did not critique the document to strengthen his argument. This is an example of a Level 2 sourcing judgment. On the other hand, Heather made a more elaborate sourcing judgment:

> "So, I guess, the first source's perspective, the author is Ron Howell, and he wrote for the Baltimore Sun, so it's a pretty credible source right there already, they were pretty popular. And yeah, in good favor in journalism I guess. So yeah, he was talking in favor really. He interviewed Mr. Quayle, Quayle, I don't know – who's one of the leading voices in the anti-displacement organization there. And so he was speaking in favor of those being displaced by other upper class citizens. And it was really to inform people about just how big the impact is going to be for a lot of Black Americans that were displaced from revitalized neighborhoods and what the impact would be the city as a whole."

Like Julius, Heather noted the basic sourcing information (e.g., author, publication) but critiqued the document to strengthen her argument. She delved into the reliability

of a journalist and their purpose ("to inform"). This is an example of a Level 4 sourcing judgment.

Similarities in Argumentation Between Co-taught and Honors 1

Students in both Mods made similar argumentative moves (see figure 3.3). The only major difference was in opposing ideas, in which students in Honors 1 did not make any such utterances. Through their argumentative utterances, students in both Mods demonstrated elements of deliberation in answering the CHQ. For example, students in Co-taught did not change their answer to the CHQ from their initial answer to their final answer. Kristen and Julius answered "no", and Bryce and Ford answered "yes". Despite not coming to a consensus, these students showed some signs of engaging in deliberation as they weighed each other's ideas, acknowledged them, and were open to others' ideas. Furthermore, they did not dispute each other by trying to persuade one another of their answers. Students in Honors 1 had varying responses initially. For example, Heather answered "both", Jake answered "no", and Taylor and Rex answered "yes", and Heather and Jake later changed their response to "yes" in their final answer, as indicated by their graphic organizers. This suggests deliberation among these students, as they constructed knowledge and their argument together and arrived at a consensus.

Below is an excerpt from Honors 1 that illustrates the class constructing an argument together.

Ms. Adams (22:45): "…Document B, what do we got? Yeah."

Student 1 (22:52): “In general, document B is pretty supportive towards it as most of the article was talking about how it's recruiting qualified minorities to join its staff.”

Ms. Adams (23:03): “Okay.”

Student 2 (23:04): “For not just the construction, but the years to come if they need workers for it.”

Ms. Adams (23:07): “Yeah. What types of jobs do you need in an aquarium?”

Taylor (23:11): “Gift shop.”

Ms. Adams (23:12): “Yeah, you got a gift shop, someone’s got to take the tickets. Someone has to clean the glass when people slobber on it. [Student 3]?”

Student 3 (23:19): “People who take care of the animals.”

Ms. Adams (23:20): “People who take care of the animals. It would be like you can't just have the gift shop running and then being like, I don't know, those fish just swim around and we don’t know what to do with them.”

In reviewing the advertisement from the Baltimore Afro-American (see Appendix A), Ms. Adams allowed the students to choose in which direction to go. Once the students discussed a topic of interest (e.g., jobs for minorities), she helped the students elaborate on the types of jobs. Students could borrow from these ideas as they built their own argument.

In both cases, one student from each pair demonstrated initiative in leading the conversation. In the Co-taught Mod, for example, Bryce and Ford contributed first at each step and at one point prompted their partners for a response. In Honors 1, Hannah and Jake played a similar role, often beginning the discussion. Figure 3.3 shows the frequency of argumentative moves between the co-taught Mod and Honors 1. Students made similar number of moves to reaffirm, elaborate, add ideas, and prompt or question each other to build an argument. Below is an excerpt of Ford,

Kristen, and Ms. Adams engaging in a sensemaking discussion in the Co-taught Mod that shifts to an argumentative discussion.

> *Ford* (00:58): "So, it's basically talking about on how the harbor should be improved in life by doing all these external things and whatnot and add them into the city."
> *Ms. Adams* (01:13): "What do you think?"
> *Kristen* (01:16): "Whether, I think they'll improve it or not?"
> *Ms. Adams* (01:19): "It can be either. What do you think the question's really asking? What could be some of the perspectives?"
> *Kristen* (01:25): "I don't think it will improve it because it was asking people to move and stuff. So, it was affecting people."
> *Ms. Adams* (01:35): "And what could've been the other argument for that? What do you think?"
> *Ford* (01:44): "Well, I think it'll improve it actually because you're expanding more to a public area where people can set up businesses and that's more job opportunities for people in cities and more income coming in, more fluctuations. I think that'll improve."

Ford and Kristen both added ideas, and Ford elaborated on his idea at the end of the excerpt. Ms. Adams prompted Kristen, who was quite shy, for her opinion. After Kristen responded that she did not believe the construction would help, Ms. Adams asked for a different argument, and Ford added his argument for the construction of Harborplace. Ford, Kristen, and Ms. Adams engaged in the beginning stages of argumentation. Similarly, in Honors 1, Jake and Taylor held an argumentative discussion, as illustrated below.

> *Jake* (00:09): "What'd you find about document A?"

> *Taylor* (00:11): “So with document A, I found that a lot of the refugees, you could say, well, essentially less fortunate people that are being kicked out of this area because their houses either being foreclosed or they're selling them. And then rich people were coming in, buying up their property, fixing it up to make it look very nice, and then selling it for a higher price. And they're doing this to neighborhoods and blocks. And they're not having any space where the less fortunate people, or for people [to] have any space to live because they don't really have any affordable housing.”
>
> …
>
> *Jake* (02:47): “So for our historical question, the construction of Harbor Place would be beneficial to the growth of Baltimore because it added tourists industry since all the actual industry got taken away. And Baltimore is known for bad things, so it'd be nice if they incorporated that. So, it would be added for good things.”
>
> *Taylor* (03:12): “For me, my personal opinion, this construction, it's done so much to people. They got kicked out of their homes, and they had to rebuild and all that stuff. But what is coming after could possibly have good impact on that city because you can bring in more people with it.”

Students in both cases demonstrated an understanding of the CHQ and identified reasons for and against the construction of the Harborplace, adding ideas and elaborating on ideas. They both identified displacement of people and some economic benefit as reasons to support or not support the construction.

However, not all discourse was argumentative. Asterhan and Schwarz (2016) made a distinction between argumentation and explanation based on the purpose of the dialogue. Thus, though the dialogue may sound similar, the distinguishing factor is the purpose. Below is an excerpt of consensual co-construction, a form of explanation:

> *Bryce*: Yeah. What do you think about construction of Harborplace in Baltimore, do you think it improves life?
>
> *Julius*: No.
>
> *Bryce*: Why not?
>
> *Julius*: I feel like money could be used towards homelessness, poverty.
>
> *Bryce*: Yeah, that's true. The money could be used towards that, but it also opens a more public area where small businesses could be opened up and provide and provide more opportunities for less fortunate people like homeless people.
>
> *Julius*: Yeah, I didn't think about it that way.

Both students provide their answer to the CHQ and reasons to support their answer. The purpose of their discussion is to explain their thinking rather than to make an argument.

Below is an excerpt of Jake and Taylor's discussion regarding whether the construction would be beneficial.

> *Jake* (08:23): "My personal opinion, the construction Harbor Place, it will not improve life."
>
> *Taylor* (08:29): "My opinion is that it will improve life because it's jobs to minorities, and also is opening up big revenue streams for the city of Baltimore. That should overall improve life."
>
> *Jake* (08:42): Yep. The reason why I say no, this will not improve is because, they're basically kicked people out of their homes this entire time just trying to build this, and so they're basically not going to be able to come back to where they used to live. So that forces them out of their home to try to find a new home when they're still limited on money. (0:18)

Jake and Taylor provided their personal opinion, but they did not engage in disputative argumentation or deliberative argumentation in this example. Rather, Jake reaffirmed Taylor's opinion in the beginning of his second utterance and then

elaborated on his opinion with a solid reason. This is an example of Asterhan & Schwarz's (2009a; 2016) consensual co-construction dialogue characteristic, which suggests that students were focused on making sense of the documents as opposed to arguing with them.

Differences in Argumentation Between Co-taught and Honors 1

The excerpts above also demonstrate that the detail provided by the students differs. Jake and Taylor provided greater argumentative detail, unprompted by Ms. Adams, than Ford and Kristen. Another difference between the pairs' discussion was Taylor alluded to reasons to support and not support the construction. Kristen suggested she would not support the construction and provided a reason when prompted by Ms. Adams, and Ford suggested he would support it. Neither Ford nor Kristen addressed a different perspective.

Similarities in Historical Thinking Between Honors 1 and Honors 2

Echoing Ms. Adams' reflection on the quality of the discussion, Mrs. Few said, "I would say that the quality of the discussion was really good, very, very strong." Students in Ms. Adams' Mod and in Mrs. Fern's Mod had a similar number of sourcing utterances (see Figure 3.6). Moreover, the quality of the contextualization judgments was similar. Hannah made the following statement during the whole class discussion, "And on the other hand, it's a pretty credible and still now pretty credible news source. And this was one of, I don't know what employment in magazines looked like for African Americans at this time. But I think that this was going out to the Black families that had been kicked out themselves, and I think... But I mean, as a

whole, it was probably pretty credible. But yeah, I would yeah." Hannah discussed the reliability but also the context of the source, as it was an advertisement for jobs in the Harborplace, and its target audience was African Americans in Baltimore. This is an example of a Level 4 contextualization judgment.

Similarly, students in Mrs. Fern's Mod made high-level contextualization judgments, as seen in the following excerpt during a whole class discussion:

> *Mrs. Fern* (06:11): "What do we know about the time period? What is some information we know about this period of time that might help us to answer the question?"
> *Paisley* (06:22): "The civil rights movement was happening and at the time a lot of middle income or wealthy White people had left because of de-industrialization. So, to get people back to the city they were trying to make over or make it be more appealing to White people again, the wealthy White people. So, they were displacing people out of their homes to try and make it seem more appealing and moving them. But that wasn't great because they were moving the minorities and a lot of it centered around the civil rights movement because there's a lot of protesting and things like that in Baltimore."
> *Mrs. Fern* (07:07): "Outstanding, outstanding contextualization. That's a massive, amazing judgment."
> *Student 1* (07:14): "There was stagflation."
> *Mrs. Fern* (07:15): "Ah, yes. Good one. Another important thing we know about the time period."

Mrs. Fern prompted students for contextual information, and Paisley provided two pieces of contextual information, the Civil Rights Movement and de-industrialization. Another student mentioned stagflation in the economy, which students knew was a bad economic situation as it was provided in the vocabulary for this HI. The excerpt

shows contextualization at a Level 4. Mrs. Few agreed that the level of historical thinking was high, adding, “I heard a couple of kids you know make judgments of maybe what you know the author was trying to convey, or what their angle might be, kids who threw out ideas that I wouldn't, that I didn't, think of. So, I would say it was a very good discussion.” In sum, both Mods showed high levels of historical thinking in their discussions.

Differences in Historical Thinking Between Honors 1 and Honors 2

Students in Honors 2 also had more utterances in corroboration, quality of facts, and using evidence. Moreover, students in Honors 2 had twice the number of contextualization judgments as the students in Honors 1. Furthermore, the quality of the sourcing in both cases varied. In Honors 1, the following excerpt demonstrates high-quality sourcing and argumentation:

> *Taylor* (19:14): “So, with ours, we read over it and read some of the footnotes. And we saw that [Ron Howell] who wrote this article works for the Baltimore Sun, which is a magazine. And also, has lived in Baltimore for a while. So, we were saying that the source is reliable, because there's nothing really for him to benefit of if he is to lie so they can [inaudible] very well. And also, since he knows the Baltimore area a lot, that he would know the most information and get the actual real story.”
> *Heather* (19:47): “Yeah. [Jake], what do you think about that?”
> *Jake* (19:49): “I think it… You said all the words I was going to say. But just because the Baltimore Sun, I don't know why he would lie about stuff that's happening in town.”
> *Heather* (19:49): “Because it was very factual.”
> *Jake* (19:49): “Yeah. Yeah.”

Taylor (20:05): "Because there's no real reason to lie."
Jake (20:06) "Because if you lie, you're caught."
Heather (20:09): "And the Baltimore Sun is a very credible newspaper even now."

In this discussion, three of the four students in Honors 1 participated substantively. Moreover, Taylor provided a Level 4 sourcing judgment by not only providing sourcing information but also judging the author's reliability. It is also evident that students engaged in deliberation, as they collaborated to construct an argument, specifically that the author was reliable because he worked at a renowned newspaper in Baltimore.

By contrast, students in Honors 2 engaged in lower-level sourcing. The sourcing excerpt is an example:

Nel (25:25): "Okay. I said I selected the perspective of the author of document B. I put document B but I spell it B-E, I don't know why I do that. Because they give excerpts from news articles."
Paisley (25:39): "I said that overall, it's like documents based off of their evidence, who the author is and what their intention may be. I selected the perspective of document D because they have a lot of statistics and research in it. And it comes from, they did a conference or a press conference where they talked all about... Is that this one where they... Yeah, they did a press conference and all of the leaders of it talked all about it. But they could be embellishing or lying about it. But because of who they are as people and their jobs, I selected document D."

Nel and Taylor attempted to source the documents they mentioned, but they did not provide the specific sourcing information. This is a Level 1 sourcing judgment. While they understood that documents have authors, they did not explicitly state that. This is

most likely due to the differences in the structure of the SAC. The SAC+G used in Honors 1 explicitly asked students who the authors of the documents were. The process support differed as the graphic organizer used in Honors 2 did not explicitly ask students to source the documents.

Argumentation in Honors 1 and Honors 2

Students in both Honors Mods also provided rebuttals, seeing both sides of the issue. Below is such an instance in Honors 1:

> *Rex* (11:04): "I believe it will, because it will for the upper class, but it won't for the lower class, since the lower class will be kicked out their homes. But the upper class would be able to find newer jobs, better place to live, a cleaner environment, and just in general, a better place to be in."
> *Heather* (11:28): "Yeah, so it would really just benefit upper class citizens by giving them the luxury of the aquariums, of all these different things, and really just judging by Document two, the only really chance... The only benefit this would bring is jobs like being a janitor, all that jazz. In the Baltimore aquarium for example. But really, that's the only benefit that they'd be getting, and that's, yeah, it really does just benefit wealthy people in the city."

Rex weighed both sides and though he said he believed the construction will improve life in the city, he added the caveat that it would only be for the wealthy since lower-income people would be displaced. In a prior step, Heather similarly weighed both sides. Both students had strong rebuttals for whichever position they ultimately took.

In Honors 2, students also weighed both sides carefully. The following excerpt demonstrates instances of students making a rebuttal:

Paisley (17:36): “Okay. With document B or document two, I agree. I would've voted for it because it's creating jobs for minorities, which kind of counter... I don't know what that word is, but it kind of evens out the playing field because the argument was that people were getting displaced and that they were losing their jobs in their area, especially minorities. But if they were creating jobs, then they could get money to move back where they used to live or just have enough money to sustain themselves. And I used, ‘If they recognize the need to hire minorities.’ And I also used, ‘If the Baltimore Aquarium is to be successful, it must be a community-oriented facility in all areas.’”

Yogi (18:32): “I said that I would not vote for the construction of the Inner Harbor or document four or document D mainly just because the whole point of the document is saying that a lot of people were questioning the financial legitimacy of the partnership with the construction company, Rouse Company. So, it was more so of a lot of people were like, eh, is this credible? We don't know. So, seeing that is if you have to question it and a lot of the public is questioning it... Eh, I don't know.”

Nel (19:06): “All right. [Isle]?”

Isle (19:08): “Yeah.”

Nel (19:09): “Wait, wait. So would you vote for it?”

Yogi (19:11): “No.”

Isle (19:14): “I said I probably wouldn't have voted for it just on document D because it really only talks about the profits people are going to make off it and there's no talking about how it's going to affect people.”

In this discussion, Paisley agreed with Nel that she would have voted for the construction of Harborplace, mainly because it would attract jobs for displaced minorities. Yogi and Isle countered with two distinct reasons; Yogi said the finances were questionable, and Isle said the construction was about making profits and not

about the people. Here again, the students created an argument with two perspectives and strong reasons for each perspective.

The number of argumentative moves shown by students in both cases was similar, as seen in Figure 3.4. They elaborated, opposed ideas, provided rebuttals and prompted or questioned each other at comparable rates. There were slight differences in the number of reaffirmations and adding ideas, with Ms. Adams' students having fewer reaffirming utterances but more adding ideas. In Honors 2, students engaged in argumentative discussion. The excerpt below is one example.

> *Paisley* (02:02): "I'm honestly kind of divided on if it will or will not, because Baltimore was suffering financially and their economy was really hurting, which has a lot of different effects because that can cause prices to go up, which people that are low income would, it would be hard for them to buy food that they need and things like that. But those people were also being displaced, especially low-income people, African American people, especially in those little suburb areas. So, it hurt people, but it also helps people. So, honestly, it's divided between are you willing to help the people that need it the most or helping the city because it intertwines, it's a cause and effect."
> *Isle* (02:57): "I kind of put something similar. The construction might help them, that might help the city to look better, but it's taking people out of their homes. So how much good is it we're really doing?"
> *Nel* (03:08): "All right. I put the construction of Harborplace in Baltimore would directly improve lives of people living nearby the harbor and indirectly improve the entire city by giving it pretty much a heart. And..."
> *Yogi* (03:24): "I said it was more of a two-sided coin, more so, because on one hand it helps the city look better. And like Nick said, it gives it a bit of a heart, but at the same time, you're also giving jobs to the poor who are being displaced in the first place. So at least they're getting jobs, but then there's the

> whole reason of why they're getting jobs in the first place because they're getting kicked out and it's destroying where they live."
>
> *Paisley* (03:51): "Oh, yeah. I didn't really think about that part."

In this exchange, each student contributed ideas to answering the CHQ, each in turn building an argument. Moreover, Paisley, Isle, and Yogi saw both sides of the argument and provided reasons for both sides. The main reason against construction, in both cases, was displacement, and the main reason for construction was job opportunities. In building their initial answer to the CHQ, Paisley, Isle, and Yogi answered "both" and Nel answered "yes". By the end, all four agreed that construction would improve life in Baltimore, just like the students in Honors 1. This is a clear example of deliberation, as students engaged in discussion to reach a better understanding of the CHQ, weigh the evidence, and ideally arrive at a consensus. Finally, Mrs. Few said, "having them take a look at each document separately, I think, force them to really think about all of the varying possibilities and answers to the question." This suggests the structure of the discussion allowed her students to listen to peers' views about all the documents.

RQ2: What can we infer about students with different incoming academic and cognitive profiles?

To answer the third research question, student performance on a US History assessment administered by the district before the study took place was used to identify differences in individual characteristics (see Table 3.2; students earned scores of mastery, threshold, and below threshold on this assessment). In addition, we report

teachers’ reflection on how students of various academic and cognitive profiles engaged in discussion.

Academic and Cognitive Differences Among Cases

The first case was a Co-taught inclusion Mod, where SWD and other students who require more support are placed. Before this study, these students were reluctant to participate substantively to develop arguments. However, in this study, these students participated at a similar rate as students who did not receive the co-teaching. For example, Ms. Adams said, “For the most part, my lower [performing] students, I think, really benefited from this. Especially because I selected some low [performing] ones for the… actual recording, and they shined… I was really pleased with how they did.” With regards to her honors students, Ms. Adams said, “I think the participation was up for honors… from small group to the big group because they… like to feel like, important. And I think, knowing that there was someone there… observing them, and that there were recording devices even nearby.” Mrs. Few had a different experience. She said that some kids who usually participate, “they seemed to clam up immediately, because we had, you know, because you were there, and we were recording them. They knew that there was a possibility, so I'm not sure why they suddenly were like, I'm not going to talk today or whatever, and that was a little frustrating.” In sum, participation in the discussion was high for both Co-taught and Honors 1 but mixed in Honors 2, with the first author’s presence perhaps having an unintended impact.

Ford and Bryce both scored below threshold on the first district quarterly US History assessment (below 50%), which was administered to students at the beginning of the year. Ford received a score of 43.75%, and Bryce received a score of 40%. Kristen and Julius did not have IEPs and scored at the threshold mark (between 50% and 70%), receiving a 60.16% and 50%, respectively. Julius had a 504 plan, though we did not receive specifics about which disability. The quality of the sourcing judgments for all four students was at Level 2, and the quality for their contextualization judgments for three students was at Level 2, with Kristen at Level 1. Taken together, the students' placement in a Co-taught Mod and their results on the district assessment indicated that their disciplinary skills were at a lower level. However, there were notable exceptions. For example, Ford showed a more elaborate corroboration judgment than Rex, and Co-taught had higher scores on sourcing judgments than Honors 2.

At times, Ford had challenges either articulating his thinking or making inferences, though it is not possible to know which. An example of this, and how Ms. Adams engaged with him, can be found in the excerpt below:

> *Ford* (19:54): "This author might have known more about the possible outcomes of what could've happened to the city around the time of the- trying to build it and whatnot."
> *Ms. Adams* (19:54): "Why do we say that?"
> *Ford* (20:13): "Because the amount of time that the person lives there and whatnot."
> *Ms. Adams* (20:16): "Okay, and do you get that from the headnote? Where do we know that?"

Ford (20:30): "No, it's a reporter that's been working there for a while and probably some of their colleagues and whatnot have possibly lived there their whole life and whatnot."

Ms. Adams (20:42): "Does it say that?"

Ford (20:42): "No."

Ford attempted to make a historical judgment, and Ms. Adams pushed him to explain his thinking. As he explained his thinking, he realized, with the help of Ms. Adams, that the judgment he made was inaccurate.

On the other hand, students in Honors 1 generally performed better on the district quarterly History assessment. Three of the four students scored at mastery (above 70%), with one student ,Jake, scoring at threshold with 51.56%, Taylor scoring at mastery with 95.31%, Rex scoring at mastery with 71.55%, and Heather scoring at mastery with 93.75%. While these scores varied, these students also provided higher level historical thinking in their discussion. For example, the quality of sourcing judgments for Heather, Rex, and Taylor was at level 4, and Jake was at Level 3. The quality of contextualization judgments for Heather, Rex, and Taylor was at level 4, and Jake did not make any contextualization judgments. In sum, the students in Honors 1 made higher-level judgments than the students in Co-taught.

Similarly, students in Honors 2 performed better on the district quarterly History assessment. Two students scored at threshold, and two scored at mastery. Paisley scored at mastery with 93.75%, Isle scored at mastery with 89.06%, Yogi scored at threshold with 64.84%, and Nel scored at threshold with 62.50%. The two lower performing students, Yogi and Nel, had 504 plans, though we did not receive specifics about any disability. Furthermore, the quality of sourcing judgments was

poor for students in Honors 2, with all students having made Level 1 judgments. This finding may be attributed to the lack of specificity regarding sourcing in the structure of the original SAC. The quality of contextualization judgments was a Level 4 for Paisley, and a Level 3 for Nel, Yogi, and Isabel. We have provided examples of historical thinking from students in Honors 1 and Honors 2. There was, therefore, some evidence that students' cognitive profiles may have affected their historical thinking skills.

Both teachers agreed that the structure embedded in the SAC and SAC+G was beneficial to lower and middle performing students. For example, Mrs. Few, who used the SAC, said:

> "And I think… that particular group of students that are sort of my middle of the road… they maybe are not going to think really deeply about content or move past just the like basic completion of an activity…. They were benefited, both because the point of the activity was to discuss with peers. So, in order to do the completion piece, they had to talk with someone else. At least they had to listen, if nothing else, to their classmates. But also, they… heard really good quality comments."

Mrs. Fern believed that the collaborative structure of the SAC was important for some students to hear from their higher performing peers. Ms. Adams also believed that lower performing students, especially SWD, benefited from the SAC+G. Ms. Adams said:

> "I do think for the [SWD], so it can be helpful…. I think the tool itself was really helpful for the lower [performing] students, because sometimes they're like, I don't know what to discuss. And I think it broke it down really nicely for them to say like, this is the part you're discussing…. They went box to

box, and I think that was really helpful for the lower to mid [performing] students"

However, she also cautioned that it may be cognitively overwhelming for some students to write responses on the graphic organizer while discussing ideas. She said, "In some ways it's harmful because they're focused so much on the writing, and just getting the questions answered on their paper as opposed to talking to each other. So, it's tricky because, like it would be nice for them to be able to write and discuss. But I just don't think that they're at that level yet." Ms. Adams made this statement with regards to honors students who may lack intrinsic motivation.

Discussion

This study sought to explore students' discussion regarding historical thinking and argumentative reasoning in-depth through a multiple-case study design (Yin, 2018). As states continue to use the C3 Framework to align their standards and create social studies curricula, engaging in inquiry is essential. This requires teachers and students to participate in discourse. Prior research indicates that teachers play an important role in student discussion (De La Paz et al., 2014, 2017; Monte-Sano et al., 2021; Reisman, 2012). In the Co-taught Mod, Ms. Adams played a crucial role in maintaining the procedural flow of the SAC+G, frequently intervening to provide instructions and support. Her facilitation was essential in guiding students through the steps of the SAC+G, and she left room for in-depth argumentation and critical analysis. In Honors 1, Ms. Adams's facilitation in the Honors classroom was less procedural and more focused on fostering critical thinking and deep discussion.

The SAC framework in Honors 2 followed a different structure, with students working in small groups before participating in whole-class discussions in four steps. Mrs. Fern provided more procedural instructions and facilitation compared to Honors 1, presumably reflecting a need for more structured guidance. She prompted students to provide explanations and elaborations, and she facilitated connections to background information such as stagflation and deindustrialization. This is consistent with what researchers have found in their work with teachers (De La Paz et al., 2014, 2017; Monte-Sano, 2008; Reisman et al., 2019). Moreover, both Ms. Adams and Mrs. Fern deftly handled what O'Connor and Snow (2018) deemed the complexities of facilitating discourse.

Student engagement varied in the Co-taught Mod, with some reliance on teacher intervention to sustain discussion. Participation was uneven, with some students dominating the conversation while others remained passive or provided minimal input. While SWD are not as likely to participate in classroom discussion (Bueso, 2022; Louick & Wang, 2021), we found that both SWD, Ford and Bryce, participated more than their non-disabled peers, including in historical thinking utterances. Asterhan and Schwarz's (2016) AFL framework highlights how individual characteristics, including cognitive differences, can enable or inhibit argumentation.

In a debriefing interview two weeks after the discussion, Ms. Adams affirmed that overall participation from SWD and struggling learners increased compared to prior discussions. Moreover, it was gratifying to see that the historical thinking utterances were quite similar, as shown in Figure 3.5. In Honors 1, engagement was uniformly high, with students actively participating and contributing meaningful

insights without teacher prompting. The SAC+G allowed students to explore the CHQ thoroughly, though it may have constrained these higher-level students. Engagement in Honors 2 was also high, with students demonstrating strong participation and interaction, supported by the teacher's facilitation. Students in Honors 2 participated more in historical thinking than students in Honors 1, as shown in Figure 3.6.

We also explored the quality of students' argumentation and historical thinking, specifically looking at their sourcing skills as measured by Nokes (2017) and their contextualization skills as measured by Monte-Sano & De La Paz (2012). We used task design and process support as outlined in the AFL framework (2016). Specifically, we problematized content through the CHQ and document set (see Appendix A), ensuring a space for argumentation. Moreover, we provided clear discourse instructions and teacher scaffolding that allowed for argumentation.

Students in all three cases engaged in disciplinary argumentation Moje (2008). Argumentation in the Co-taught Mod was generally competent, with students providing basic arguments. Three of the four students produced rebuttals, demonstrating an ability to weigh arguments. In addition, the quality of historical analysis was limited, reflecting the students' varied academic abilities. Students made partial sourcing and contextualization judgments but made several judgments that showed high engagement levels. However, Ford made a substantively better corroboration judgment compared to his peer in Honors 1, Rex.

In Honors 1, students demonstrated sophisticated argumentation and historical thinking, incorporating sourcing, contextualization, and quality of evidence into their

discussions. Their ability to construct and defend nuanced arguments was markedly superior. Argumentation in Honors 2 was similarly robust, with students engaging in high-level deliberation and argumentation. Mrs. Fern's prompts for sourcing and contextualization enhanced the depth of discussion, though her students did not demonstrate high-quality sourcing skills. This result was surprising given the high-quality argumentation and other historical thinking skills (e.g., contextualization, corroboration). This demonstrates the importance of explicit instruction in historical thinking skills because sourcing does not come natural to students, as supported by prior research (De La Paz et al., 2014, 2017; Monte-Sano et al., 2014, 2017, 2021).

We also explored whether students engaged in deliberative or disputative argumentation (Asterhan & Schwarz, 2016; Felton et al., 2009, 2015). It is important to look at how students held their discussion to determine whether they engaged in deliberative or disputative argumentation. We see the students never engaged in disputative argumentation or had a combative discussion in any of the Mods. Rather, they understood each other's perspectives and remained respectful throughout, consistent with characteristics of deliberative argumentation (Asterhan & Schwarz, 2016; Felton et al., 2009, 2015). Students in the Co-taught Mod and Honors 1 also engaged in consensual co-construction, which is distinct from argumentation because students seek to explain or clarify more than form arguments as defined by Asterhan and Schwarz (2009a; 2016). Though the quality of argumentation differed between cases, all students worked together to understand and construct an argument, which leads to greater learning (Felton et al., 2009, 2015; Johnson & Johnson, 1979, 1985, 1988; Johnson et al., 1996). The graphic organizers serve as a pre-writing tool. We

think that by deliberating, students can better develop their written argumentation (Graff, 2003).

Finally, there were certainly differences among the students with varying academic and cognitive profiles. This pertains to the "individual characteristics" category within Asterhan and Schwarz's (2016) inhibitors and enablers of argumentation. For example, though Ms. Adams' students in Co-taught and Honors 1 participated at a similar rate, the quality of their discussion differed by their academic and cognitive ability. Ford and Bryce, both SWD, had lower-level quality historical judgments. They also had the lowest scores on the district History assessment. Moreover, students who need more support benefit from scaffolding and can use time in small groups to prepare for participation in whole class discussions (Monte-Sano et al., 2021). Specifically, Ford participated extensively in the whole class discussion. Perhaps with more practice, these students will improve their oral argumentation and historical thinking.

Students in Honors 1 and Honors 2 had the highest scores on the district History assessment and generally had higher quality historical judgments. The only exception was Honors 2, which we attributed to a difference in emphasis on sourcing. Ms. Adams repeatedly emphasized the importance of properly sourcing the documents, as outlined in the SAC+G. The proper sourcing of the documents was not explicitly outlined in the SAC used by students in Honors 2.

Limitations

Due to the case methodology, causal outcomes cannot be made. The findings reported here are based on qualitative analyses of 12 students from two teachers at one school. However, this in-depth case study offers a rich description of how knowledge is constructed and negotiated in argumentative discourses. It also provides insight into how argumentation affects learning historical thinking skills. Second, there was no comparison group to other types of structured discourse (e.g., Socratic seminar) used in social studies. Finally, as this cross-case analysis lacked formal controls, exogenous factors could have influenced the reported results.

Implications for Practice

We believe that by modifying inhibitors and enablers of argumentation, (e.g., task design and process support), a common discourse structure can be made into an explicit routine and used at scale for SWD and other struggling learners. The two main principles found here are discussion and explicit instruction, which are HLPs. Students in Ms. Adams' Co-taught Mod appeared to benefit from the step-by-step structure of the discussion, which in turn allowed all students to participate. We believe that starting students in pairs with the aim to deliberate is helpful for their comfort and confidence to speak. Ending the discussion with a whole class debrief allows students to hear new ideas, and it allows the teacher to clarify ideas or prompt students to elaborate on their thinking. The SAC+G may be useful for students in Honors classes, though teachers should monitor whether the structure is as flexible as some higher-level students may require. The use of turn and talk (Swanson et al.,

2019) is particularly useful when students are in pairs. Finally, the graphic organizers as pre-writing planning tools allow students to formulate their arguments before writing their essay, which is an effective process support to improve writing (De La Paz et al., 2014, 2017; Monte-Sano et al., 2017; Wissinger & De La Paz, 2016).

Recommendations for Future Research

While we provide an in-depth exploration of student discourse, future research could follow our SAC+G and coding to use discourse analysis to provide an even richer description of how students engaged in argumentation, historical thinking, or both. Future research could also focus on quantitative findings stemming from the findings in this study. For example, discourse could be analyzed quantitatively to examine whether engaging in deliberative argumentation improves student learning as measured by historical content knowledge. In addition, future research could quantitatively examine whether the use of a SAC+G and video modeling by school-aged students improves the participation and quality of historical thinking among SWD and other struggling learners. Future research could also examine whether a SAC+G has an effect on student oral argumentation, written argumentation, or both. Future research could also examine how historical thinking translates from discussion to writing, looking specifically at sourcing, contextualization, corroboration, and rebuttal. This can also be extended to examine how argumentation skills translate from oral communication to written communication. This study was conducted at a majority White school. Future research should examine whether a SAC+G as outlined

in this study also improves the performance of culturally and linguistically diverse students.

Chapter 4: Improving Students with Disabilities and Other Struggling Learners' Argumentative Discussions with Guided Support for Structured Academic Controversies

Bryce: All right, so I had document B about minority applications, so basically this talks about the construction of the Baltimore Aquarium and how a bunch of new job opportunities will be opened and six minority applications from the Afro-American. So, it's pro African American in this time, which is huge, so they just want more job opportunities for the less fortunate. Then again, all that money that they're building an aquarium for entertaining could also be putting back in the community, which they are. So, I think they were pro, for it.
Julius: On the author's perspective, it's from Ron Howell, a reporter for the Baltimore Sun, and his answer is no to whether or not he supports the building of the Harborplace in Baltimore because this quote from the source, he said that this project will have an effect on the future racial socioeconomic makeup of several Baltimore neighborhoods.
Bryce: Well, document A was about people buying out people's homes and then modernizing homes just for cheap profit.
Bryce: Yeah. What do you think about construction of Harborplace in Baltimore, do you think it improves life?
Julius: No.
Bryce: Why not?
Julius: I feel like money could be used towards homelessness, poverty.
Bryce: Yeah, that's true. The money could be used towards that, but it also opens a more public area where small businesses could be opened up and provide and provide more opportunities for less fortunate people like homeless people.
Julius: Yeah, I didn't think about it that way.

Many children enjoy the social aspect of talking with friends at school. Indeed, discussion is a social act that occurs naturally when people interact. However, it is a different matter altogether when students must meaningfully engage in disciplinary classroom discussions, and this topic has been of longstanding interest to teachers. According to the College, Career, and Civic Life (C3) Framework for Social Studies, students are expected to engage in discussion and argumentation through inquiry to cultivate disciplinary practices and ways of thinking (NCSS, 2013). As such, discussion is a complex instructional practice that can pose significant challenges for teachers (Mercer, 2004). The excerpt above comes from two students in a ninth grade US History co-taught classroom. Bryce has an IEP and Julius has a 504 plan, but they are both able to have substantive disciplinary discussions.

We believe it is vital that all students participate in discussion and not just the few who tend to dominate the conversation. Research demonstrates that some students have little prior experience with argumentation and those with limited background knowledge are likely to need different supports than peers who understand how to engage in argumentation. For example, students with disabilities participate at lower levels than their fellow non-disabled peers (Louick & Wang, 2021). It is also true that the type of argumentation matters, as students in disputative conditions have been found less likely to share incomplete understandings with dialogue partners than those in deliberative conditions (Asterhan & Babichenko, 2015). Prior research has also demonstrated that when students engage in deliberative argumentation rather than other argumentative dialogue (e.g., disputative

argumentation), they increase participation, produce better writing, and learn more about topics (Asterhan & Schwarz, 2016; Felton et al., 2009).

Social studies inquiry provides engaging opportunities for discussion as students explore complex and debatable questions. Students analyze diverse sources to consider multiple perspectives, weigh often contradictory evidence, and develop well-supported conclusions about unresolved social and historical issues (Grant et al., 2017). Throughout this process, students communicate and learn from each other as they work toward conclusions collectively (Parker, 2006).

Furthermore, to hold effective discussions, teachers should give students sufficient time to exchange ideas and reasoning about texts with their peers and teachers (Kamil et al., 2008). It is recommended that teachers: (a) carefully and purposefully prepare for the discussion; (b) ask follow-up questions to sustain and extend the discussion as needed; (c) provide a task or format that students can easily follow while discussing texts in small groups; and (d) implement a specific discussion protocol that consistently provides specific guidelines to clarify procedures. In this way, student and teacher roles are defined, and both can effectively engage in classroom discussion despite its challenging nature.

Many states have aligned their social studies standards with the C3 Framework. As a result, social studies teachers need the knowledge and tools to facilitate discussions through inquiry. This approach is particularly effective in history classrooms, where historical inquiry allows students to engage in meaningful discussion by investigating complex and disputable questions. The effects of the war on drugs is one example. After digesting the question, students analyze various

sources with contradictory evidence and consider multiple perspectives to form conclusions about historical controversies.

Structured Academic Controversy

Historical arguments are a response to a historical question, or controversy, and an appropriate historical question provides students with more than one possible response. Historical arguments can be more engaging when students have some sort of interest in the event or if the event evokes emotion, so teachers may want to consider using current events to introduce a structured academic controversy (SAC). Consider the following potentially contentious question: "How much collateral damage is acceptable in war?" For specific examples, we can reference the atomic bombs in Japan during WWII. To history buffs, this may be compelling and intriguing. To adolescents today, it is ancient history (it has been almost 80 years). We need buy-in from the students to have the discussion. Therefore, a modern example like the current Israeli-Gaza war, could work better. While some teachers may wince at the thought of teaching such a hot-button issue, that is the point of SAC. Teachers can use SAC to engage students in argumentative discussion whereby students collaborate to make well-formed arguments about complex and contentious issues after examining at least two perspectives (Johnson & Johnson, 1988).

The SAC is a great classroom practice that is relatively simple to implement, so that even novice teachers can use it. Moreover, a SAC can also be used in civics classes with controversial topics. In a SAC, the teacher specifies the objectives, prepares for the exercise, explains and organizes the task and controversy, monitors student performance and intervenes as necessary, and evaluates and processes the

lesson. Social studies teachers should be well-equipped to conduct discussions involving academic controversies to help their students improve academically. The SAC also has benefits beyond the classroom, as it facilitates the social aspect of learning through discussion. By using a SAC, students can increase achievement, self-esteem, problem solving, and decision making.

Historical Investigation

Students need explicit instruction in planning, and composing, alongside explicit instruction in historical reading skills, and a cognitive apprenticeship approach to instruction is useful to teach these skills. Wineburg (1991a) specified three heuristics that historians use and that history teachers can readily adopt when reading historical documents: sourcing, contextualization, and corroboration. We developed a 9th grade US History curriculum with six historical inquiry lessons in which students are taught historical thinking skills as well as specific historical reading and writing skills (De La Paz et al., 2023). We used a Central Historical Questions (CHQ) to guide this historical investigation, "Will the construction of Harborplace in Baltimore improve Life in the city?".

To engage students in discussion, we initially provided teachers with a SAC graphic organizer (see Figure 4.1) meant for discussion in small groups. Throughout the SAC, the teacher debriefs with the whole class for a quick discussion to ensure understanding. After discussing, students write down their final stance on the Central Historical Question. The teacher's main tasks during the group discussions are to ensure students understand the historical question and that they are participating in the discussion with their peers. However, we noticed through observing classrooms in

our studies that very few students were engaged in discussion. This led us to rethink how to use the SAC.

Guided Structured Academic Controversy

We know from prior research that students with disabilities do not participate as much as their non-disabled peers (Bueso, 2022). The revised, guided SAC we developed is intended for teachers seeking additional supports for disciplinary discussions with these and other struggling students. The goal of the supportive SAC is to help students engage in productive, authentic, and deliberative argumentative discourse. Deliberative argumentation encourages students to collaborate on arguments and try to build consensus. As such, the revised SAC provides more guidance and support than the original SAC, as students move from discussion in pairs to small groups to whole class.

We call the result SAC+G because of the enhanced guidance for both teachers and students. Moreover, the SAC+G includes more questions to drive evaluations of judgments related to historical thinking, including sourcing, contextualization, and corroboration. These questions are integrated into a routine for teachers and students, modeled on turn and talk (Swanson et al., 2009). In addition, this routine breaks down step-by-step how students comprehend and evaluate sources, which we found to be helpful in eliciting more abstract reasoning (see Chapter 3). The accompanying graphic organizer provides a structure for writing ideas and evidence. It includes sentence starters to help students plan how to write the evaluation of sources in their essays. Moreover, the SAC+G contains six step that will be described in detail in the

next section: (a) warm up; (b) identify perspectives, (c) initial hypothesis; (d) find evidence; (e) evaluate each document; (f) whole class discussion.

Preparation

We believe it is important to provide teachers with detailed instructions on how to properly use the SAC+G to maximize its success in the classroom. However, our instructions are recommendations. Teachers have flexibility to adapt the SAC+G as they see fit for their students. Our recommendations are based on research that has shown promising results, and we hope teachers adopt the SAC+G for their classroom discussions. Figure 3.2 shows the SAC+G graphic organizer, which can be completed digitally or on paper.

Warm Up

The warm up consists of questions meant to engage students in recalling the central historical question and recognizing that there are differing opinions about it. Teachers should take this opportunity to remind students that history is not fixed, and the purpose of the question is to have students think like historians. Students have two options for the warm up, prompting both a concrete and an interpretive response. Teachers choose one of the questions and then model the question they did not ask students to complete. Students have three to five minutes to complete the warm up, and when time is up, teachers should direct students to move on to the first step.

Step 1: Identify Perspectives

In this step, teachers show students the first two rows in the graphic organizer that contain Questions 1 and 2. These two questions ask students to determine the perspective of the document, and it is vital to remind students to write the name of the

documents (e.g., “Urban Renewal’s Refugees”) as opposed to referring to the document number or letter (e.g., Document 1, Document A). Teachers then set a timer for two to three minutes while students work in pairs. One partner should focus on the first source and the other partner should focus on the second source. Then, teachers ask students to share their responses with their partner. This should take approximately five minutes, though some students may finish in less time while others may not finish in the allocated time. The teacher should circulate the classroom to comment, elaborate, and give constructive feedback.

Step 2: Initial Hypothesis

In this step, teachers show students the third row in the graphic organizer and instruct students to answer Question 3. It is important that teachers explain this is students’ initial hypothesis and that their response is not fixed and may change based on the subsequent discussions. Students should be prompted to provide a reason for their claim. As teachers circulate the classroom, they should ensure that students are writing down their response after discussing their thoughts. This activity should take two minutes, and then teachers should promptly move to the next step.

Step 3: Find Evidence

In this step, teachers show students the fourth row and explain that Question 4 is meant to provide evidence that supports students’ responses to the historical question in Question 3. Teachers should encourage students to talk with each other to critique the evidence and clarify that the purpose of this activity is to push students to explain their reasoning. Students should be given five minutes to complete the activity. The teacher should continue to circulate the classroom to comment and

elaborate or provide constructive feedback in case students provide inaccurate evidence to a claim. More importantly, teachers should ensure other students weigh in and provide more accurate evidence in this case. We acknowledge this can be difficult, but that is why we encourage circulating the classroom to ensure student engagement. This activity should take two minutes, and then teachers should promptly move to the next step.

Steps 4 and 5 Evaluate Each Document

In step 4, teachers show students the fifth and sixth rows and explain that for Questions 5 and 6, teachers will assign pairs to work on the first or the second document. Another option is to allow pairs who are working at a table of four students to decide who is evaluating the first document and who is evaluating the second document. Students have five minutes to complete this activity. The pair first chooses one historical judgment question from the list of critical questions in the SAC+G graphic organizer. If students are comfortable with sourcing, encourage them to focus on contextualization or the important details. Teachers should also encourage students to engage in corroboration using background information or another document. After each pair evaluates a document orally, they should proceed to write down their answers. This activity should take five minutes.

In step 5, teachers then ask pairs to share evaluations with another pair and encourage students to critique and evaluate each other's ideas. The discussion should be ongoing to make the best possible argument. Teachers provide feedback as to how evaluations can be improved, and other students may join in. This should take five minutes. When time is up, teachers ask pairs to share their ideas and comment,

elaborate, provide constructive feedback, and try to arrive at a consensus. This should take two minutes. Here, it is intended and expected that students engage in deliberative argumentation and collaborate to arrive at a consensus. Teachers should advise students to support the side that has the most evidence, regardless of personal opinion.

Step 6: Whole Class Discussion

At the end of the lesson, teachers bring the class together for a final whole class discussion. The purpose of this step is to decide which evidence is strongest, regardless of students' initial hypothesis, and to share evaluations of documents. Teachers should ask students if the evidence is stronger for one perspective than another perspective. This should take between five to 10 minutes, depending on how the discussion flows and student engagement. Then, teachers solicit answers and tell students to decide for themselves which perspective is strongest. This is the final hypothesis, and teachers should ask students to reflect on whether their final hypothesis differs from their initial hypothesis and why or why not. Teachers provide students five minutes to complete this step.

Planning to Write

After the discussion, teachers show students the seventh row and tell students to complete each part to plan their essays. Teachers should review with students that they should select ideas using a star or checkmark and number quotes and ideas in the order they will use for their essay. If time permits, teachers can solicit student plans. Teachers should circulate among students as needed to ensure they complete this step, as it will facilitate their essay writing. Teachers encourage students to use the

sentence starters to evaluate the sources if they notice students struggling. This step should take five minutes.

Student Discussion in Action

We provide excerpts from students in two of Ms. Adams' classes, one co-taught and one honors. This serves as a contrast to show how the SAC+G impacted discussion among students with varying academic profiles. The first group of students was comprised of Ford, Bryce, Kristen, and Julius, and they were in Ms. Adams' co-taught class. Their historical thinking skills were lower than their peers in the same teacher's honors class, as measured by the district quarterly History assessment administered at the beginning of the school year, and teacher feedback. Ford and Bryce had Individual Education Programs (IEP) for a LD and other health impairment, respectively, and Julius had a 504 Plan. Kristen was very shy and mostly spoke when prompted by the teacher or a peer.

The discussion below occurred when two pairs merged to discuss students' evaluation of a historical document. Students were encouraged to make historical judgments based on sourcing, contextualization, corroboration, or the quality of facts.

> *Ford*: "We did document A. What was yours about?"
> *Julius*: "So, I said I think the source might be a little one-sided because it says, 'Baltimore Aquarium seeks minority applications.' And this is by the African American newspaper, so I think it's just over viewing the positives for like the minority at the time."
> *Ms. Adams*: "So maybe there is some underlying bias there. Maybe it's one-sided. What did you all say?"

Ford: “This author might have known more about the possible outcomes of what could've happened to the city around the time of the – trying to build it and whatnot.”

Ms. Adams: “Why do we say that?”

Ford: “Because the amount of time that the person lives there and whatnot.”

Ms. Adams: “Okay, and do you get that from the headnote? Where do we know that?”

Ford: “No, it's a reporter that's been working there for a while and probably some of their colleagues and whatnot have possibly lived there their whole life and whatnot.”

Ms. Adams: “Does it say that?”

Ford: “No.”

Ms. Adams: “That he lived there? So, we can't make that judgment, because we don't know that. So, we can't make a judgment that we don't know about. We have to judge what we have in our information. So, what do we know about him? We know he's a reporter. So, what do you think?”

Ford: “Yeah he was part of the news anchor and whatnot.”

Ms. Adams: “So what type of judgment could be make based off of that?”

Kristen: “That they know more because they have more resources.”

Ms. Adams (21:10): “So maybe they have more research, maybe more access too. What do you think about your source? Do you trust it? Do you not trust it?”

Julius: “I'm a little in between.”

Ms. Adams: “Okay, why?”

Julius: “Because, well, I mean, the author, he's on site and his quotations and stuff. On the other hand, he's just putting the good stuff, not the bad stuff, like there is always downsides to things.”

The students and teacher engaged in historical thinking through sourcing. Ford did not specify the author’s name, but he did know the author is a reporter. He

then made an inference that the author lived in Baltimore, and rather than immediately correct him, Ms. Adams prompted Ford to explain his thinking. Ultimately, Ford realized that there was no information regarding where the author lived, and Ms. Adams pressed forward. Kristen's sole contribution was that the author had more access to resources and was therefore a reliable source. Julius understood the author had aspects that made him reliable, but he also noticed that the author could be biased. This excerpt illustrates a historical thinking discussion with high engagement by two students with disabilities, facilitated by the SAC+G. This is promising because students with disabilities and other struggling learners tend to participate at lower levels than their peers without learning challenges. In addition, we share Ford's completed graphic organizer as an example (see Figure 4.2).

The second group of students was comprised of Heather, Rex, Taylor, and Jake, and they were in Ms. Adams' honors class. Their historical thinking skills were higher than their peers in Ms. Adams' co-taught class, as measured by the district quarterly History assessment and Ms. Adams' feedback. Ms. Adams did not participate in this discussion, as she had 31 students in her honors class and was unable to reach all groups in the allotted time. Students were encouraged to make historical judgments based on sourcing, contextualization, corroboration, or the quality of facts. Below is a discussion of students' evaluation of a historical document during the small group discussion where the two pairs merged to discuss.

> *Jake*: "What did y'all discover from your document?"
> *Heather*: "We talked mainly about the reliability of the document two. So, Ryan pointed out that there might be some bias as it was primarily, or it was a African American owned newspaper, and due to displacement of a lot of

Black families, it could have been biased to events. However, it was and still is a very notable and credible newspaper."

Jake: "Yeah, okay."

Heather: "It's there."

Jake: "You want to go over what we've figured out?"

Taylor: "So, with ours, we read over it and read some of the footnotes. And we saw that the guy who wrote this article works for the Baltimore Sun, which is a magazine. And also, has lived in Baltimore for a while. So, we were saying that the source is reliable, because there's nothing really for him to benefit of if he is to lie so they can – very well. And also, since he knows the Baltimore area a lot, that he would know the most information and get the actual real story."

Heather: "Yeah. [Jake], what do you think about that?"

Jake (19:49): "I think it… You said all the words I was going to say. But just because the Baltimore Sun, I don't know why he would lie about stuff that's happening in towns."

Heather: "Because it was very factual."

Jake: "Yeah. Yeah."

Taylor: "Because there's no real reason to lie."

Jake: "Because if you lie, you're caught."

Heather: "And the Baltimore Sun is a very credible newspaper even now."

In this excerpt, three of the four students participated in the conversation. Although Rex did not participate, in the prior step, he and Heather discussed the sourcing and reliability of their document. This excerpt demonstrates higher historical thinking skills around the sourcing and reliability of the documents. The students critiqued the documents based on sourcing information but provided more detail than the students in the prior excerpt. Finally, Ford and Taylor made a similar inference of the author's knowledge of Baltimore due to him residing in Baltimore, which Ms.

Adams said could not be supported by information in the document. Taylor, however, proceeded to add to his judgment that (a) lying is not beneficial to the author, and (b) the author gets accurate information about Baltimore because he knows the city from covering it as a reporter.

Teacher Reflection

Upon reflecting on the implementation of the SAC+G in her co-taught and honors classrooms, Ms. Adams stated:

> "I liked it for myself…. I thought it was helpful… it gave a good model for them to kind of know what they were writing. I think the students also did really well…. I do think for the [students with disabilities], it can be helpful. "

Ms. Adams concluded that some students, specifically students with disabilities, benefitted from a more structured discussion. She also cautioned that some students, specifically in honors classes, may worry more about writing than discussing or may be unable to handle completing both. Teachers should be aware of this and ensure that their scaffolding and structure are not cognitively overwhelming, and that appropriate time is allotted for discussion and for writing. As for the quality of the discussion, Ms. Adams stated:

> "I think this one – the discussion was really good. I would say like compared to like regular lesson discussions…. They did really nice, especially in small groups. I really thought the small groups those were the ones that were like doing it, were really good. What I heard sounded really great, and it's not like it was exactly what I wanted to hear from – but not everyone. Some of them were not the best, but for the most part, like those who I was worried about, and maybe didn't know how they would do, I thought they did really well. I was really quite pleased with a lot of those discussions."

Ms. Adams reported that with the SAC+G, the quality of the discourse was better than other discussions throughout the year. Thus, it appears that the SAC+G was helpful for students, especially for struggling learners.

Conclusion

The idea that students lack motivation and will not engage in discussion is pervasive and discourages teachers from having discussions in the classroom, especially with students with disabilities and other struggling learners. We developed the SAC+G to engage and support all students, regardless of academic and linguistic ability. We found that through its use, students of all academic profiles can participate in an engaging and thoughtful discussion. Moreover, when students are asked to collaborate in grappling with a historical controversy, we encourage deliberative argumentation. This is especially important because prior research has found that students who engaged in deliberative discussions had better learning outcomes. Finally, as students learn through inquiry-based instruction, they must be prepared to communicate what they know orally and in writing. For this reason, we specifically crafted our graphic organizer to prepare students for a written final product and to equip teachers with a valuable tool for their toolkit. Through their discussion, students find evidence to support their claims, evaluate authors' perspectives and credibility, and contextualize and corroborate documents as they collaboratively learn. After exchanging ideas and writing down what they discussed with their peers and teachers, students are better equipped to write argumentatively. Creating a structured and supportive environment for discussion is a tall order, but the results are worth it.

Chapter 5: Conclusion

The use of argumentative discourse in education dates to the great philosophers of Ancient Greece. Today, argumentative discourse continues to be promoted in primary and secondary classrooms through new standards and frameworks that encourage inquiry, such as the College, Career, and Civic Life Framework (C3 Framework) in social studies, which encourages teachers to engage students in inquiry that promotes thoughtful discourse about debatable and multifaceted questions (NCSS, 2013). Prior research indicates that deliberative discourse improves student participation and learning outcomes (Asterhan & Schwarz, 2016; Felton et al., 2009). When students are faced with complex academic problems, they can achieve better results when they collaborate, and teachers can use a structured academic controversy (SAC) to guide students through effective discourse (Johnson & Johnson, 1988).

First, I conducted a research synthesis to investigate the impact of disputative and deliberative argumentation discourse on students' learning outcomes and to understand the instructional components utilized by teachers in primary and secondary science and social studies classrooms (Chapter 2). The findings of this synthesis identified important instructional components related to discourse that can yield positive results for student learning. Based on my synthesis, I conducted a descriptive multiple-case study (Chapter 3). Specifically, I explored three cases through a cross-case analysis to provide insightful descriptions of the relationship between student discourse and discourse structure, as well as the quality of the

discourse through a revised, guided SAC. I collected and analyzed data from two teachers in three classes, with each class comprising a case. I wrote a practitioner manuscript after concluding the study that outlines how teachers can use a guided SAC to increase the participation and quality of all students' argumentative discourse in history classrooms.

I will summarize the findings and limitations of my research synthesis and multiple-case study in this chapter. In addition, I will describe how the practical implications of the synthesis and multiple-case study helped inform the practitioner manuscript. Furthermore, I will review the guidance offered in this manuscript. Finally, I will conclude by exploring future research directions based on the synthesis and multiple-case study.

Research Synthesis

As more academic disciplines approach instruction through inquiry-based learning, it is essential that educators understand how to hold effective disciplinary argumentative discourse with their students. Following Asterhan & Schwarz's (2016) Argumentation for Learning (AFL) conceptual framework, deliberative and disputative argumentation are the two main types of argumentative discourse. Thus, my purpose in conducting the research synthesis was to determine the impact of disputative and deliberative argumentation and inhibitors and enablers of argumentation (e.g., task design, process support) that aid students' learning outcomes in primary and secondary science and social studies classrooms. Specifically, I identified four learning outcomes that included disciplinary

knowledge, disciplinary writing, argumentation structure, and holistic writing. Thirty studies from regions across the world (e.g., North America, South America, Europe) qualified for inclusion in my synthesis.

The synthesized findings of these studies revealed that that students benefited from discourse when they were in more than one instructional grouping (e.g., small groups, whole class), demonstrating improved learning outcomes. Students also benefitted from explicit instruction, teacher modeling, and graphic organizers, and while technology use is promising, it requires further research. The findings further demonstrated that deliberative discourse is more effective than disputative discourse in improving students' disciplinary knowledge and disciplinary writing, including for students with disabilities (SWD) and multilinguals (ML).

In addition, students' argumentation improved, especially when learning was measured through writing – and there were no differences in findings between academic domains. One limitation of the research synthesis was that various studies did not clarify whether the explicit instruction and modeling related to holding a discussion or for another purpose (e.g., content knowledge). Moreover, the description of the discourse was often not explored in-depth. Thus, I conducted my study to address these gaps and extend the findings of my research synthesis through a qualitative study.

Multiple-Case Study

I conducted a descriptive multiple-case study in the Spring of 2023 using cross-case analysis to explore how argumentation inhibitors and enablers moderate dialogue characteristics and learning outcomes (Asterhan & Schwarz, 2016) and to

richly describe the quality of the discourse in ninth-grade US History culturally and academically diverse students. Specifically, I analyzed the students' argumentation and historical thinking skills. The first case involved four students and their teacher in a co-taught class, the second case involved the same teacher and four other students in an honors class, and the third case involved four students and a different teacher in an honors class. The first two cases (Co-taught and Honors 1) used a revised, guided SAC termed SAC+G, and the third case (Honors 2) used an original SAC. I compared the co-taught and Honors 1 classes to explore how students with differing academic and cognitive profiles used a modified SAC during a historical investigation (HI). Similarly, I compared the Honors 1 and Honors 2 classes to explore how students with similar academic and cognitive profiles used different discourse structures.

The SAC+G included three instructional groupings, comprised six steps, and had an accompanying graphic organizer. In steps 1 through 4, students in pairs discussed the meaning of the central historical question (CHQ), provided an answer to the CHQ (i.e., the claim), found supporting textual evidence, and evaluated at least one source using historical thinking skills (e.g., sourcing, contextualization). In step 5, student pairs merged with a different pair to share and expand upon their historical evaluations. In step 6, the teacher conducted a whole class discussion to debrief what students discussed. I collected data including teacher and student audio recordings, permanent products (written graphic organizers, digital work), semi-structured interviews with teachers after data collection, and field notes. I utilized software to transcribe all audio recordings. Sixteen codes emerged after analyzing student and teacher utterances. The codes were organized into two main types, argumentative and

historical thinking. I then compared argumentative and historical thinking utterances for quality across the co-taught and Honors 1 classes and Honors 1 and Honors 2 classes. In addition, I analyzed the teacher interviews and identified the following themes: (a) structure increased struggling learners' engagement in discussion; (b) the quality of discussion was overall higher quality compared to other lessons; and (c) writing and discussing may be overwhelming for some students.

Findings revealed that both teachers played vital roles in their classroom discussions. The teacher played an important role in the co-taught Mod by maintaining the procedural flow of the SAC+G and frequently intervening to provide instructions and support. The teacher likewise played a crucial role in facilitating the discourse in Honors 1, but her facilitation focused less on procedural flow and more on promoting critical thinking and deep discussion. In both cases, she encouraged a higher level of academic discourse by encouraging students to elaborate on their arguments, grapple with different perspectives, and critically evaluate historical evidence. In Honors 2, the teacher provided more procedural instructions and facilitation compared to Honors 1, suggesting a need for more structured guidance.

Findings also revealed that participation in the co-taught class was similar to honors classes, though it was more uneven among the four students. Surprisingly, the two SWD participated more than their peers without disabilities in the co-taught class, suggesting such purposeful structure can be beneficial for SWD. Moreover, Ms. Adams added that she believed the SAC+G particularly helped struggling learners, especially SWD.

Students demonstrated varied historical analysis skills. In the co-taught class, while students attempted to use historical thinking skills such as sourcing and contextualization, they demonstrated limited historical thinking skills compared to students in Honors 1. However, one student in the co-taught made a more elaborate corroboration judgment than a student in Honors 1. Students in Honors 1 showed high-level argumentation and historical thinking, as they incorporated sourcing, contextualization, and quality of evidence in their discussions. In Honors 2, students demonstrated advanced argumentation and historical thinking skills, specifically in contextualization. However, these students did not source documents properly, perhaps reflecting a need for the teacher to prompt students for sourcing information. Both teachers said they heard high level discussions and were pleased with much of what teachers discussed. In sum, this demonstrates the importance of explicit instruction in historical thinking skills because they do not come natural to students.

Furthermore, findings revealed that all students engaged in deliberative argumentation, though students in the co-taught class also engaged in consensual co-construction. Asterhan & Schwarz (2016) state that while the talk may resemble each other, consensual co-construction is considered explanation rather than argumentation because the purpose of each is different. Moreover, none of the students engaged in disputative argumentation, and they remained respectful and constructive throughout. This was not surprising as the SAC+G was developed for students to engage in deliberative argumentation rather than disputative argumentation.

Finally, the differences among students with various academic and cognitive profiles were evident. The quality of students' discussions varied according to their

academic and cognitive abilities. The students in the co-taught class scored lower on a district History assessment and exhibited lower-quality historical judgments compared to students in Honors 1, apart from one student's corroboration judgment. Students in honors classes achieved higher scores on the district assessment and showed better historical thinking, apart from the sourcing judgments in Honors 2. Ms. Adams said that while participation and quality improved, there was still room for improvement for her struggling learners.

Practical Implications

The multiple-case study of the current dissertation has several practical implications. First, the study demonstrated that provided with explicit instruction, proper structure, and teacher facilitation, all students can engage in argumentative discussion regarding a historical controversy. Though SWD typically do not participate as much as their non-disabled peers (Bueso, 2022; Louick & Wang, 2022), their high participation level in this study is encouraging. Thus, both special educators and general educators should consider purposefully structuring and guiding their discussion activities, particularly the use of group formation (e.g., pairs, small groups, whole class), to involve all students. Teachers should be flexible in their group formation, as some students may not require as much support as other students, and further support may in fact impede discussion for students with higher-level academic profiles. This may require professional development on how to structure classroom discussions to effectively engage diverse learners.

Second, historical controversies are naturally argumentative. Prior research shows that deliberative argumentation yields better learning outcomes than

disputative argumentation (Asterhan & Schwarz, 2016; Felton et al., 2009). Teachers should consider developing a collaborative and respectful environment that lends itself to deliberation rather than dispute. Modeling how to hold deliberative discussions could be beneficial, and avoiding debates may be helpful when resolving historical controversies. Moreover, teachers should consider video modeling, as presented in this study, to ensure the modeling demonstrates meta-cognitive thinking, is consistent, and to facilitate reteaching.

I wrote a practitioner manuscript that offers guidance to social studies teachers on how to effectively hold a more supportive and guided SAC (SAC+G) to foster higher levels of participation and quality in argumentative discourse for all students based on these practical implications and the practical implications of my research synthesis. In the manuscript, I explain that teachers can plan for discussion in historical controversies by thoughtfully planning a structure with supports and providing explicit instruction to students. Specifically, teachers were provided with a graphic organizer and instructions on how to effectively hold the SAC+G, including suggested pacing.

After providing detailed instructions on how to prepare for the SAC+G and use an accompanying graphic organizer, I provide student discourse excerpts, a student’s completed graphic organizer, and a participating teacher’s reflection. These are meant to encourage teachers to hold argumentative discussions in their classrooms through careful and strategic planning. Given the complexities of discourse in classrooms, it is important that teachers have procedural and disciplinary information,

including strategies and tools to incorporate in their planning and delivering of instruction involving discourse.

Future Directions

The current dissertation has essential findings and implications that can be addressed in future research. For example, Chapter 3 offers a rich description of how teachers and students engage in historical thinking with the SAC+G. Taking the qualitative results from the study, two quantitative research questions that could be addressed in future research include: (a) To what extent does engaging in deliberative argumentation during the SAC+G impact students' historical content knowledge? and (b) What is the effect of using the SAC+G and video modeling on the participation and historical thinking quality of SWD and other struggling learners? Answers to these questions would increase our understanding of instructional approaches that address how structuring discourse could benefit students' understanding of disciplinary knowledge in history.

The practical implications of Chapter 3 suggest it is important to strategically structure discourse to elicit higher participation and historical thinking from students. Given that the study took place in a majority White high school, these results could be expanded to schools whose populations are more culturally and linguistically diverse. Two research questions that could be addressed in future research include: How does strategically structured discourse impact the participation and historical thinking skills of students who are culturally and linguistically diverse? and (b) What is the effect of strategically structured discourse on the level of participation and the quality of historical thinking among students from culturally and linguistically diverse

backgrounds? This research could extend the discourse literature and identify specific strategies that educators could use with culturally and linguistically diverse learners.

Finally, the SAC+G graphic organizer was designed to not only encourage active and high-quality participation, but as a preparation for a culminating written product. The study's HI called for an answer to the CHQ by stating a claim, providing reasons supported by textual evidence, explaining how the evidence connects to the claim, evaluating the historical sources, and providing a rebuttal within a five-paragraph argumentative essay. Thus, two possible research questions that could be addressed in future research include: (a) How do elements of historical thinking, such as sourcing, contextualization, and corroboration, translate from classroom discussion to individual writing? and (b) How do argumentative skills developed through oral communication in discussions transfer to written communication in argumentative writing? Examining these questions through a mixed-methods study could result in strong quantitative contributions that further connect discourse and writing, as well as new qualitative descriptions and insights.

Conclusion

The current dissertation explored argumentative discourse and historical thinking in ninth-grade US History teachers and students from various academic profiles. In contemporary education, argumentative discourse remains a focal point in primary and secondary classrooms, bolstered by frameworks that advocate for inquiry-based learning and encourage meaningful dialogue on complex questions. Research underscores the benefits of deliberative discourse for enhancing student participation and learning outcomes (Asterhan & Schwarz, 2016; Felton et al., 2009).

Furthermore, collaborative approaches, particularly through SAC, can facilitate effective discourse on complex academic problems (Johnson & Johnson, 1988).

This dissertation presented a research synthesis examining the effects of disputative and deliberative argumentation on learning outcomes, along with instructional strategies in science and social studies classrooms (Chapter 2). The synthesis identified key instructional components that positively influence learning. Subsequently, a descriptive multiple-case study (Chapter 3) explored three cases, offering a cross-case analysis of the interplay between structured student discourse with a more supportive and guided SAC and students' argumentative and historical thinking discourse with students of varying academic and cognitive profiles. Data from two teachers in three classes provided a foundation for a practitioner manuscript (Chapter 4), detailing a specific structure and guided supports to enhance student participation and discourse quality in high school history classrooms.

Expanding on this dissertation may equip researchers and teachers with greater access to knowledge, instructional practices, and resources to help argumentative discourse thrive in secondary classrooms. Finally, continued research in this area may improve the argumentation outcomes, both orally and in writing, of students from academically, culturally, and linguistically diverse populations.

Appendix A

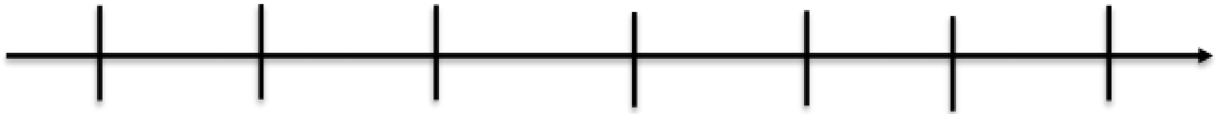
INVESTIGATION #5:
WILL THE CONSTRUCTION OF HARBORPLACE IN BALTIMORE IMPROVE LIFE IN THE CITY? DECIDE IF YOU WOULD HAVE VOTED TO APPROVE THE CONSTRUCTION OF THE INNER HARBOR.

In the 19th and early 20th centuries, Baltimore's manufacturing blossomed, contributing to transportation and urban development across the nation. Oil refineries, steel plants, as well as shipbuilding, lumber, iron, coal, and canning companies decked the waterfront. Business especially boomed during both World Wars, as the military needed airplanes, ships, and weapons.

Between the 1950s and 1970s, Baltimore began to experience deindustrialization. This happened in part because President Nixon’s newly formed economic relationship with China contributed to the exporting of manufacturing jobs. As a result, many wealthier, mostly White citizens, fled cities across the United States for the suburbs leaving poorer, mostly Black, and other minority residents in poor urban areas. Mayors across the United States began to implement programs to revitalize these areas.

To attract wealthier people to Baltimore, officials began to seize what they considered ruined areas, which were typically occupied by black and minority populations, for large scale construction projects. Mayor William Donald Schaefer and developers had a grand vision to build an attractive entertainment and retail district that would draw tourists, dollars, and national notice. At that time, in Baltimore, the question was whether to build a major attraction, “Harborplace” based on who it would benefit and who stood to lose. Therefore, our central question for this investigation is *Would you have voted to approve the construction of the Inner Harbor?* Take out your pens and get ready to investigate.

Construct your own timeline here:



Document A: Urban Renewal's Refugees

Head Note*:* Ron Howell, a reporter for the *Baltimore Sun,* wrote that while the construction of the Inner Harbor was praised for revitalizing the city in the 1970s, many city leaders and activists began to question the negative consequences of urban revival, such as the displacement of lower-income families and individuals.

There has been much talk in recent days, at various community meetings around the city, about an issue some claim will have significant effects on the future racial and socioeconomic makeup of several Baltimore neighborhoods.

The issue is displacement – the relocation of poor people from city neighborhoods that are experiencing "revitalization," or, the "return of the middle class."

Vincent P. Quayle, Director of the Saint Ambrose Housing Aid Center, has said that speculators in some neighborhoods increasingly are evicting tenants in order to sell the houses to middle class families or to other speculators.

A former Jesuit who holds a real estate brokers license, Mr. Quayle is one of the city's leading anti-displacement voices. He originally established the Saint Ambrose Housing Aid Center several years ago to help low-income people become homeowners.

Speculators [investors], he says, are trying to sell predominantly black Harwood and Barclay-Greenmount neighborhoods as extensions of Charles Village in order to attract white middle-class homeowners to those areas.

"Two years ago the 2500, 2600 and 2700 blocks of Guilford avenue were virtually all black," Mr. Quayle says…. Evictions of the former renters occurred in the process, he said, and the three blocks now each contain about six white families.

Mr. Quayle has predicted that largely crowded Park Heights in the largely black areas around it will become "dumping grounds" for blacks displaced from revitalized neighborhoods. He has maintained that landlords already are increasing the numbers of units in northwest Baltimore buildings.

Source: Excerpt from The Baltimore Sun, pg. K1**, by Ron Howell, March 12, 1978.**

Document B: Baltimore Aquarium Seeks Minority Applicants

Head Note: As construction of the Inner Harbor began in the 1970s, city leaders who were in charge of the construction and management of the Aquarium announced that they would actively recruit Black and other minorities for staff positions. This ad appeared in the *Baltimore Afro-American*, an African American owned newspaper from 1892 to the present.

The Baltimore Aquarium, currently under construction at Pier 3 of Baltimore's Inner Harbor, is actively recruiting qualified minorities to join its staff.

As part of its effort, the Aquarium invited minority community leaders to a special recruitment luncheon held at the World Trade Center Oct. 17.

During the luncheon, minority leaders were asked to assist the Aquarium in recruiting minority candidates for the various positions that will be available over the coming months and up until the Aquarium officially opens during the spring of 1981.

Aquarium director James S. Kepley, told the group of local as well as national minority leaders that the Baltimore Aquarium recognizes the need to hire minorities.

He recently said, "If the Baltimore Aquarium is to be successful, it must be a community oriented facility in all areas from hiring to educational programs, to volunteers."

Kepley adds, "In the future, we plan to participate in career seminars and other related activities to inform minorities of the job opportunities at the Baltimore Aquarium and opportunities that exist in other similar facilities."

Accountants, building service workers, cashiers, stenographers, membership coordinator, membership secretary, gift shop manager, security officer and chief engineer are just some of the positions that will be filled.

Source*:* Excerpt from The Baltimore Afro-American pg. 6, November 10, 1979.

Document C: Letter to the Editor: Inner Harbor Plan

Head Note: Martin Daugherty, a born and raised Baltimore citizen, wrote a letter to the editor of The Baltimore Sun, Baltimore's most prestigious newspaper. In this letter, he voiced his opinion for the construction of Harborplace, a hotly debated issue at the time.

Sir: I have waited as long as I can to speak out in public on the proposed Rouse [this is the name of the company] development of the Inner Harbor – Harborplace. Opponents have talked about the Inner Harbor as if it were carved in stone, a religious retreat or an ancient burial ground. Believe me, it's none of the above.

Back in the Sixties, in an effort to restore Baltimore to the position it held in the U.S. in the Nineteenth Century, when it was the second largest city in the country, the citizenry of Baltimore voted to rehabilitate the disheveled Inner Harbor. As a matter of fact, we voted to develop the whole west shore of the Inner Harbor.

Mr. Rouse and his enlightened firm, seeing this would deny the view, submitted a proposal to build using the northwestern corner only. This proposal left a beautiful view from the western shore, a large portion of the open space and an economically self-supporting business, drawing from the business district as well as tourists and convention-goers.

Source: The Baltimore Sun, pg. A16, August 31, 1978.

Document D: Furor over Harborplace continues around City Hall

Head Note: As city officials introduced the Harborplace construction plan, many began to question the financial legitimacy behind the city's partnership with the construction company, the Rouse Company.

Questions and answers about that financial agreement between Baltimore and Rouse Company, which wants to build a pair of market pavilions at the Inner Harbor, led to a moving press conference yesterday in and around City Hall.

Estimated income for the city in the first year, not including some state tax revenue, is projected at $628,000.

The agreement provides the company an initial 10 per cent profit. The company would pay full real and personal property taxes and ground rent of $100,000. The city would share 25 per cent of the developer's profits after certain costs – including debt service, taxes and ground rent – are deducted.

Representatives of Citizens for Preservation of the Inner Harbor, which is working to defeat the project, came to Mr. Pressman's conference to challenge the basis on which the projections were made.

The projected revenue data, the opponents persisted, are suspect because they were put together by the city and the Rouse Company. They asked, for example, if there had been any independent appraisal of the 3.2 acres of land at the corner of Pratt and Light streets where the project is to be constructed.

[Walter S. Orlinsky, the City Council president] suggested that the city has been hiding the real details of the agreement and that it may not have gotten the best possible deal for the taxpayers. Mr. Rouse, he said, appears to be getting a large profit – $400,000 – at small risk of his own capital.

Source: Excerpt from The Baltimore Sun, pg. C4, November 3, 1978.

Appendix B

Table 2.1

Description of Studies

Authors	Design	Grade	Participants (n)	Duration (sessions)	Discourse Type	Discourse Setting
1. Akbaş et al., 2019	MMR	7	94 students	24	Deliberation	SG, WC
2. Casado-Ledesma et al., 2023	Quasi	9	151 students	8	Deliberation	SG
3. Chen et al., 2016	MMR	5	22 students, 1 teacher	11	Deliberation	SG, WC
4. De La Paz & Wissinger, 2017	RCT	6 – 7	22 students (6th), 17 (7th)	15	Deliberation	SG
5. De La Paz et al., 2014	MMR	8	310 students, 13 teachers	18	Deliberation	-
6. De La Paz et al., 2017	Quasi	8	1,029 students, 36 teachers	54	Deliberation	SG
7. De La Paz et al., 2021	Quasi	10 – 12	74 students (10th), 138 (11th), 20 (12th)	5	Deliberation	SG, WC
8. De La Paz et al., 2022	Quasi	8	110 students, 1 teacher	12	Deliberation	SG, WC
9. De Vries et al., 2002	MMR	11	15 students	2	Deliberation	P
10. Del Favero et al., 2007	Quasi	8	100 students	7	Deliberation	SG, WC

11.	Felton et al., 2009	Quasi	7	101 students	8	Deliberation, Disputative	P
12.	Felton et al., 2015	Quasi	7	70 students	8	Deliberation, Disputative	P
13.	Ferretti et al., 2001	Quasi	5	87 students	33-37	Deliberation	SG, WC
14.	Garcia-Mila et al., 2013	Quasi	7	65 students	8	Deliberation, Disputative	P
15.	Gronostay, 2019	Quasi	8 – 9	221	4	Disputative	WC
16.	Grooms et al., 2018	RCT	10-11	63 students, 2 teachers	N/A	Deliberation	SG, WC
17.	Iordanou et al., 2019	RCT	Study 1: 5-6, Study 2: 6 – 8	Study 1: 88 students, Study 2: 107 students	Study 1: 9, Study 2: 9	Deliberation, Disputative	P, SG, WC
18.	Kent et al., 2015	RCT	11	24 students, 5 teachers	45	Deliberation	SG, WC
19.	Larraín et al., 2018	Quasi	4	61 students	9	Deliberation	SG, WC
20.	Larraín et al., 2021	Quasi	4	20 teachers, 502 students	30	Deliberation	SG, WC
21.	MacArthur et al., 2002	Quasi	6	31 students	25	Disputative	SG
22.	Munneke et al., 2006	RCT	10-11	175 students	6	Deliberation	P
23.	Murphy et al., 2018	Quasi	9-12	301 students, 7 teachers	24	Deliberation	SG
24.	Nussbaum & Edwards, 2011	MMR	7	30 students	13	Deliberation	WC

25. Reisman, 2012	Quasi	11	236 (182) students	105	Deliberation	WC
26. Relyea et al., 2022	Quasi	6	135 students, 6 teachers	20	Deliberation	SG, WC
27. van Drie et al., 2005	Quasi	11	72 students	6	Deliberation	P
28. Wanzek et al., 2015	RCT	8	6 teachers	45	Deliberation	SG, WC
29. Wissinger & De La Paz, 2016	RCT	6-7	151 students	15	Deliberation	SG, WC
30. Zhang et al., 2023	MMR	6	137 students, 3 teachers	20	Deliberation, Disputative	SG

Note. Participants in parentheses represent final sample used for analyses. MMR = mixed methods research, Quasi = quasi-experimental, RCT = randomized control trial study design, L = longitudinal, SG = small group, WC = whole class, P = pairs

Table 2.2

Demographics of Studies

Citation	Setting	Discipline	SWD	ML	GEN
1. Akbaş et al., 2019	Turkey	Social studies	-	-	94
2. Casado-Ledesma et al., 2023	Spain	Science	0	-	151
3. Chen et al., 2016	USA	Science	2	-	22
4. De La Paz & Wissinger., 2017	USA	Social studies	24	-	*15
5. De La Paz et al., 2014	USA	Social studies	-	-	310^{+}
6. De La Paz et al., 2017	USA	Social studies	71	54	482, 269*, 265**
7. De La Paz et al., 2021	USA	Social studies	2	7	218
8. De La Paz et al., 2022	USA	Science	32	-	78
9. de Vries et al., 2002	France	Science	-	-	15
10. Del Favero et al., 2007	Italy	Social studies	-	-	100
11. Felton et al., 2009	Spain	Science	-	-	101
12. Felton et al., 2015	Spain	Science	-	-	70
13. Ferretti et al., 2001	USA	Social studies	28	-	59

14. Garcia-Mila et al., 2013	Spain	Science	0	0	65
15. Gronostay, D., 2019	Germany	Social studies	-	-	221
16. Grooms et al., 2018	USA	Science	0	0	63
17. Iordanou et al., 2019	1: Cyprus, and 2: USA	1: Science, 2: social studies	-	-	1: 88 2: 107
18. Kent et al., 2015	USA	Social studies	24	-	-
19. Larraín et al., 2018	Chile	Science	-	-	61
20. Larraín et al., 2021	Chile	Science	-	-	502
21. MacArthur et al., 2002	USA	Social studies	9	-	22
22. Munneke et al., 2006	Netherlands	Science	-	-	175
23. Murphy et al., 2018	USA	Science	-	-	301
24. Nussbaum & Edwards, 2011	USA	Social studies	-	-	30
25. Relyea et al., 2022	USA	Science	-	70	65
26. Reisman, 2012	USA	Social studies	-	26	156, 54*
27. van Drie et al., 2005	Netherlands	Social studies	-	-	72**
28. Wanzek et al., 2015	USA	Social studies	36	-	322
29. Wissinger & De La Paz, 2016	USA	Social studies	20	0	151

30. Zhang et al., 2023	USA	Science	-	74	63

Note: Includes all participants, not just those used in a final sample. SWD = students with disabilities, ML = multilingual learners, GEN = general education students, a = struggling learners, b = advanced learners, c = specific demographics unknown.

Table 2.3

Effect Size by Treatment Condition

Study	Dependent Measure	IV	ES
Akbaş et al., 2019	Content learning	Argumentation-based science learning (ABSL)	$\eta^2 = 0.30$
De La Paz & Wissinger., 2017	Historical knowledge	Disciplinary Discussion	$\eta_p^2 = 0.23$
	Persuasive quality		$\eta_p^2 = 0.43$
	Historical thinking		$\eta_p^2 = 0.40$
	Historically thinking (maintenance)		$\eta_p^2 = 0.19$
De La Paz et al., 2014	Historical arguments	Historical thinking curriculum	$d = 0.56$
	Holistic quality		$d = 0.10$
De La Paz et al., 2017	Historical writing	Historical thinking curriculum	$\eta_p^2 = 0.45$
	Holistic quality		$\eta_p^2 = 0.32$
De La Paz et al., 2021	Historical writing	Historical thinking curriculum	$d = 0.33$
	Contextualized thinking		$d = 0.44$
	Use of evidence		$d = 0.32$
De La Paz et al., 2022	Written arguments (near transfer)	Cognitive apprenticeship curriculum	$d = 1.08$
	Written arguments (far transfer)		$d = 0.76$
	SWD Written arguments (far transfer)		$d = 0.73$
Del Favero et al., 2007	Understanding of historical inquiry	Discussion-based problem solving	$\eta_p^2 = 0.06$
	Situational interest		$\eta_p^2 = 0.05$
Felton et al., 2009	Content learning	Deliberation	$\eta^2 = 0.11$
	Argument quality		$\eta^2 = 0.13$
Felton et al., 2015	Argumentative exchange length	Deliberation	$d = 0.74$
Ferretti et al., 2001	Unit knowledge	Social Studies Project-Based Learning Unit	*g = 1.86 g = 2.33
	Historical content		*g = 1.85 g = 2.81
	Historical inquiry		*g = 0.53

			g = 1.29
Garcia-Mila et al., 2013	Quality of Argumentation (Level 4, Consensus)	Consensus v Persuasion	*r* = 0.43
	Claim-Data (Persuasion)		*d* = 0.53
	Claim Rebuttal (Consensus)		*r* = 0.37
	Claim-Data-Rebuttal (Consensus)		*r* = 0.315
	Claim-Data-Warrant-Rebuttal (Consensus)		*r* = 0.34
	Claim-Data-Warrant-Backing-Rebuttal (Consensus)		*r* = 0.33
Grooms et al., 2018	Scientific performance task (familiar task)	Assessment of Scientific Argumentation in the Classroom	*d* = 0.59
Iordanou et al., 2019	Functional evidence-based idea units	Q&A cards	g = 0.33
	Weaken-other position statements		g = 0.50
Kent et al., 2015	Assessment of Social Studies Knowledge	Team-Based Learning (TBL)	g = 0.50
	Vocabulary knowledge		g = 1.01
	Content knowledge		g = 0.38
Larraín et al., 2018	Content knowledge test (maintenance)	Peer-group argumentation	$\eta^2 = 0.31$
	Group-work argumentative utterances		$\eta^2 = 0.20$
	Time devoted to group work		$\eta^2 = 0.85$
Larraín et al., 2021	Group-work argumentative utterances	Computer-enhanced curriculum	g = 1.48
	Group-work argumentative utterances	Standard support	g = 1.31
MacArthur et al., 2002	Knowledge posttest	Special Education group	g = 0.49
Munneke et al., 2007	Argumentative Activity	CSCL Diagram	$\eta^2 = 0.41$

	Breadth of argument		$\eta^2 = 0.31$
	Depth of argument		$\eta^2 = 0.14$
Murphy et al., 2018	Written argumentation	Quality Talk	$\eta^2 = 0.06$
	Conceptual understanding		$\eta^2 = 0.03$
Nussbaum & Edwards, 2011	Mean frequency of weighing stratagems	Argument from sign test	$d = 1.14$
	Increased weighing of arguments		$d = 1.56$
	Decrease in Pseudo-integration		$d = 1.19$
Relyea et al., 2022	Academic vocabulary knowledge	Dialogic Inquiry for Socio-scientific and Conceptual Understanding in School Science (DISCUSS)	$d = 0.43$
	Scientific argumentation		$d = 0.40$
Reisman, 2012	Historical thinking	Document-based curriculum	$\eta_p^2 = 0.09$
	Historical thinking (maintenance)		$\eta_p^2 = 0.08$
	Factual knowledge		$\eta_p^2 = 0.03$
	Reading comprehension		$\eta_p^2 = 0.05$
van Drie et al., 2005	Chat contributions	Diagram argument organizer	g = 1.12
Wanzek et al., 2015	Written essay: content	Team-Based Learning (TBL)	g = 0.31
	Written essay: support		g = 0.16
	Social studies knowledge		g = 0.07
Wissinger & De La Paz, 2016	Historical knowledge	Experimental condition	$\eta_p^2 = 0.44$
	Substantiation		$\eta_p^2 = 0.24$
	Mature rebuttals		$\eta_p^2 = 0.41$
Zhang et al., 2023	Total words	Dialogic Inquiry for Socio-scientific and Conceptual Understanding in School Science (DISCUSS)	g = 0.74
	Total idea units		g = 0.92

Note: Effect size interpretation: η_p^2 and η^2 - 0.01 = small, 0.06 = medium, 0.14 = large; d and g - 0.2 = small, 0.5 = medium, 0.8 = large. a represents SWD.

Table 3.1

Student Demographic Information

Baseline characteristic	Adams: Co-Taught Mod SAC			Adams: Honors Mod SAC		Fern: Honors SAC	Total
	Mod 1	Mod 2	Mod 3	Mod 4	Mod 5	Mod 6	
Gender							
Female (N= 2 or 3)	4	3	5	7	4	7	30
Male (N = 2 or 3)	8	9	7	5	8	5	42
Race							
White	5	6	5	6	9	7	38
Black	3	5	5	4	0	3	20
Hispanic	3	1	1	0	0	2	7
Asian	0	0	0	2	1	0	3
Multi Racial	1	0	1	0	2	0	4
IEP	7	3	2	1	0	2	15
504	1	1	2	0	2	2	8
ELL	0	1	0	0	0	0	1
FARM	2	2	2	1	1	0	8

Note. N = 72. “Mod SAC” refers to modified SAC used in Ms. Adams’ Class periods.

Table 3.2

Selected Students Characteristics

Student	MOD	Teacher	Gender	Race	IEP	504	Attendance %	District Score
Yogi	Honors 2	Fern	M	Black	N	Y	96	Threshold
Nel	Honors 2	Fern	M	White	N	Y	95	Threshold
Isle	Honors 2	Fern	F	White	N	N	98	Mastery
Paisley	Honors 2	Fern	F	White	N	N	96	Mastery
Ford	Co-taught	Adams	M	White	Y	N	97	Below threshold
Kristen	Co-taught	Adams	F	Black	N	N	94	Threshold
Julius	Co-taught	Adams	M	Black	N	Y	93	Threshold
Bryce	Co-taught	Adams	M	Hispanic	Y	N	97	Below threshold
Jake	Honors 1	Adams	M	White	N	N	98	Threshold
Taylor	Honors 1	Adams	M	White	N	N	95	Mastery
Rex	Honors 1	Adams	M	Asian	N	N	100	Mastery
Heather	Honors 1	Adams	F	White	N	N	97	Mastery

Note. District Scores were for a history assessment and are described as follows: “Mastery” is a score between 70% - 100%, “Threshold” is a score between 50% - 70%, and “Below threshold” is a score below 50%.

Table 3.3

Overview of the Six Historical Investigations

Investigation	Central Historical Question	Historical, political, and local connections
#1- Progressivism	Who was a stronger advocate for African Americans during the early 1900s, Booker T. Washington or W.E.B Du Bois?	Analyzing what happens when individuals work within existing societal structures vs. suggesting that society make fundamental changes.
#2- Great Depression	What path offered the best chance for a better life for the "Okies" in the 1930s: remaining in the Dust Bowl or migrating to the West?	Analyzing how individuals within larger social groups made economic decisions in times of need
#3- Containment of Communism	Why did the U.S. lead the 1954 coup in Guatemala?	Analyzing what happens when governments prioritize its economic interests in the world.
#4- Civil Rights for Latinos	To what extent was the UFW a successful movement?	Analyzing how social movements, such as the BLM and other movements (e.g., March for Our Lives), push for lasting change in America
#5- Deindustrialization	Will the construction of Harborplace in Baltimore improve life in the city?	Analyzing how individuals within a community interact with local officials and businesses in revitalizing an urban area.
#6- War on Drugs and Mass Incarceration	Did President Ronald Reagan's "Intensified War on Drugs" have its intended consequences?	Analyzing the origins and impact of the "War on Drugs" on policing and mass incarceration

Table 3.4

Sourcing Historical Judgment Rubric

Level	Description	Example
1	Student does not mention the source of the documents	"Document B says that it's going to help."
2	Student mentions source but does not use source information to evaluate documents	"Document A was written by Ron Howell from the Baltimore Sun."
3	Student mentions source and evaluates the document using ahistorical reasoning	"I was going to say that the author's reliable because it's during the time period but also, he's a reporter, so he's obviously writing this for his own benefit. So, it could be biased."
4	Student critiques document using source information and uses their critique to strengthen their argument.	"I agree with what she was saying with document C being written by just the citizen who might not have that much information as the city officials who wrote or who were part of document D…"

Note. Modified from Nokes (2017) writing rubric for discourse purposes.

Table 3.5

Contextualization Historical Judgment Rubric

Score	Description	Example
1	(a) No context mentioned OR (b) Student uses anachronisms (e.g., makes a chronological mistake or uses information from another time period w/o noting the different era)	"I think it was like right after WWII."
2	(a) Mentions the time, place, or audience in the documents (e.g., "In document 2, a letter to the commander in the 3rd district . . ." "After the Civil War . . .")	"Because it was beneficial to the majority of the citizens in Baltimore at the time."
3	(a) Describes contextual details alongside the evidence without explicitly connecting them.	"What [he] said, stagflation was going on in the economy."
4	(a) Integrates context and evidence in an explanation or conclusion. OR (b) Uses context and evidence together to draw a conclusion or make an inference.	"The civil rights movement was happening and at the time a lot of middle income or wealthy White people had left because of de-industrialization."

Note. Modified from Monte-Sano & De La Paz (2012) writing rubric for discourse purposes.

Figure 2.1

Prisma Diagram

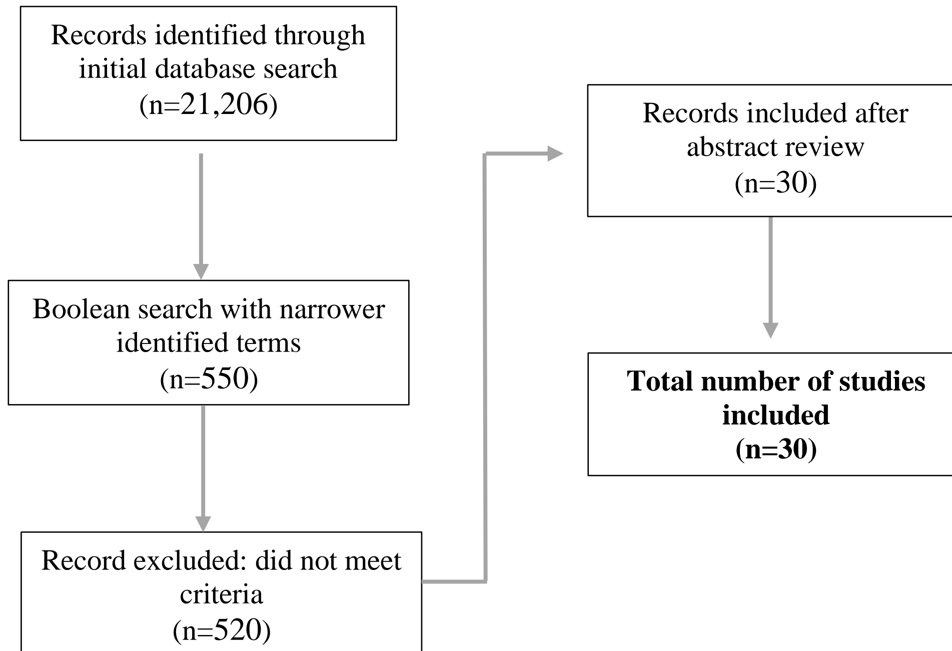


Figure 3.1

The Three-node Argumentation for Learning framework (Asterhan & Schwarz, 2016).

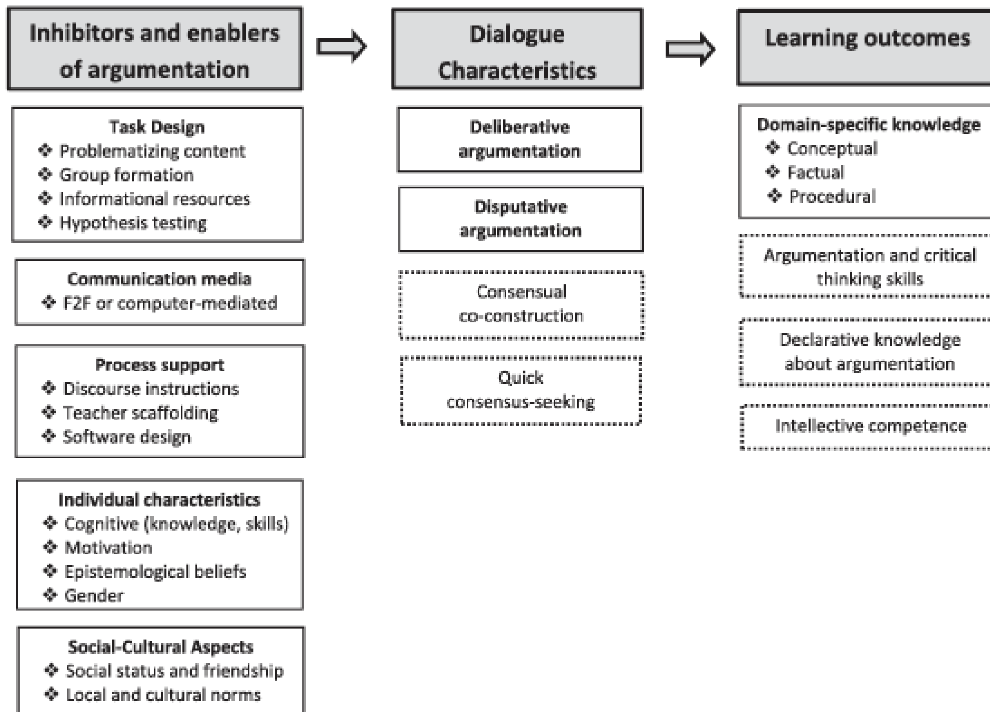


Figure 3.2

SAC+G Graphic Organizer

<table>
<tr><td colspan="3">WARM UP:
What is the historical question about? Or, why did people have different opinions on this issue?</td></tr>
<tr><td>1. What is the 1st source's perspective? Write WHO it is…</td><td>Answer:</td><td>Quote that shows the source's perspective:</td></tr>
<tr><td>2. What is the 2nd source's perspective? Write WHO it is…</td><td>Answer:</td><td>Quote that shows the author's perspective:</td></tr>
<tr><td colspan="3">3. Do you think the construction of Harborplace in Baltimore will improve life in the city?</td></tr>
<tr><td>4. Look for evidence that supports what you think.</td><td colspan="2">Quotes to support what I figured out:</td></tr>
<tr><td>5. Evaluate the 1st source/document.

Use 1 EAD question (see foldable or list)</td><td colspan="2">Answer 1:</td></tr>
<tr><td>6. Evaluate the 2nd source/document.

Use 1 EAD question (see foldable or list)</td><td colspan="2">Answer 1:</td></tr>
<tr><td>7. Decide which perspective you believe, based on the evidence.

Put a check mark or star by author 1 or 2, above.</td><td colspan="2">Organize and write:
A. Put a star or ☑ by quotes you will use and number (1, 2, 3) the quotes to show the order you will use them.
B. Star or ☑ which evaluations that you will use.</td></tr>
</table>

	C. Write your essay, using “sentence starters” from the next page to introduce judgments.
Critical Questions Sentence Starters for Evaluating Sources & Evidence for Question 4 & 5	
Author’s reliability (E of IREAD) • I trust/doubt this author because…. • This (author) is reliable or unreliable because… • I think the audience is… • This (author) might have known more about… • The author probably believes…	
Influence of context (A of IREAD) • Compared to other people’s views at the time/ place, the (author) • Given what else was going on at this time/place (source’s) facts makes sense because… • the source may have been affected by (this event from history) ….	
Quality of the author’s facts (D of IREAD) • (Author) supports this point by showing that… • The evidence used to support the author’s claims is… • This document might not give me the whole picture because… • Another document to consider might be…	
Corroboration: • Given what else was going on at the time (or in name of place) • Based on what I know from the background information • These documents all agree/disagree about . . . • Another document to consider might be . . .	

Figure 3.3

Argumentative Utterances for Co-taught and Honors 1

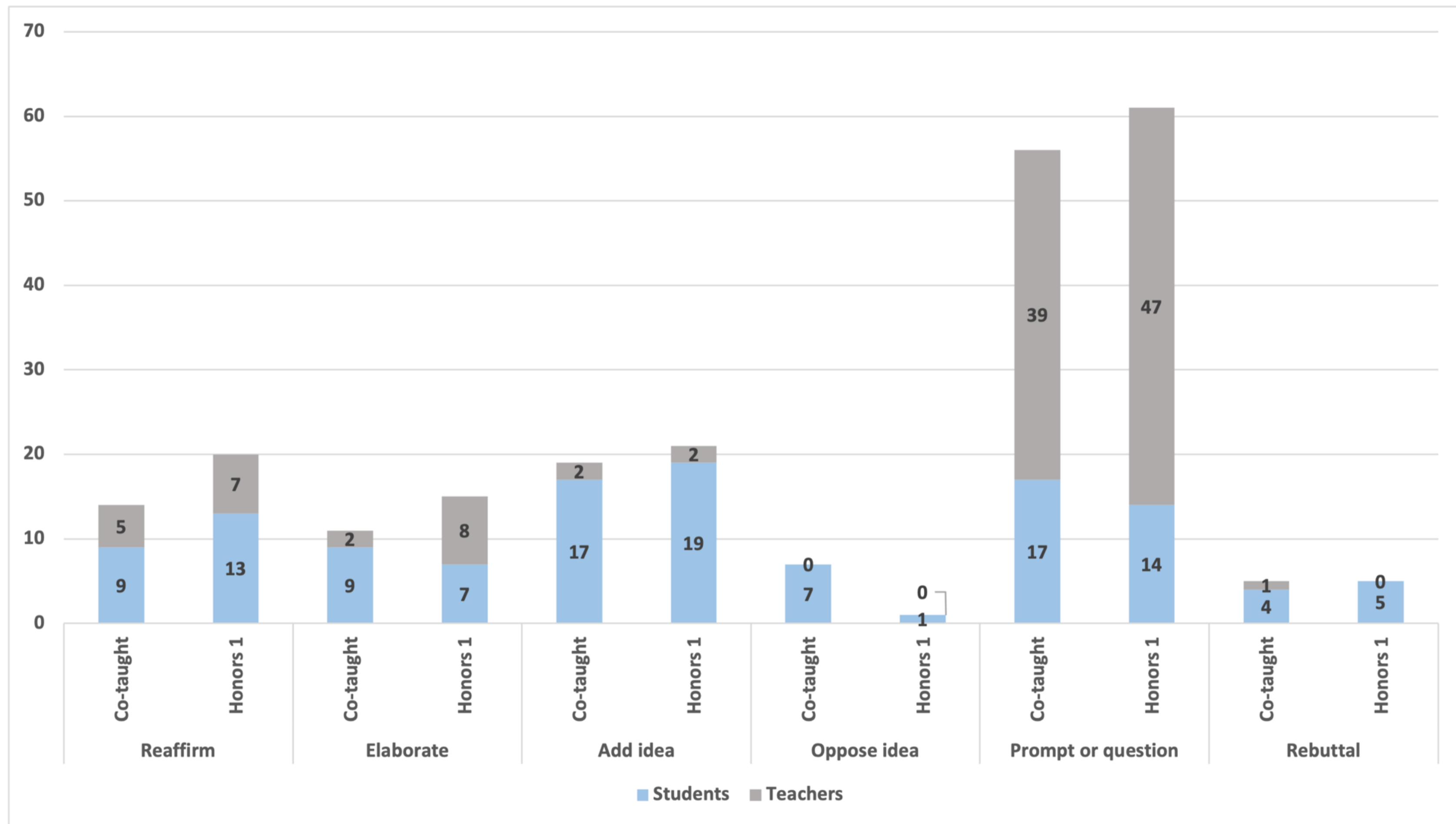


Figure 3.4

Argumentative Utterances for Honors 1 and Honors 2

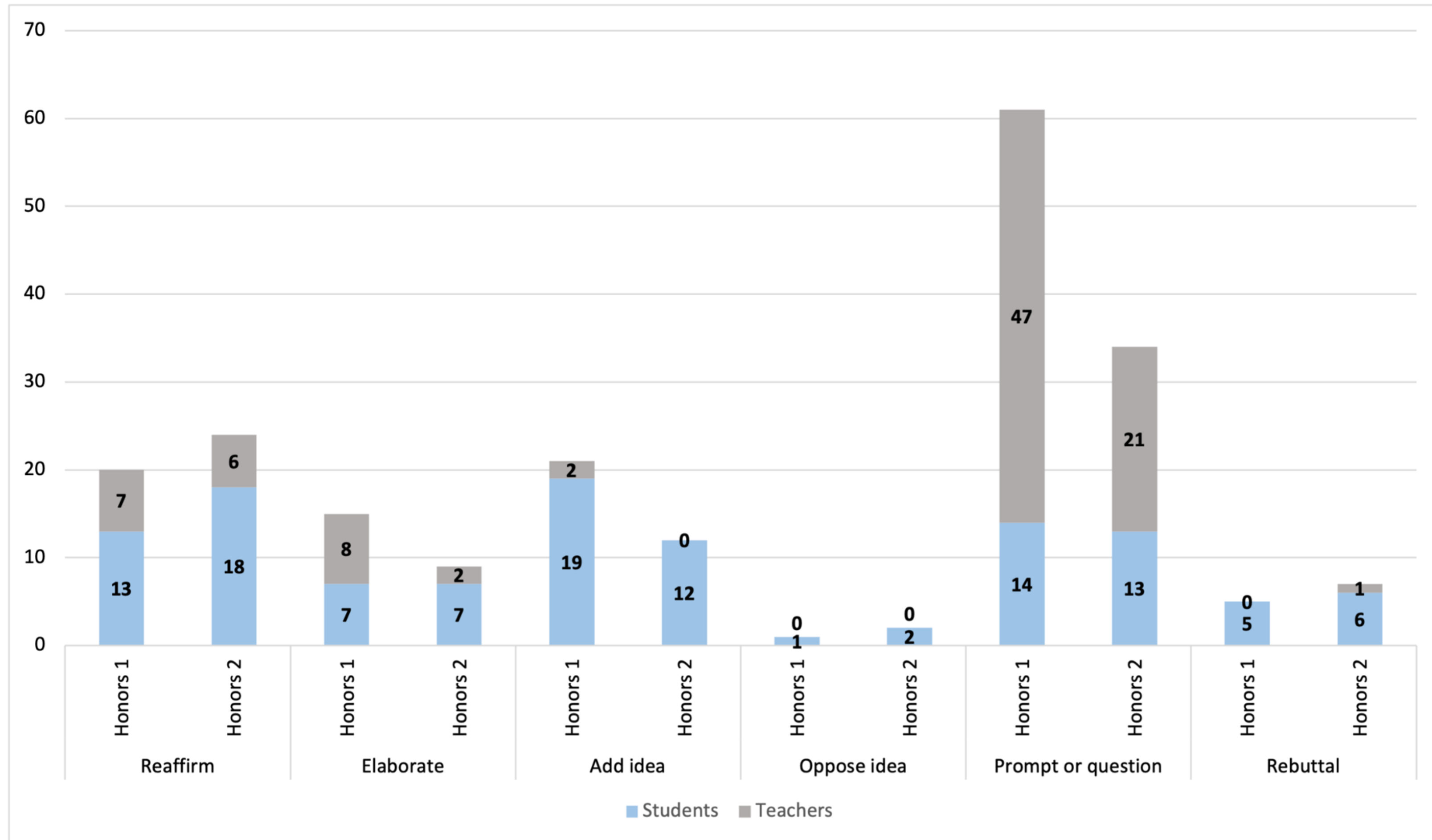


Figure 3.5

Historical Thinking Utterances in Co-taught and Honors 1

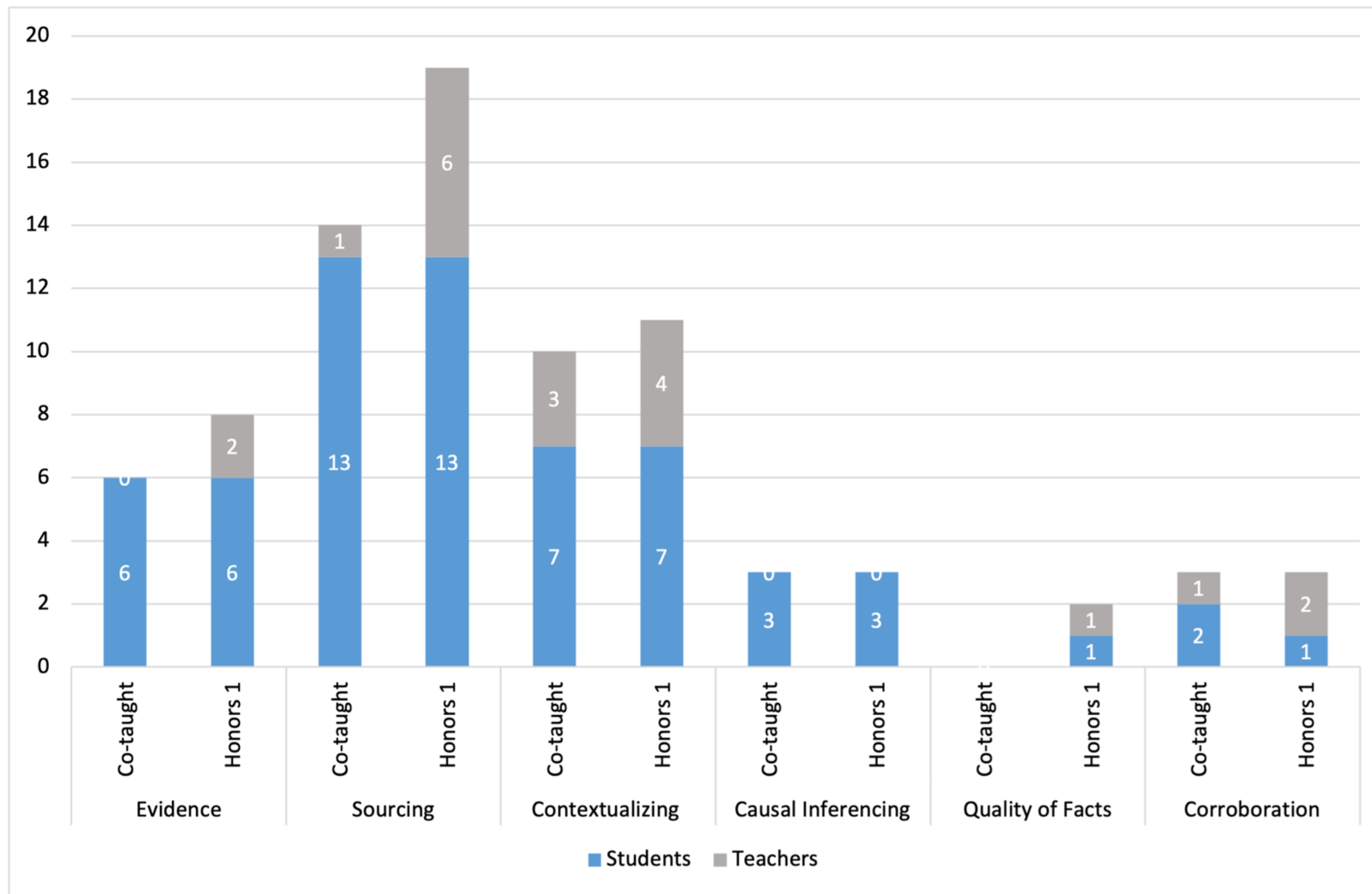


Figure 3.6

Historical Thinking Utterances in Honors 1 and Honors 2

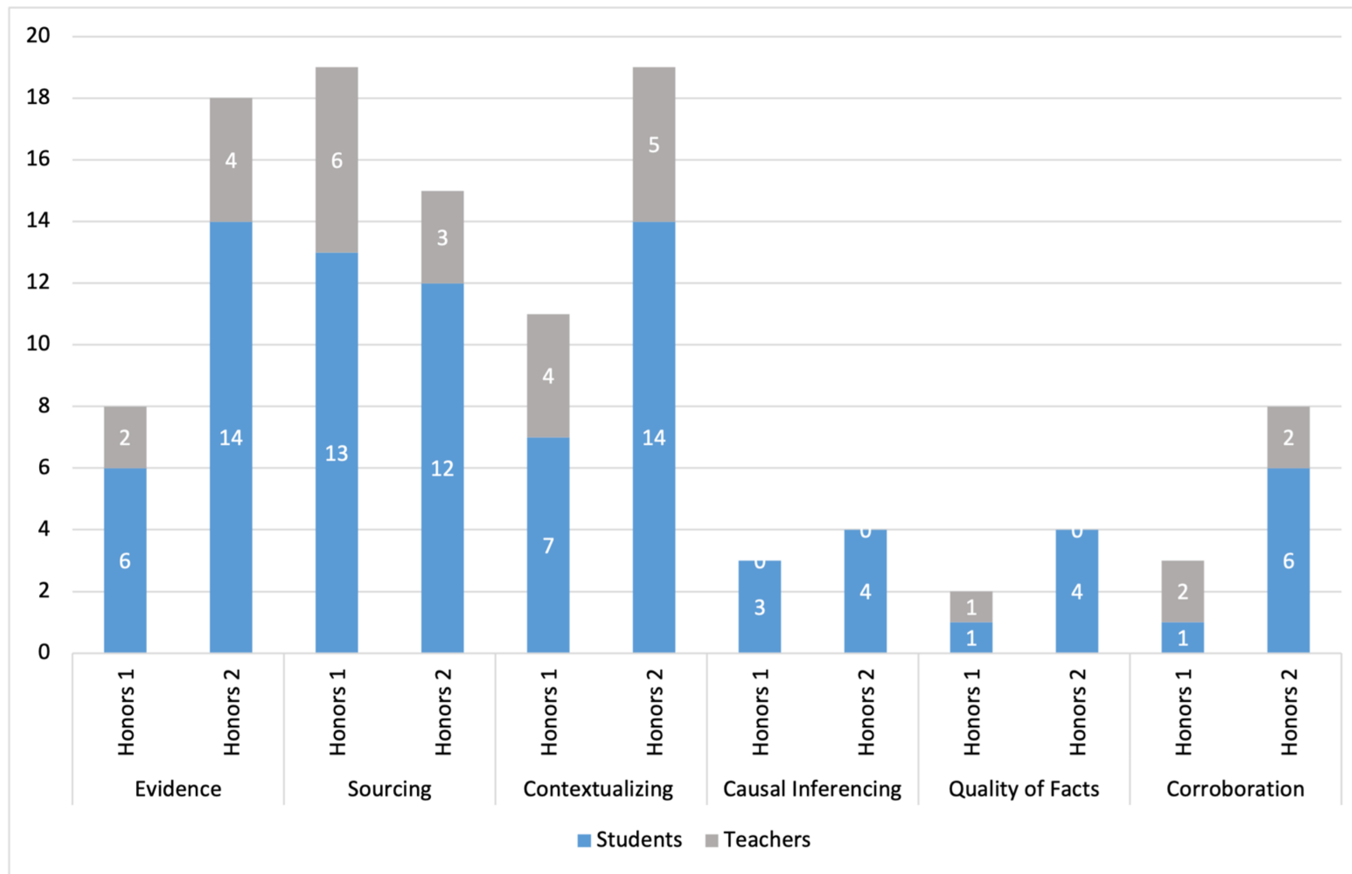


Figure 4.1

SAC Graphic Organizer (De La Paz et al., 2014; 2017; Monte-Sano et al., 2017).

State your group's initial ideas on what the CHQ is about

In the box state your initial ideas about the controversy, using background information, timeline, etc.

Document #1 or #3

Did the US help lead the 1954 coup in Guatemala to contain communism?	Select evidence (quotations) from this document that help you understand why the US helped lead the 1954 coup to contain communism.

Document #2 or #4

Did the US help lead the 1954 coup in Guatemala due to private interests?	Select evidence (quotations) from this document that help you understand why the US helped lead the 1954 coup due to private interests.

Discuss with your group:

Why did you select the perspective of one author or the other?

What did you decide about the CHQ? Write your interpretation:

Critical Questions

What is the historical question asking? What do you know about the time period and setting?

What is each author's point of view?
Do the documents agree? If not, why?
What evidence is important to you, from each author? Why?
What judgments did you make?
Did we have different judgments?

Figure 4.2

Student Sample of SAC+G Graphic Organizer

Modified SAC graphic organizer.

WARM UP:
What is the historical question about?
Or, why did people have different opinions on this issue?
• That it will improve life.

1. What is the 1st source's perspective? ***Write WHO it is…***	Answer:	Quote that shows the source's perspective:
Ron Howell		"Mr… largely black areas around it will become "dumping grounds"…"
2. What is the 2nd source's perspective? *Write WHO it is…*	**Answer:**	**Quote that shows the author's perspective:**
Baltimore Afro-American pg. 6 Nov. 10, 1979		"If the Baltimore Aquarium is to be succesful, it must be a community oriented facility in all areas from hiring to educational programs, to volunteers"

3. Do you think the construction of Harborplace in Baltimore will improve life in the city? Yes due to people wanted jobs and opertunities.

4. Look for evidence that supports what you think.	Quotes to support what I figured out:
	"Accountants, building servic workers,… are just some of the positions that will be filled."

5. Evaluate the 1st source/document. Use 1 EAD question (see foldable or list)	Answer 1: This author might have known more about the possible outcome to the city.
6. Evaluate the 2nd source/document. Use 1 EAD question	Answer 1:
7. Decide which perspective you believe, based on the evidence. Put a check mark or star by author 1 or 2, above.	Organize and write: A. Put a star or ☑ by <u>quotes</u> you will use **and** number (1, 2, 3) the quotes to show the order you will use them. A. Star or ☑ **which** <u>evaluations</u> that you will use. A. Write your essay, using "sentence starters" from the next page to introduce judgments.

References

* indicate studies included in the analysis of Chapter 2: Research Synthesis.

*Akbaş, Y., Sahin, I. F., & Meral, E. (2019). Implementing argumentation-based science learning approach in social studies: Academic achievement and students' views. *Review of International Geographical Education Online*, *9*(1), 209–245.

Alexander, P. A. (2020). Methodological guidance paper: The art and science of quality systematic reviews. *Review of Educational Research, 90*(1), 6–23. https://doi-org.proxy-um.researchport.umd.edu/10.3102/0034654319854352

Alexander, P. A., Fusenig, J., Schoute, E. C., Singh, A., Sun, Y., & van Meerten, J. E. (2023). Confronting the challenges of undergraduates' argumentation writing in a "Learning How to Learn" course. *Written Communication*, *40*(2), 482-517. https://doi-org.proxy-um.researchport.umd.edu/10.1177/07410883221148468

Andriessen, J., & Baker, M. J. (2014). Arguing to Learn. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 439–460). Cambridge: Cambridge University Press.

Asterhan, C. S., & Babichenko, M. (2015). The social dimension of learning through argumentation: Effects of human presence and discourse style. *Journal of Educational Psychology, 107*(3), 740–755. http://dx.doi.org.proxy-um.researchport.umd.edu/10.1037/edu0000014

Asterhan, C., & Schwarz, B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of*

Educational Psychology, 99, 626–639. https://doi-org.proxy-um.researchport.umd.edu/https://content.apa.org/journals/edu/99/3/626

Asterhan, C. S. C., & Schwarz, B. B. (2009a). The role of argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialogue. *Cognitive Science, 33*, 373–399.

Asterhan, C., & Schwarz, B. (2016). Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist, 51*, 164-187. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00461520.2016.1155458

Asterhan, C. S. C., Butler, R., & Schwarz, B. B. (2010). Goals for learning and interaction in argumentation and conceptual change. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010) - Volume 1*. Chicago IL: International Society of the Learning Sciences.

Bae, C. L., Mills, D. C., Zhang, F., Sealy, M., Cabrera, L., & Sea, M. (2021). A systematic review of science discourse in K–12 urban classrooms in the United States: Accounting for individual, collective, and contextual factors. *Review of Educational Research*, *91*(6), 831-877. https://doi-org.proxy-um.researchport.umd.edu/10.3102/00346543211042415

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*(1), 29-58. https://doi-org.proxy-um.researchport.umd.edu/https://www.aera.net/publications/?id=319

Bennett, J., Hogarth, S., Lubben, F., Campbell, B., & Robinson, A. (2010). Talking science: The research evidence on the use of small group discussions in science teaching. *International Journal of Science Education*, *32*(1), 69-95. https://doi-org.proxy-um.researchport.umd.edu/10.1080/09500690802713507

Berland, L. K. (2011). Explaining variation in how classroom communities adapt the practice of scientific argumentation. *Journal of the Learning Sciences, 20*(4), 625–664. https://doi-org.proxy-um.researchport.umd.edu/10.1080/10508406.2011.591718

Berland, L. K., & Lee, V. R. (2012). In pursuit of consensus: Disagreement and legitimization during small-group argumentation. *International Journal of Science Education, 34*(12), 1857-1882. https://doi-org.proxy-um.researchport.umd.edu/10.1080/09500693.2011.645086

Bolick, C. M., Adams, R., & Willox, L. (2010). The marginalization of elementary social studies in teacher education. *Social Studies Research and Practice, 5*(2), 1-22.

Boyd, M. P., & Rubin, D. L. (2002). Elaborated student talk in an elementary ESoL classroom. *Research in the Teaching of English, 36*(4), 495–530.

Brooks, S. (2009). Historical empathy in the social studies classroom: A review of the literature. *Journal of Social Studies Research*, *33*(2), 213-234. https://doi-org.proxy-um.researchport.umd.edu/https://www.thejssr.com/

Brownell, M. T., Benedict, A. E., Leko, M. M., Peyton, D., Pua, D., & Richards-Tutor, C. (2019). A continuum of pedagogies for preparing teachers to use

high-leverage practices. *Remedial and Special Education*, *40*(6), 338-355. https://doi-org.proxy-um.researchport.umd.edu/10.1177/0741932518824990

Bueso, L. (2022). Civic equity for students with disabilities. *Teachers College Record, 124*(1), 62-86. https://doi-org.proxy-um.researchport.umd.edu/10.1177/01614681221086092

Capobianco, B. M., DeLisi, J., & Radloff, J. (2018). Characterizing elementary teachers' enactment of high-leverage practices through engineering design-based science instruction. *Science Education*, *102*(2), 342-376. https://doi-org.proxy-um.researchport.umd.edu/10.1002/sce.21325

*Casado-Ledesma, L., Cuevas, I., & Martín, E. (2023). Learning science through argumentative synthesis writing and deliberative dialogues: A comprehensive and effective methodology in secondary education. *Reading and Writing*, *36*(4), 965-996. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11145-021-10191-0

Cervetti, G. N., DiPardo, A. L., & Staley, S. J. (2014). Entering the conversation: Exploratory talk in middle school science. *Elementary School Journal, 114*(4), 547–572. https://doi-org.proxy-um.researchport.umd.edu/10.1086/675638

*Chen, Y. C., Hand, B., & Park, S. (2016). Examining elementary students' development of oral and written argumentation practices through argument-based inquiry. *Science & Education, 25*(3/4), 277–320. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11191-016-9811-0

Colley, C., & Windschitl, M. (2016). Rigor in elementary science students' discourse: The role of responsiveness and supportive conditions for talk. *Science*

Education, *100*(6), 1009-1038. https://doi-org.proxy-um.researchport.umd.edu/10.1002/sce.21243

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American educator, 15*(3), 6-11.

Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & J. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 3–16). Russell Sage Foundation.

Darnon, C., Butera, F., & Harackiewicz, J. M. (2007). Achievement goals in social interactions: Learning within mastery vs. performance goals. *Motivation & Emotion, 31*, 61–70. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11031-006-9049-2

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

*De La Paz, S., & Wissinger, D. R. (2017). Improving the historical knowledge and writing of students with or at risk for LD. *Journal of Learning Disabilities, 50*(6), 658–671. https://doi-org.proxy-um.researchport.umd.edu/10.1177/0022219416659444

De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. (2012). Adolescents' disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written*

Communication, *29*(4), 412-454. https://doi-org.proxy-um.researchport.umd.edu/10.1177/0741088312461591

*De La Paz, S., Felton, M., Monte-Sano, C., Croninger, B., Jackson, C., Deogracias, J., & Hoffman, B.P. (2014). Developing historical reading and writing with struggling adolescent readers: Implementation and learning outcomes. *Theory and Research in Social Education, 42*(2), 228–274. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00933104.2014.908754

*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. https://doi-org.proxy-um.researchport.umd.edu/10.1002/rrq.147

*De La Paz, S., Wissinger, D. R., Gross, M., & Butler, C. (2021). Strategies that promote historical reasoning and contextualization: a pilot intervention with urban high school students. *Reading & Writing, 35*(2), 353–376. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11145-021-10183-0

De La Paz, S., Butler, C., Levin, D., Felton, M. (2022). Effects of a cognitive apprenticeship on transfer of argumentative writing in middle school science. *Learning Disability Quarterly*. 1-14. https://doi-org.proxy-um.researchport.umd.edu/10.1177/07319487221119365

De La Paz, S., Otarola, J., & Butler, C. (2023). Using virtual teacher PD to support implementation of a historical literacy curriculum. *Social Studies Research*

and Practice, 18(2), 164-179. https://doi-org.proxy-um.researchport.umd.edu/10.1108/SSRP-08-2023-0044

*de Vries, E., Lund, K., & Baker, M. (2002). Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *The Journal of the Learning Sciences, 11*(1), 63-103. https://doi-org.proxy-um.researchport.umd.edu/10.1207/S15327809JLS1101_3

*Del Favero, L.D., Boscolo, P., Vidotto, G., & Vicentini, M. (2007). Classroom discussion and individual problem-solving in the teaching of history: Do different instructional approaches affect interest in different ways? *Learning and Instruction*, *17*(6), 635-657. https://doi-org.proxy-um.researchport.umd.edu/10.1016/j.learninstruc.2007.09.012

Dockrell, J. E., & Connelly, V. (2016). The relationships between oral and written sentence generation in English speaking children: The role of language and literacy skills. *Written and spoken language development across the lifespan: Essays in honour of Liliana Tolchinsky*, 161-177.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and instruction*, *20*(4), 399-483.

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. https://doi-org.proxy-um.researchport.umd.edu/10.1002/sce.20012

*Felton, M., Garcia-Mila, M., & Gilabert, S. (2009). Deliberation versus dispute: The impact of argumentative discourse goals on learning and reasoning in the science classroom. *Informal Logic, 29*(4), 417-446.

*Felton, M., Garcia-Mila, M., Villarroel, C., & Gilabert, S. (2015). Arguing collaboratively: Argumentative discourse types and their potential for knowledge building. *British Journal of Educational Psychology*, 85, 372-386. https://doi-org.proxy-um.researchport.umd.edu/10.1111/bjep.12078

*Ferretti, R. P., MacArthur, C. D., & Okolo, C. M. (2001). Teaching for historical understanding in inclusive classrooms. *Learning Disability Quarterly*, *24*(1), 59-71.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, *35*(1), 39-50.

Ford, M. J. (2008). Grasp of practice' as a reasoning resource for inquiry and nature of science understanding. *Science & Education, 17*(2), 147–177. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11191-006-9045-7

*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. https://doi-org.proxy-um.researchport.umd.edu/10.1002/sce.21057

Gargiulo, R. M., & Metcalf, D. (2023). *Teaching in Today's Inclusive Classrooms: A Universal Design for Learning Approach* (4th Edition).

Gotwals, A. W., Wright, T. S., Domke, L., & Anderson, B. (2022). Science talk in elementary classrooms: A synthesis of the literature. *The Elementary School*

Journal, *122*(4), 642-673. https://doi-org.proxy-um.researchport.umd.edu/10.1086/719410

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

Grant, S. G., Swan, K., & Lee, J. (2017). *Inquiry-based practice in social studies education: Understanding the inquiry design model.* Routledge.

Gronostay, D. (2016). Argument, counterargument, and integration? Patterns of argument reappraisal in controversial classroom discussions. *Journal of Social Science Education*, *15*(2), 42-56.

*Gronostay, D. (2019). To argue or not to argue? The role of personality traits, argumentativeness, epistemological beliefs and assigned positions for students' participation in controversial political classroom discussions. *Unterrichtswissenschaft, 47*(1), 117-135.

*Grooms, J., Sampson, V., & Enderle, P. (2018). How concept familiarity and experience with scientific argumentation are related to the way groups participate in an episode of argumentation. *Journal of Research in Science Teaching, 55*(9), 1264–1286. https://doi-org.proxy-um.researchport.umd.edu/10.1002/tea.21451

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching: Theory and Practice*, *15*(2), 273-289. https://doi-org.proxy-um.researchport.umd.edu/10.1080/13540600902875340

Hackling, M., Smith, P., & Murcia, K. (2010). Talking science: Developing a discourse of inquiry. *Teaching Science: The Journal of the Australian Science Teachers Association, 56*(1), 17-22. https://doi-org.proxy-um.researchport.umd.edu/https://www.asta.edu.au/resources/teachingscience

Herrenkohl, L. R. (2006). Intellectual role taking: Supporting discussion in heterogeneous elementary science classes. *Theory Into Practice, 45*(1), 47–54. https://doi.org/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. https://doi-org.proxy-um.researchport.umd.edu/10.1080/10508406.2013.799475

Hyland, K. (2014). Disciplinary discourses: Writer stance in research articles. In *Writing: Texts, processes and practices* (pp. 99-121). Routledge.

Hynd, C. R. (1999). Teaching students to think critically using multiple texts in history. *Journal of Adolescent & Adult Literacy*, *42*(6), 428-436.

*Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction*, *63*. https://doi-org.proxy-um.researchport.umd.edu/10.1016/j.learninstruc.2019.05.004

Johnson, D. W., & Johnson, R. T. (1975). *Learning together and alone: Cooperation, competition, and individualization*. Englewood Cliffs, NJ: Prentice-Hall.

Johnson, D. W., & Johnson, R. T. (1979). Conflict in the classroom: Controversy and learning. *Review of educational research*, *49*(1), 51-69.

Johnson, R. T., & Johnson, D. W. (1985). Student-student interaction: Ignored but powerful. *Journal of teacher education*, *36*(4), 22-26.

Johnson, D. W., & Johnson, R. T. (1988). Critical thinking through structured controversy. *Educational leadership*, *45*(8), 58-64.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1996). Academic controversy: enriching college instruction through intellectual conflict. *ASHE-ERIC Higher Education Report*, *25*(3), 1–123.

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc.

Kelly, G. J. (2007). Discourse in science classrooms. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 443–469).

Kelly, G. J., & Bazerman, C. (2003). How students argue scientific claims: A rhetorical-semantic analysis. *Applied Linguistics*, *24*(1), 28-55.

*Kent, S., Wanzek, J., Swanson, E. A., & Vaughn, S. (2015). Team-based learning for students with high-incidence disabilities in high school social studies classrooms. *Learning Disabilities Research & Practice, 30*(1), 3–14. https://doi-org.proxy-um.researchport.umd.edu/10.1111/ldrp.12048

Kroll, B.M. (2005). Arguing differently. *Pedagogy, 5*(1), 37-60.

Kuhn, D. (1991). *The skills of argument*. Cambridge, England: Cambridge University Press. https://doi-org.proxy-um.researchport.umd.edu/10.1017/CBO9780511571350

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

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science, 22*(4), 545-552. https://doi-org.proxy-um.researchport.umd.edu/10.1177/0956797611402512

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*(3), 309–328.

Lambert, R., & Sugita, T. (2016). Increasing engagement of students with learning disabilities in mathematical problem-solving and discussion. *Support for Learning*, *31*(4), 347-366. https://doi-org.proxy-um.researchport.umd.edu/10.1111/1467-9604.12142

Langley, G. J., Moen, R. D., Nolan, K. M., Nolan, T. W., Norman, C. L., & Provost, L. P. (2009). *The improvement guide*. San Francisco, CA: Jossey-Bass.

*Larraín, A., Freire, P., Grau, V., Lopez, P., Salvat, I., Silva, M., & Gastellu, V. (2018). The effect of peer-group argumentative dialogue on delayed gains in scientific content knowledge. In V. Grau & D. Whitebread (Eds.), Relationships between classroom dialogue and support for metacognitive, self-regulatory development in educational contexts. *New Directions for Child and Adolescent Development, 162*, 67–87. https://doi-org.proxy-um.researchport.umd.edu/10.1002/cad.20263

*Larraín, A., Singer, V., Strasser, K., Howe, C., López, P., Pinochet, J., & Villavicencio, C. (2021). Argumentation skills mediate the effect of peer argumentation on content knowledge in middle-school students. *Journal of Educational Psychology, 113*(4), 736-753. https://doi-org.proxy-um.researchport.umd.edu/10.1037/edu0000619

Larson, B.E. (2021). Guiding principles for using classroom discussion. In J. C. Lo, *Making classroom discussions work: Methods for quality dialogue in the social studies series* (11-26). Teachers College Press.

Lehrer, R. (2009). Designing to develop disciplinary dispositions: Modeling natural systems. *American Psychologist, 64*(8), 759–771.

Levin, D., De La Paz, S., Lee, Y., & Nadal, E. (2021). Supporting students' construction and critique of scientific explanations and arguments: A cognitive apprenticeship for teaching scientific writing in middle school inclusion classrooms. *Learning Disabilities: A Multidisciplinary Journal, 26*(1), 58-72. *https://doi.org/10.1177/073194872211119365*

Louick, R., & Wang, M. (2021). Classroom discourse and disability: Interactional opportunities for development of self-determination beliefs. *The Journal of Educational Research, 114*(1), 52-63. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00220671.2021.1872475

Matos, F. (2021). Collaborative writing as a bridge from peer discourse to individual argumentative writing. *Reading and Writing, 34*(5), 1321-1342. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11145-020-10117-2

*MacArthur, C. A., Ferretti, R., & Okolo, C. (2002). On defending controversial viewpoints: Debates of sixth graders about the desirability of early 20th century American immigration. *Learning Disabilities Research and Practice*, *17*, 160–172.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & research in social education*, *46*(2), 165-193. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00933104.2017.1416320

McLeskey, J. (2017). High-leverage practices in special education. Arlington, VA: Council for Exceptional Children.

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

Mercer, N. (2000). *Words and minds: How we use language to think together* (1st ed.). Taylor & Francis Group.

Mercer, N. (2004). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*, 57–74. https://doi-org.proxy-um.researchport.umd.edu/10.1017/S0140525X10000968

Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education. Revised and Expanded from" Case Study Research in Education."*. Jossey-Bass Publishers, 350 Sansome St, San Francisco, CA 94104.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. sage.

Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of adolescent & adult literacy*, *52*(2), 96-107. https://doi-org.proxy-um.researchport.umd.edu/10.1598/JAAL.52.2.1

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. https://doi-org.proxy-um.researchport.umd.edu/10.3102/0002831208319733

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

Monte-Sano, C., De La Paz, S., & Felton, M. (2014). Implementing a disciplinary-literacy curriculum for US history: Learning from expert middle school teachers in diverse classrooms. *Journal of Curriculum Studies, 46*(4), 540–575. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00220272.2014.904444

Monte-Sano, C., De La Paz, S., & Felton, M. (2014). *Reading, thinking, and writing about history: Teaching argument writing to diverse learners in the Common Core classroom, grades 6-12*. Teachers College Press.

Monte-Sano, C., De La Paz, S., Felton, M., Piantedosi, K. W., Yee, L. S., & Carey, R. L. (2017). Learning to teach disciplinary literacy across diverse eighth-grade

history classrooms within a district-university partnership. *Teacher Education Quarterly*, *44*(4), 98-124.

Monte-Sano, C., Schleppegrell, M., Sun, S., Wu, J., & Kabat, J. (2021). Discussion in diverse middle school social studies classrooms: Promoting all students' participation in the disciplinary work of inquiry. *Teachers College Record*, *123*(10), 142-184. https://doi-org.proxy-um.researchport.umd.edu/10.1177/01614681211058971

*Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2007). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behavior*, *23*(3), 1072-1088. https://doi-org.proxy-um.researchport.umd.edu/10.1016/j.chb.2006.10.003

Murphy, P. K., Wilkinson, I. A., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of educational psychology*, *101*(3), 740-764. https://doi-org.proxy-um.researchport.umd.edu/10.1037/a0015576

Murphy, P. K., Firetto, C. M., & Greene, J. A. (2017). Enriching students' scientific thinking through relational reasoning: Seeking evidence in texts, tasks, and talk. *Educational Psychology Review, 29*(1), 105–117. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s10648-016-9387-x

*Murphy, P. K., Greene, J. A., Allen, E., Baszczewski, S., Swearingen, A., Wei, L., & Butler, A. M. (2018). Fostering high school students' conceptual understanding and argumentation performance in science through Quality

Talk discussions. *Science Education*, *102*(6), 1239–1264. https://doi-org.proxy-um.researchport.umd.edu/10.1002/sce.21471

Nathan, E., & Lee, C. K. E. (2004). Using structured academic controversies in the social studies classroom. *Teaching and Learning, 25*(2), 171-188.

National Center for Education Statistics. (2024). Students with Disabilities. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved May 20, 2024, from https://nces.ed.gov/programs/coe/indicator/cgg.

National Center for Education Statistics. (2024). English Learners in Public Schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved May 20, 2024, from https://nces.ed.gov/programs/coe/indicator/cgf.

National Center for Education Statistics. (2024). Racial/Ethnic Enrollment in Public Schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved May 20, 2024, from https://nces.ed.gov/programs/coe/indicator/cge.

National Council for the Social Studies. (2013). College, career, and civic life (C3) framework for social studies state standards. NCSS.

National Research Council. (2013). *Next Generation Science Standards: For states, by states.* National Academies Press.

National Research Council. (2014). *Literacy for science: Exploring the intersection of the Next Generation Science Standards and Common Core for ELA standards: A workshop summary.* National Academies Press.

Newell, G. E. (2006). Writing to learn. *Handbook of Writing Research*, 235-247.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220.

Nokes, J. D. (2017). Exploring patterns of historical thinking through eighth-grade students' argumentative writing. *Journal of Writing Research*, *8*(3), 437-467.

Nokes, J. D., Dole, J. A., & Hacker, D. J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology, 99*(3), 492–504. https://doi-org.proxy-um.researchport.umd.edu/https://content.apa.org/journals/edu/99/3/492

Nokes, J. D., & De La Paz, S. (2018). Writing and argumentation in history education. *The Wiley international handbook of history teaching and learning*, 551-578.

Nokes, J. D., & Kesler-Lund, A. (2019). Historians' social literacies: How historians collaborate and write. *The History Teacher, 52*(3), 369-410. https://doi-org.proxy-um.researchport.umd.edu/https://www.societyforhistoryeducation.org/pdfs/M19_Nokes_and_Kesler-Lund.pdf

Nokes, J. D., & De La Paz, S. (2023). Historical Argumentation: Watching Historians and Teaching Youth. *Written Communication, 40*(2), 333-372. https://doi-org.proxy-um.researchport.umd.edu/10.1177/07410883221148679

Nussbaum, E. M. (2021). Critical integrative argumentation: Toward complexity in students’ thinking. *Educational Psychologist*, *56*(1), 1-17. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00461520.2020.1845173

*Nussbaum, E. M., & Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. *Journal of the Learning Sciences, 20*(3), 443-488. https://doi-org.proxy-um.researchport.umd.edu/10.1080/10508406.2011.564567

Nussbaum, M., Dove, I., & Putney, L. (2023). Bridging dialogic pedagogy and argumentation theory through critical questions. *Dialogic Pedagogy: An International Online Journal*, *11*(3), A7-A25.

Nystrand, M., Gamoran, A., Kachur, R., & Prendergast, C. (1997). *Opening dialogue* (pp. 30-61). New York: Teachers College Press.

O'Connor, C., & Snow, C. (2018). Classroom discourse: What do we need to know for research and for practice? In M. F. Schober, D.N. Rapp, & M. A. Britt (Eds.), *The Routledge Handbook of Discourse Processes* (pp. 315–342). Routledge. https://doi-org.proxy-um.researchport.umd.edu/10.4324/9781315687384-17

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J. Grimshaw, J.M., Hrobjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *British Medical Journal, 372*(71), 1-9. https://doi-org.proxy-um.researchport.umd.edu/10.1371/journal.pmed.1003583

Parker, W. C. (2021). Structured academic controversy: What it can be. In J. C. Lo, *Making classroom discussions work: Methods for quality dialogue in the social studies series* (73-89). Teachers College Press.

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

Reisman, A. (2015). Entering the historical problem space: Whole-class text-based discussion in history class. *Teachers College Record, 117*(2), 1–44. https://doi-org.proxy-um.researchport.umd.edu/https://www.tcrecord.org/Content.asp?ContentId=17783

Reisman, A., & Wineburg, S. (2008). Teaching the skill of contextualizing in history. *The Social Studies, 99,* 202–207. https://doi-org.proxy-um.researchport.umd.edu/10.3200/TSSS.99.5.202-207

Reisman, A., Cipparone, P., Jay, L., Monte-Sano, C., Kavanagh, S. S., McGrew, S., & Fogo, B. (2019). Evidence of emergent practice: Teacher candidates facilitating historical discussions in their field placements. *Teaching and teacher education*, *80*, 145. https://doi-org.proxy-um.researchport.umd.edu/10.1016/j.tate.2018.12.014

*Relyea, J. E., Zhang, J., Wong, S. S., Samuelson, C., & Wui, M. G. L. (2022). Academic vocabulary instruction and socio-scientific issue discussion in urban sixth-grade science classrooms. *The Journal of Educational Research,*

115(1), 37–50. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00220671.2021.2022584

Reynolds, T. (2018). Understanding and embracing contradictions: An exploration of high school English teachers' beliefs about whole-class discussion. *Language and Education, 32*(4), 371-385. https://doi-org.proxy-um.researchport.umd.edu/10.1080/09500782.2018.1445265

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education, 92*(3), 499–525. https://doi-org.proxy-um.researchport.umd.edu/10.1002/sce.20264

Saddler, B., Asaro-Saddler, K., Moeyaert, M., & Ellis-Robinson, T. (2017). Effects of a summarizing strategy on written summaries of children with emotional and behavioral disorders. *Remedial and Special Education*, *38*(2), 87-97. https://doi-org.proxy-um.researchport.umd.edu/10.1177/0741932516669051

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*(5), 513–536. https://doi-org.proxy-um.researchport.umd.edu/10.1002/tea.20009

Sadler, T. D., & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching, 42*(1), 112-138. https://doi-org.proxy-um.researchport.umd.edu/10.1002/tea.20042

Sandoval, W. A., Enyedy, N., Redman, E. H., & Xiao, S. (2019). Organising a culture of argumentation in elementary science. *International Journal of Science*

Education, 41(13), 1848–1869. https://doi-org.proxy-um.researchport.umd.edu/10.1080/09500693.2019.1641856

Schwarz, B.B. (2009). Argumentation and Learning. In: Muller Mirza, N., Perret-Clermont, A.N. (eds) *Argumentation and Education*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-98125-3_4

Seixas, P. (2006). Benchmarks of historical thinking: A framework for assessment in Canada. *The Center for the study of Historical Consciousness. Recuperado el*, *16*, 1-12.

Shanahan, C. (2014). Disciplinary comprehension. In *Handbook of research on reading comprehension* (pp. 264-284). Routledge.

Shi, Y., Matos, F., & Kuhn, D. (2019). Dialog as a bridge to argumentative writing. *Journal of Writing Research, 11*(1), 107-129. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s11145-020-10117-2

Swanson, E., Stevens, E. A., & Wexler, J. (2019). Engaging students with disabilities in text-based discussions: Guidance for general education social studies classrooms. *TEACHING Exceptional Children*, *51*(4), 305-312. https://doi-org.proxy-um.researchport.umd.edu/10.1177/0040059919826030

TeachingWorks (2023). *Leading a discussion*. TeachingWorks Resource Library. https://library.teachingworks.org/curriculum-resources/teaching-practices/leading-a-discussion/

The Teacher Toolkit. (n.d.). *Turn and talk*. Retrieved from http://www.theteachertoolkit.com/index.php/tool/ turn-and-talk

van Drie, J., & van de Ven, P. H. (2017). Moving ideas: An exploration of students' use of dialogue for writing in history. *Language and education*, *31*(6), 526-542. https://doi-org.proxy-um.researchport.umd.edu/10.1080/09500782.2017.1326504

*van Drie, J., van Boxtel, C., Erkens, G., & Kanselaar, G. (2005). Using representational tools to support historical reasoning in computer-supported collaborative learning. *Technology, Pedagogy and Education*, *14*(1), 25–41. https://doi-org.proxy-um.researchport.umd.edu/10.1080/14759390500200191

Voss, J. F., &Wiley, J. (2006). Expertise in history. In Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 569–584). Cambridge, UK: Cambridge University Press.

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

Walqui, A., & Heritage, M. (2018). Meaningful classroom talk: Supporting English learners' oral language development. *American Educator*, *42*(3), 18-23.

*Wanzek, J., Kent, S. C., Vaughn, S., Swanson, E. A., Roberts, G., & Haynes, M. (2015). Implementing team-based learning in middle school social studies classes. *The Journal of Educational Research, 108*(4), 331-344. https://doi-org.proxy-um.researchport.umd.edu/10.1080/00220671.2014.893224

Waring, S. M. (Ed.). (2024). *Teaching with primary sources for cultural understanding, civic mindedness, and democracy*. Teachers College Press. https://doi-org.proxy-

um.researchport.umd.edu/https://www.tcpress.com/teaching-with-primary-sources-for-cultural-understanding-civic-mindedness-and-democracy-9780807769713

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.

Wineburg, S. (1991a). 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.

Wineburg, S. (1991b). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal, 28*(3), 495–519.

Wineburg, S. S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*. Philadelphia, PA: Temple University Press.

Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record, 121*(11), 1-40. https://doi-org.proxy-um.researchport.umd.edu/https://www.tcrecord.org/content.asp?contentid=22806

*Wissinger, D. R., & De La Paz, S. (2016). Effects of critical discussions on middle school students' written historical arguments. *Journal of Educational Psychology, 108*(1), 43-59. https://doi-org.proxy-um.researchport.umd.edu/10.1037/edu0000043

Wissinger, D. R., De La Paz, S., & Jackson, C. (2021). The effects of historical reading and writing strategy instruction with fourth-through sixth-grade students. *Journal of Educational Psychology, 113*(1), 49-67. doi:10.1037/edu0000463

Yin, R. (2018). *Case study research and applications designs and methods*. Sage.

*Zhang, J., Wui, M. G. L., Nam, R., Relyea, J. E., & Wong, S. S. (2023). Improving argumentative writing of sixth-grade adolescents through dialogic inquiry of socioscientific issues. *Journal of Writing Research 14*(3), 375-419.

Zolkower, B., & Shreyar, S. (2007). A teacher's mediation of a thinking-aloud discussion in a 6th grade mathematics classroom. *Educational Studies in Mathematics, 65*, 177-202. https://doi-org.proxy-um.researchport.umd.edu/10.1007/s10649-006-9046-0