UC Irvine
UC Irvine Electronic Theses and Dissertations

Title

Academic Language Features in the Argumentative Writing of Second through Eighth Grade Students

Permalink

https://escholarship.org/uc/item/95p4437k

Author

Taylor, Karen Suzanne

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Academic Language Features in the Argumentative Writing
of Second through Eighth Grade Students

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

Karen Suzanne Taylor

Dissertation Committee:
Professor Carol Connor, Chair
Associate Professor Joshua Lawrence
Professor Young-Suk Kim
Associate Professor Penelope Collins
Professor Lisa Pearl

2019

DEDICATION

To

My family and friends

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my two academic advisors over the course of my time at UC Irvine, Dr. Joshua Lawrence and Dr. Carol Connor. Your mentorship was a gift I will always treasure. Your expertise, guidance, availability, and kindness are just a few of the ways to describe your consistent support during the doctoral program and dissertation process. Thank you also to Dr. Lawrence for the incredible opportunity of inviting me to work with and learn from you at UC Irvine. And thank you to Dr. Connor for the invaluable opportunity to work with and learn from you in the context of your Individualizing Student Instruction lab. I really appreciated both of your investments into my growth as a researcher, and besides that, you are wonderful people I enjoyed getting to know.

My sincere gratitude also goes to my dissertation committee members, Dr. Young-Suk Kim, Dr. Penelope Collins, and Dr. Lisa Pearl. Your expertise and supportive suggestions have strengthened my dissertation, and I felt very lucky to have you on my committee. Thank you very much for your time and effort to help make this dissertation possible.

I also thank Dr. Catherine Snow, a wonderful mentor I had from Harvard Graduate School of Education. You have provided so many opportunities for me to learn and grow as a researcher. Thank you for each one, from the immersive WordGen research practicum as a masters student, to the ongoing research and collaboration with you as a doctoral student, as well as the chance to spend this past fall at HGSE as a visiting graduate student researcher. I am deeply grateful for each of these opportunities, and for your time and expertise you have so graciously shared.

I thank each faculty and staff member at UC Irvine School of Education I had the chance to learn from and who supported us throughout the doctoral program. Thank you to my inspiring cohort mates and lab mates in the ISI lab. You all have made these past five years a truly wonderful experience. Many thanks also to Sarah Ingebrand for assistance with the methods in Chapter 2, and to Chelsea Parlett-Pelleriti for assistance with the methods in Chapter 3. I thank the many undergraduate research assistants for their dedication to helping with essay coding: Ashley Chang, Alice Foy, Montana Henry, Robin Jang, Kasey Kim, Odalys Orellana, Ante Xu, and Maria Yanez (Chapter 2); Kasey Kim (Chapter 3); and Breanna Briggs, Andrea Byng, Ragiah El-Shantaly, and Samaan Nur (Chapter 4).

Lastly and most importantly, I thank my family and friends, to whom I have dedicated this dissertation. You have been there for all of my mountain climbing endeavors, and I am especially full of love and gratitude for your support during these five years. Your love, laughs, and encouragement helped sustain my endurance. To my parents, your faithful love throughout my life means the world to me. Thank you.

CURRICULUM VITAE

KAREN S. TAYLOR

Updated: August 2019
Google Scholar Profile: https://scholar.google.com/citations?hl=en&user=AIe_1YwAAAAJ

EDUCATION

Ph.D., Education, *Language, Literacy, and Technology.*
    University of California, Irvine, August 2019.
Ed.M., Education, *Learning and Teaching.*
    Harvard Graduate School of Education, May 2013.
M.A., Education, *Education and Psychology.*
    Pepperdine Graduate School of Education and Psychology, July 2011.
B.A., Liberal Arts, *Emphasis Psychology.*
    Pepperdine University, April 1995.

RESEARCH EXPERIENCE

August 2018 – December 2018	*Visiting Doctoral Researcher*, Harvard Graduate School of Education. Word Generation academic vocabulary program. (PI: Catherine Snow).
August 2018 – December 2018	*Consulting Researcher,* Harvard Graduate School of Education. Model of Reading Engagement (MORE) project. (PI: James Kim).
June 2016 – September 2017	*Research Assistant,* University of California, Irvine. Individualizing Student Instruction Lab: Optimizing Learning Opportunities for Students project. (PIs: Carol Connor and Deborah Vandell).
January 2016 – June 2016	*Research Assistant,* University of California, Irvine. Learning to Learn from Mathematics Teaching project. (PI: Rosella Santagata).

RESEARCH INTERESTS

Adolescent literacy, Argumentative writing, Academic language, Reasoning, Literacy learning and teaching, Educational technology, Teacher education and professional development

JOURNAL ARTICLES

Taylor, K. S., Lawrence, J. F., Connor, C. M., & Snow, C. E. (2019). Cognitive and linguistic features of adolescent argumentative writing: Do connectives signal more complex reasoning? *Reading and Writing, 32*(4), 983-1007.

Connor, C. M., Day, S. L., Zargar, E., Wood, T. S., Taylor, K. S., Jones, M. R., & Hwang, J. K. (2019). Building word knowledge, learning strategies, and metacognition with the Word-Knowledge e-Book. *Computers & Education, 128,* 284-311.

Santagata, R., & Taylor, K. (2018). Novice teachers' use of student thinking and learning as evidence of teaching effectiveness: A longitudinal study of video-enhanced teacher preparation. *Contemporary Issues in Technology and Teacher Education, 18*(1), 11-28.

Lin, A. R., Lawrence, J. F., Snow, C. E., & Taylor, K. S. (2016). Assessing adolescents' communicative self-efficacy to discuss controversial issues: Findings from a randomized study of the Word Generation program. *Theory & Research in Social Education, 44*(3), 316-343.

UNIVERSITY TEACHING EXPERIENCE (2014-2019)
Teaching Assistant, **University of California, Irvine**

Undergraduate Courses:	Complex Pedagogical Design
	Classroom Interactions
	Theories of Development and Learning in Education

Masters Courses:	Reading and Writing in Secondary Classrooms
	Outcomes on Schooling and Student Assessment
	Learning from Teaching
	Identifying Dilemmas of Practice

K-6 TEACHING EXPERIENCE (1995 – 2014)
Oak Hills Elementary School, Oak Park Unified School District, Oak Park, CA
August 1995 – June 1996

Red Oak Elementary School, Oak Park Unified School District, Oak Park, CA
August 1996 – June 1997

Montecito Union School, Montecito Union School District, Santa Barbara, CA
August 1997 – June 2014
(At Harvard Graduate School of Education during 2012 – 2013 school year)

PROFESSIONAL AFFILIATIONS
Society for the Scientific Study of Reading (SSSR)
Society for Research in Child Development (SRCD)
Society for Research on Educational Effectiveness (SREE)

ABSTRACT OF THE DISSERTATION

Academic Language Features in the Argumentative Writing

of Second through Eighth Grade Students

By

Karen Suzanne Taylor

Doctor of Philosophy in Education

University of California, Irvine, 2019

Professor Carol M. Connor, Chair

Three studies within this dissertation investigated 2nd through 8th grade students' productive use of academic language in their argumentative essays (N=741 essays). Despite the prominence of argumentative writing in the Common Core State Standards (Common Core State Standards Initiative, 2010), little research has been conducted on specific reasoning strategies and the lexical precision to signal arguments in elementary and early adolescent writing. Using quantitative and natural language processing methods, I examined argumentative reasoning and connective types (and their relations) in students' independently written essays.

In Study 1, I analyzed 385 argumentative essays written by a sample of 385 2nd and 3rd grade students (one essay per student). All students wrote about the same topic, and essays were coded for arguments, argument sophistication, and connectives. Controlling for grade and essay length, reading comprehension had a significant positive association with argument sophistication in the essays. When connective types were added to the model, and again

controlling for essay length and grade, reading comprehension and adversative connectives both had significant positive associations with argument sophistication in the essays.

In Study 2, I analyzed 198 argumentative essays written by a sample of 66 4th and 5th grade students (3 essays per student). Fourth and 5th grade students wrote about different topics, and essays were coded for arguments, argument sophistication, connectives, and connectives sophistication. Controlling for topic, essay length, and ethnicity variables, reading comprehension had a significant positive association with argument sophistication. When a measure of connectives sophistication was added to the model, reading comprehension no longer had a significant association with argument sophistication, and connectives sophistication also did not have a significant association with argument sophistication. However, adding connective types to the model explained an additional 7% of the variance, and adversative connectives had a significant positive association with argument sophistication in the essays.

In Study 3, I analyzed 158 argumentative essays written by a sample of 40 6th through 8th grade students (3 or 4 essays per student). Students in all grades wrote about the same four topics. Essays were coded for arguments and argument sophistication, and a natural language processing tool (Tool for the Automated Analysis of Cohesion, TAACO; Crossley, Kyle, & McNamara, 2016) calculated categories of connectives. Controlling for length and topic, adversative connectives had a significant positive association with argument sophistication.

Findings from this dissertation underscore that 2nd through 8th grade students marshal academic language in their argumentative writing in specific ways that are unlikely to be elucidated in holistic measures of writing quality. Findings also indicate that certain linguistic features (i.e., adversative connectives) are related to more complex arguments in these students' essays.

References

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

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods, 48,* 1227-1237. doi:10.3758/s13428-015-0651-7

CHAPTER 1

Introduction

Argumentative writing, though featured in the Common Core State Standards (CCSSI, 2010), is a difficult writing genre to learn (Ferretti & Fan, 2016) and one in which a majority of students in the US have yet to show proficiency. Indeed, only 25% of students' argumentative essays were considered competent in the 2012 NAEP Writing Report Card (National Center for Educational Statistics, 2012). The nature of reasoning inherent within skilled argumentative writing is challenging, as it involves complex cognitive processes such as problem solving (Bereiter & Scardamalia, 1987), as well as counterfactual reasoning or considering alternative viewpoints (Kuhn, 1992; Kuhn & Udell, 2007; van Eemeren, Grootendorst, & Henkemans, 1996; Walton, 1989) and possibly even the more sophisticated, integrative reasoning (Kuhn & Crowell, 2011; Nussbaum, 2011). In addition, argumentative writing is a challenging task linguistically, since it incorporates features of an academic register of the English language (i.e., academic language) such as syntactic complexity (e.g., dense clause packaging) and lexical precision (e.g., connectives to signal textual relations) (Snow & Uccelli, 2009; Uccelli et al., 2015a, b).

Research on elementary school students' writing has historically been less of a focus than other subjects such as reading (Graham, McKeown, Kiuhara, & Harris, 2012; National Commission on Writing, 2004), with argumentative writing also being understudied (Anderson, Torgerson, Low, & McGuinn, 2009; Newell, Beach, Smith, & VanDerHeide, 2011). However, research on argumentative writing in general has begun to receive greater attention in recent years (Ferretti & Fan, 2016; Allen, Likens, & McNamara, 2019). Some have examined a robust set of research on adolescent writing instruction (i.e., grades 4-12; Graham & Perin, 2007).

Researchers also call for greater attention to research on students' argumentative writing over time (Newell, Beach, Smith, & VanDerHeide, 2011), as well as the development of more researchers with an extended program of research centered on adolescent writing (Graham & Perin, 2007). In addition, researchers point to "unique synergies" when research on argumentative writing takes an interdisciplinary approach, drawing on linguistic, sociocultural, and cognitive perspectives (Ferretti & Graham, 2019, p. 1353). It is an interdisciplinary approach to this important area of argumentative writing, through textual analysis of children's and early adolescents' single-occasion as well as multi-occasion writing samples, that is the goal of this work.

The studies in this dissertation are, first of all, situated within the field of educational linguistics (Uccelli & Snow, 2008), as they are concerned with the language children and adolescents' produce in their argumentative writing samples. Specifically, these studies draw upon conceptualizations of an academic register of English often referred to as *academic English* or *academic language* (Bailey, 2007; Scarcella, 2003; Schleppegrell, 2004; Snow & Uccelli, 2009; Uccelli et al., 2015a, b). Given the more challenging linguistic features of texts students will be tasked with comprehending and producing in elementary school and middle school, research on students' productive capabilities with academic language is important for research and practice (Snow & Uccelli, 2009; Fillmore & Snow, 2000). In particular, this work presents a conceptual model of language for sub-goals of argumentative writing that is primarily guided by Snow and Uccelli's (2009) pragmatics-based approach to academic language, in which they offer a series of nested figures representing different communicative purposes as an initial attempt to answer the question: "What are the communicative challenges to which the features of academic language are meant to respond?" (p. 122). The figures aim to move beyond inventories of

academic language so as to help advance an overall rationale for using academic language. For example, specific purposes are shown nested within levels of language; that is, the clause level (which incorporates the lexical level), the discourse level, the message, and finally, the self and audience. Together, these comprise a straightforward display coordinating different levels of academic language. Applying this type of visualization to the present studies, I will show the different levels of language – and where the language functions as reasoning – within a similar representation.

Figure 1.1 Model of language for argumentative reasoning micro genre

Representing the macro message

- Text type: Argumentative writing/ Macro genre: Persuasive essay (Genre theory; SFL; Writing as product)
- Argumentative writing as a problem solving and goal-directed activity (Cognitive; Writing as process)
- Argument Schema Theory (General understanding of argument) (Cognitive and sociocultural; Writing as a product of thinking, following the process of dialogic argumentation)

Representing the micro message

- **Micro genre: Argument/counterargument (Genre theory; SFL; Writing as product)**
- Setting relevant content and rhetorical subgoals for building toward an overall argument (Cognitive; Writing as process)
- Using reasoning to construct/represent complex ideas (e.g., counterfactual reasoning) (Writing as process and product)
- Using discourse markers/connectives to signal textual relations (i.e., additive, adversative, causal, temporal relations) (Linguistic; Writing as product)

Producing language

- Choosing precise lexicon

- Selecting grammatical structures

Surface code

Word

Situation Model

Discourse level

Textbase

Sentence (T-unit) level

Reading Comprehension

3

To contextualize the role of reasoning within the overall task of argumentative writing, aspects of cognitive models of writing have influenced the different levels of language represented in the model. Whereas these cognitive models describe multilevel language *processes* for writing (and relations among them), the scope of the present studies distill aspects of them into a generalized map for language of argumentation, with language to describe complex reasoning (e.g., argument/counter-argument) being the grain-sized outcome. First, in Berninger and colleagues' (2002) theoretical model of the writing system, they specify the components of *text generation* and *language representation* both as multilevel (i.e., word-sentence-discourse). This model, as well as Juel, Griffith, and Gough's (1986) model were known as the simple view of writing, in which writing is the product of transcription and ideation or text generation. The simple view of writing did not account for foundational or higher order cognitive skills, such as reasoning or perspective taking. However, even this baseline of empirical evidence for text generation and language representation as this multilevel, word-sentence-discourse level, process, is helpful to draw from in presenting a representation of argumentative writing specifically. Later theoretical models have expanded on the simple view of writing, in order to describe the component skills (and relations among them) of transcription and text generation. Most recently, Kim and Schatschneider's (2017) Direct and Indirect Effects Model of Developmental Writing (DIEW) found that discourse-level oral language skills were directly related to writing quality for young children, and the component skills of discourse-level oral language were indirectly related. That is, the relation of foundational oral language skills of vocabulary and grammar and the higher order cognitive abilities of inference and perspective taking to writing were mediated by discourse-level oral language skills. Transcription skills of spelling and writing fluency were also directly related to writing (although to a lesser extent),

and working memory was found to be a foundational cognitive skill for all of the component discourse-level oral language skills and transcription skills. The contribution of the DIEW model to research on children's writing is a notable expansion to earlier conceptualizations of writing, in that components of discourse-level oral language are operationalized, and their direct and indirect relations to writing are specified. Although the model was fit for first grade narrative writing quality, their findings related to the role of perspective taking in particular lend support to the importance of examining higher-order cognitive skills such as reasoning (e.g., considering multiple viewpoints or alternative perspectives) represented in children's and early adolescents' argumentative writing. Although perspective taking is a component skill of discourse-level oral language, in the above conceptual model of language for argumentative reasoning, reasoning is operationalized at the sentence-level, because the unit of analysis for identifying particular types of reasoning in children's and adolescents' writing samples is essentially at the sentence level (i.e., reasoning was coded at the T-unit, or idea-unit level).

An important foundation for operationalizing the component skills of discourse-level oral language in the DIEW model was Kim's (2016) conceptual model of discourse-level oral language itself, which incorporates construction-integration models (e.g., Kintsch, 1988) of discourse comprehension and production represented as three levels: the surface code of foundational language and cognitive skills; above that, the textbase including higher order cognitive skills; and at the top, the situation model encompassing discourse comprehension and production. This is a further influence of a cognitive perspective on the model of language for argumentative reasoning, as seen in the aspects of language also mapped onto these levels. Specifically, we could expect that the components of lexical precision (e.g., academic vocabulary and connectives) might fall under the surface code of academic language for the genre of

argumentative writing. Next, we could assume that higher-order cognitive skills such as constructing complex reasoning (e.g., argument and counter argument) might fit within the textbase. For the highest level of the situation model, we could see this as the overall structure of the argumentative piece of writing, which lies in the macro-genre of the persuasive essay for the children and adolescents in the present studies.

The present studies also draw from interdisciplinary lines of research that incorporate cognitive and sociocultural perspectives for purposes of argumentation and argumentative writing. In these education interventions, dialogic argumentation often precedes occasions for argumentative writing with the goal of fostering individual argumentative reasoning (e.g., Kuhn & Crowell, 2011; Reznitskaya et al., 2001, 2002). However, as Ferretti and Fan (2016) point out, the process of writing itself can be thought of as a social act: "Argumentative writing is always dialogic because it involves communication between real or imagined interlocutors" (p. 302). Reznitskaya and Anderson's (2002) theoretical model, called Argument Schema Theory (AST), is represented in the language for argumentative reasoning model as the overarching, highest level which accounts for the discourse level and the overall genre (i.e., *macro genre,* Paltridge, 2014; Woodward-Kron, 2005) of a persuasive essay. Two of the studies in this dissertation investigated writing samples that took place in the context of an academic vocabulary program, Word Generation, that was driven by a dialogic teaching approach toward integrating academic language, perspective taking, and complex or argumentative reasoning hypothesized to impact deep reading comprehension (Donovan, Snow, & Daro, 2013; LaRusso et al., 2016). Although the outcome in these studies is not *overall* argumentative structure of the students' essays, it is still the broader frame in which students could demonstrate knowledge of a generalized argument structure, or whether they had developed an *argument schema*. It is through

"socialization into the discourse of reasoned argumentation" (Reznitskaya et al., 2012, p. 290) that dialogic teaching approaches, such as the Word Generation curriculum, are hypothesized to support the development of students' individual competencies in argumentation.

At the center of the model of language for argumentative reasoning lies the outcome of the studies in this dissertation, which is the micro-genre (Woodward-Kron, 2005) of argument and counterargument. Because of the affordances of quantitative methods of textual analysis, I undertook a more fine-grained approach to analyzing students' productive uses of academic language and investigated what would be signposts or components toward an argument schema or toward the overall genre of the persuasive essay: the individual arguments inside of the essay as representations of students' reasoning. When argumentative essays are assessed with a holistic measure, it would make sense that the rubric would capture the sophistication of students' reasoning within the essay to some extent. What is less clear with one-score measures, however, is the extent to which students deployed individual types of arguments (displaying different levels of reasoning) in their essays. In my dissertation studies, I utilize an approach to measuring argumentation in students' writing that aligns with Reznitskaya's (2009) analytic method affording a means to "separately evaluate each proposition made by students in their essays" (p. 220). Thus, I can quantify types and amounts of specific arguments within and across students' essays, rather than assigning one, qualitative score to each essay. As an initial exploration of students' reasoning toward the development of an argument schema, I analyze these facets of their writing, while development or utilization of composite measures capturing an overall argument schema or argument structure within their persuasive essays is slated for future research.

To further conceptualize students' argumentative reasoning in their writing, I also consider connections between reading and writing (Fitzgerald & Shanahan, 2000; Graham et al., 2018). Research has shown that writing interventions can impact reading (e.g., Graham & Hebert, 2011a, b), and a recent, robust meta-analysis (Graham et al., 2018) has demonstrated that reading interventions can enhance students' writing. Two of the dissertation studies were able to incorporate a measure of reading comprehension as an independent variable in the regression models with the highest level of reasoning as the outcome. I anticipate that reading comprehension may indeed be associated with the level of argument sophistication in students' essays, based on models of reading and writing described earlier. If reading comprehension involves an understanding of text at all three levels, surface-code, textbase, and situation model, then reading comprehension should be related to a feature of writing that occurs at the textbase level, which is the level of sophistication of reasoning. If an association is found between reading comprehension and the reasoning outcome in the writing, other explanations should be considered as well. While such an association could be due to shared reading and writing processes, it may also have to do with a variable not yet considered, which is that of the home literacy environment. Research has shown that home factors such as shared reading can influence literacy skills (e.g., Bus, van IJzendoorn, & Pellegrini, 1995; Payne, Whitehurst, & Angell, 1994). Further research has demonstrated a host of literacy practices in the home literacy environment that are related to emergent literacy (e.g., Whitehurst & Lonigan, 1998; Lonigan, Schatschneider, Westberg, & the National Early Literacy Panel, 2008; Storch & Whitehurst, 2001). It is important to keep in mind other factors that also are likely to contribute to a relation between reading ability and writing outcomes.

Overview of the Dissertation

This dissertation examines academic language use in the argumentative writing of three analysis samples of 2nd through 8th grade students (2nd and 3rd; 4th and 5th; and 6th through 8th grades). There are two aims for this proposed study. The first aim is to investigate two features of academic language in the students' writing: (1) specific argumentative moves that display various levels of sophistication in reasoning, and (2) the major classes of connectives, which signal different cohesive functions within a text (additive, causal, adversative, and temporal; Halliday & Hasan, 1976). The second aim is to analyze whether there is a relation among connectives and argument sophistication. Given the importance of learning how to take a reasoned point of view and defend it in the 21st Century, our limited understanding of how children and young adolescents develop these skills points to a vital area for research in educational linguistics. Studying the features of academic language and complex reasoning produced by different samples and ages of students, and within different contexts, will provide insights into features of academic writing that students are capable of employing on their own. Having a better understanding of these facets from a descriptive standpoint, as well as the relations among them, can then begin to inform interventions and/or curricula with the aim of improving elementary and middle school students' writing outcomes.

Overview of the studies

Study 1. Argumentative reasoning and connectives in second- and third-grade students' persuasive essays

In the first study of my dissertation, I examine a sample of 385 persuasive essays produced by second and third grade students (1 essay each) as a fall, baseline writing measure in the context of a larger study of a literacy intervention. Students were from two schools in the

Southwest US that were considered high poverty schools. The writing prompt was a relatively simple, on-demand writing prompt on a common, everyday topic. In this way, it was hoped that background knowledge would not be an issue, since the writing was not completed within the context of a particular content area or unit of study. The outcome was argument sophistication in the essays, assessed by the highest argument used according to an argumentative reasoning coding scheme (Kuhn & Crowell, 2011). A general reading comprehension measure (Gates-MacGinitie Reading Test [MacGinitie, MacGinitie, Maria, & Dreyer, 2006]) was included to investigate the role of reading comprehension. Argument types in the essays were assessed, as well as students' uses of major classes of connectives (Halliday & Hasan, 1976). Frequencies of connectives were independent variables. To address my research questions, I first utilized multiple procedures for descriptive statistics to show how students used these cognitive and linguistic features in their writing. I also employed linear regression to test for relations among variables with overall argument sophistication in the essays. The results from Study 1 describe detailed ways these children used two domains of academic language in their writing (argumentative reasoning and connecting ideas logically), along with the relation of reading comprehension and specific linguistic features they used to express more complex reasoning moves.

Study 2. Fourth- and fifth-grade argumentative writing: An analysis of argumentative reasoning and connectives

In the second study of my dissertation, I examine a sample of 198 essays written by fourth and fifth grade students (3 essays for each of 66 students) in the context of a cross-content area academic vocabulary intervention, Word Generation. Students were from urban schools in two states in the US. The writing prompts consisted of the Word Generation unit topics (each

grade had different topics), which were contestable, real world issues that students had discussed and debated in class during the previous week. The outcome was argument sophistication in the essays, assessed by the highest argument used according to an argumentative reasoning coding scheme (Kuhn & Crowell, 2011). A measure of deep reading comprehension (Global Integrated Scenario-based Assessment, GISA [O'Reilly, Weeks, Sabatini, Halderman, & Steinberg, 2014]) was included to investigate the role of reading comprehension. Argument types were assessed in the essays, as well as students' uses of major classes of connectives (Halliday & Hasan, 1976) and overall sophistication of connectives, as measured by my own researcher-created rubric. Frequencies of connectives and sophistication of connectives were independent variables. To address my research questions, I first employed descriptive statistics to show how students used these cognitive and linguistic features in their writing, followed by linear regression to test for relations among variables with overall argument sophistication in the essays. The results from Study 2 describe detailed ways these upper elementary students used two domains of academic language in their writing (argumentative reasoning and connecting ideas logically), along with the relation of deep reading comprehension and specific linguistic features they used to express more complex reasoning moves.

Study 3. Cognitive and linguistic features of adolescent argumentative writing: Do connectives signal more complex reasoning?

The third study examines 158 essays written by 60 sixth through eighth grade students (all but 2 students wrote 4 essays each) in the context of a cross-content area academic vocabulary intervention, Word Generation. This paper is currently published (Taylor, Lawrence, Connor, & Snow, 2019). Students were from urban schools in the West Coast of the US. All students wrote about the same Word Generation unit topics, which were contestable, real world

issues that students had discussed and debated in class during the previous week. The outcome was again argument sophistication in the essays, assessed by the highest argument used according to an argumentative reasoning coding scheme (Kuhn & Crowell, 2011). Two different topic types (binary and open ended) were compared. Argument types were assessed in the essays. In this study, students' uses of major classes of connectives were calculated by the Tool for the Automated Analysis of Cohesion (TAACO; Crossley, Kyle, & McNamara, 2016), a natural language processing tool. Densities of connectives per essay were independent variables. To address my research questions, I first employed descriptive statistics to show how students used these measures of cognitive and linguistic features in their writing, followed by linear regression to test for relations among connectives and individual argument types, as well as among connectives variables with overall argument sophistication in the essays. The results from Study 3 provide information on topic differences and describe detailed ways these middle school students used two domains of academic language in their writing (argumentative reasoning and connecting ideas logically), along with specific linguistic features they used to express more complex reasoning moves.

Taking the three studies together helps to paint a developmental picture of how argumentative writing develops. Although not longitudinal data, all studies use the same outcome, highest level of argumentative reasoning used, to examine how argumentative writing differs for students at different developmental phases – early elementary, late elementary, and middle school.

Structure of the Dissertation

The three studies within this dissertation are each presented as a separate manuscript. They will include study-specific introductions, research questions, methods, discussions, and

conclusions. I will end the dissertation with a conclusion that summarizes results and discusses

implications and directions for future research.

References

Allen, L. K., Likens, A. D., & McNamara, D. S. (2019). Writing flexibility in argumentative essays: A multidimensional analysis. *Reading and Writing, 32,* 1607-1634.

Andrews, R., Torgerson, C., Low, G., & McGuinn, N. (2009). Teaching argument writing to 7- to 12-year-olds: An international review of the evidence of successful practice. *Cambridge Journal of Education, 39*(3), 291-310. doi:10.1080/03057640903103751

Bailey, A. L. (Ed.)(2007). *The language demands of school: Putting academic English to the test.* New Haven, CT: Yale University Press.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Berninger, V. W., Abbott, R. D., Abbott, S., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities, 35*(1), 39-56.

Bus, A. G., van IJzendorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1-21.

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

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods, 48,* 1227-1237. doi:10.3758/s13428-015-0651-7

Donovan, M. S., Snow, C. E., & Daro, P. (2013). The SERP approach to problem-solving

    research, development, and implementation. *National Society for the Study of Education*

    Yearbook, 112(2), 400-425.

Ferretti, R. P., & Fan, Y. (2016). Argumentative writing. In C. A. MacArthur, S. Graham, & J.

    Fitzgerald (Eds.), *Handbook of writing research* (pp. 301-315). New York, NY: The

    Guilford Press.

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and

    instruction. *Reading and Writing, 32, 1345-1357.*

Fillmore, L. W., & Snow, C. E. (2000). What teachers need to know about language.

    Washington, DC: U.S. Department of Education.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development.

    Educational Psychologist, 35(1), 39-50.

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., & Aitken, A. (2018). Reading for writing:

    A meta-analysis of the impact of reading interventions on writing. *Review of Educational*

    Research, 88(2), 243-284.

Graham, S., & Hebert, M. (2011a). *Writing to read: Effective strategies to improve writing of*

    adolescents in middle and high schools. New York: Carnegie Corporation of New

    York.

Graham, S., & Hebert, M. (2011b). Writing to read: A meta-analysis of the impact of writing and

    writing instruction on reading. *Harvard Educational Review, 81*(4), 710-144.

Graham S., McKeown, D., Kiuhara, S., & Harris. K. R. (2012). A meta-analysis of writing

    instruction for students in the elementary grades. *Journal of Educational Psychology,*

    104(4), 879-896.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools – A report to Carnegie Corporation of New York.* Washington, DC: Alliance for Excellent Education.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Longman Group Ltd.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243-255.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95,* 163-182.

Kim, Y.-S. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology, 141,* 101-120. http://dx.doi.org/10.1016/j.jecp.2015.08.003

Kim, Y.-S., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology, 109*(1), 35-50.

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

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

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

LaRusso, M., Kim, H. Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., Donovan, S., &

Snow, C. (2016). Contributions of academic language, perspective taking, and complex

 reasoning to deep reading comprehension. *Journal of Research on Educational*

 Effectiveness, 9(2), 201-222.

Lonigan, C. J., Schatschneider, C., Westberg, L., & the National Early Literacy Panel (2008).

 Impact of code-focused interventions on young children's early literacy skills (pp. 107-

 152). *Developing literacy: Report of the National Early Literacy Panel.* National Institute

 for Literacy.

MacGinitie, W.H., MacGinitie, R.K., Maria, K., & Dreyer, L. G. (2006). Gates-MacGinitie

 reading tests (4th ed.)[Kit]. Itasca, IL: Riverside.

National Center for Educational Statistics. (2012). *The Nation's Report Card: Writing 2011*

 (NCES 2012-470). Washington, D.C.: Institute of Education Sciences, U.S. Department

 of Education.

National Commission on Writing. (2004). *Writing: A ticket to work. Or a ticket out.* New York:

 College Board.

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning

 argumentative reading and writing: A review of research. *Reading Research Quarterly,*

 46(3), 273-304.

Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: Alternative

 frameworks for argumentation research in education. *Educational Psychologist, 46*(2),

 84-106.

O'Reilly, T., Weeks, J., Sabatini, J., Halderman, L., & Sternberg, J. (2014). Designing reading

 comprehension assessments for reading interventions: How a theoretically motivated

assessment can serve as an outcome measure. *Educational Psychology Review, 26*(3), 403-424.

Paltridge, B. (2014). Research timeline: Genre and second-language academic writing. *Language Teaching, 47*(3), 303-318.

Payne, A. C., Whitehurst, G. J., & Angell, A. L. (1994). The role of literacy environment in the language development of children from low-income families. *Early Childhood Research Quarterly, 9,* 427-440.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2-3), 155-175. doi: 10.1207/S15326950DP3202&3_04

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction* (pp. 319-334). New York: Guilford.

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

Reznitskaya, A., Kuo, L., Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual differences, 19,* 219-224.

Scarcella, R. (2003). *Academic English: A conceptual framework* (Tech. Rep. No. 2003-1). Santa Barbara, CA: University of California, Linguistic Minority Research Institute.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective.* Mahwah, NJ: Erlbaum.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In N. Torrance & D. R.

 Olson (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge

 University Press.

Storch, S. A., & Whitehurst, G. J. (2001). The role of family and home in the literacy

 development of children from low-income backgrounds. In P. R. Britto & J. Brooks-

 Gunn (Eds.), *The role of family literacy environments in promoting young children's*

 emerging literacy skills. New directions for child and adolescent development (pp. 53-

 71). San Francisco, CA: Jossey-Bass.

Taylor, K. S., Lawrence, J. F., Connor, C. M., & Snow, C. E. (2019). Cognitive and linguistic

 features of adolescent argumentative writing: Do connectives signal more complex

 reasoning? *Reading and Writing, 32,* 983-1007.

 https://doi.org/10.1007/s11145-018-9898-6

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015a).

 Core academic language skills: An expanded operational construct and a novel

 instrument to chart school-relevant language proficiency in preadolescent and adolescent

 learners. *Applied Psycholinguistics,* 1-33. doi:10.1017/S014271641400006X

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015b). Beyond

 vocabulary: Exploring cross-disciplinary academic-language proficiency and its

 association with reading comprehension. *Reading Research Quarterly, 50*(3), 337-356.

 doi:10.1002/rrq.104

Uccelli, P., & Snow, C. E. (2008). A research agenda for educational linguistics. In B. Spulsky

 & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 626-642). Malden,

 MA: Blackwell Publishing.

van Eemeren, F., Grootendorst, R., & Henkemans, F. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments.* Mahwah, NJ: Lawrence Erlbaum Associates Inc.

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

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 68,* 848-872.

Woodward-Kron, R. (2005). The role of genre and embedded genres in tertiary students' writing. *Prospect, 20*(3), 24-41.

CHAPTER 2

Study 1. Argumentative reasoning and connectives in second and third grade

students' persuasive essays

National statistics show that many children and early adolescents, who are still in the process of acquiring the language skills needed for academic reading and writing tasks, are consistently behind their more linguistically mature peers in writing achievement (National Center for Education Statistics, 2012). In addition, achievement gaps in overall writing achievement between children from low socioeconomic status (SES) families, English Learners (EL), and under-represented minority groups, and their more affluent majority peers, is a serious issue, considering that academic writing skills are foundational for success in college and/or future civic involvement (National Center for Education Statistics, 2012; National Commission on Writing, 2004). Researchers, practitioners, and policy makers have attempted to foreground the skill of argumentative writing in the Common Core State Standards, currently adopted by 42 of the states in the U.S. (Common Core State Standards Initiative, 2010, 2018; Duncan, 2009). Indeed, special emphasis is given to the text type of "argument" in the standards, where even in grades K-5, the term "opinion" is used instead to refer to a "developing form of argument" (Common Core State Standards Initiative, 2010, Appendix A, p. 23).

Although argumentative writing has a place of prominence in the standards, the standards do not operationalize the nature of reasoning that is inherent in argumentative writing, nor do they detail just *how* teachers can ensure that these standards are met (Kuhn & Crowell, 2011). Argumentative writing is not only a cognitively demanding skill to learn; it is also a linguistically demanding skill to learn. It entails knowledge of a special register of the English language functional linguists refer to as *academic language*; that is, language deployed for

school purposes (Bailey, 2007; Scarcella, 2003; Schleppegrell, 2004; Snow & Uccelli, 2009). Whereas common ways of assessing classroom writing utilize a holistic rubric that generates a global score (Llosa, Beck, & Zhao, 2011), analytic methods of scoring reasoning in student writing have been found to a reliable means of elucidating fine-grained information about differences in students' argumentation (Reznitskaya, Kuo, Glina, & Anderson, 2009; Kuhn & Crowell, 2011). A recent study has even combined analytic scoring of argumentative moves with computational tools to measure linguistic features in middle school students' essays, in order to investigate relations among the cognitive features and linguistic features (Taylor, Lawrence, Connor, & Snow, 2019). More research is needed, however, and more needs to be known about quantitative methods of assessing detailed information about students' argumentation and related linguistic features for elementary students as well. Thus, using quantitative methods, the current study investigates the argumentative reasoning and linguistic feature of connectives within a sample of second and third grade students' persuasive essays.

The persuasive writing context

Persuasive writing tasks have been used as single-occasion writing prompts in large-scale assessments for holistically evaluating a range of writing skills in upper elementary grades and beyond (Olinghouse, Santangelo, & Wilson, 2012). Studies of writing instruction in elementary school have also used persuasive writing tasks to provide an overall writing quality measure for pretest and posttest assessments (e.g., Graham, Harris, & Mason, 2005). In addition, persuasive writing samples afford evaluating elementary students' writing from a lens of development of genre-specific writing elements (Ingebrand, 2016). A few studies have analyzed specific academic language features in persuasive essays, but they investigated middle school (Dobbs, 2014) or high school students (Uccelli, Dobbs, & Scott, 2013). The outcomes in these two

studies were measures of overall writing quality. Because argumentative reasoning appears to be an element of academic language (see Snow & Uccelli, 2009), even among middle schoolers (Uccelli et al., 2015a, b), it is worthwhile to investigate the reasoning moves within students' persuasive writing. With Common Core State Standards emphasizing supplying reasons to support opinions in the opinion text type for even younger students (i.e., both 2nd and 3rd grades; CCSSI, 2010), it is important to begin to identify young students' command of argument types that display particular types of reasoning. Since holistic measures of writing quality are not a conducive method for this, a more fine-grained approach is needed. One purpose of the current study was to utilize quantitative text analysis to elucidate these young writers' argumentative reasoning. In addition, there is little research on use of particular linguistic features in young students' writing, such as the cohesive devices used in arguments, and again, it has investigated this aspect of productive academic language use in older elementary students (Crowhurst, 1987). Thus, a second purpose of this study was to deploy quantitative text analysis to explore the 2nd and 3rd graders' uses of coordinating uses of connectives (e.g., *in addition, but, since*) at the T-unit level (i.e., an independent clause with any dependent clauses, per Hunt, 1965).

An analytic approach to assessing academic language in writing

Studies of young writers may rely on familiar models of writing such as Bereiter and Scardamalia's (1987) *knowledge-telling model* in order to analyze children's knowledge-telling processes involved in their manipulation of the *content* of students' writing, or researchers may call upon their more complex, *knowledge-transforming model* as a means of encompassing children's problem solving processes along with the content of their writing. The Flower and Hayes model (Hayes & Flower, 1986) is another prominent model of the writing process that aims to explain the stages of writing, along with more recent additions to this model that include

aspects outside of the writing process that influence different cognitive levels (Hayes, 2012; Hayes & Olinghouse, 2015). Both the original model and the more recent iteration, however, aim to describe the writing processes of adults. Hayes and Olinghouse (2015) have proposed adaptations to the Hayes (2012) model in order to apply it to young and developing writers, and for the purpose of suggesting improvements to the Common Core State Standards in writing. Hayes and Olinghouse point out that at the control level, the Standards omit the development of writing schemas, such as the selecting the forms needed for a particular genre. At the process level, they suggest that the internal cognitive process of translation would be an internal cognitive process that could be added to the standards. Here, they suggest that the writing standards should incorporate linguistic knowledge, instead of keeping them as separate language standards as they currently stand. For the widest reach of studying these aspects of the writing process – that is, genre knowledge of argumentative writing, for example, and the translation process of "turning ideas into language" (Hayes & Olinghouse, 2015, p. 491) – study design methods would need to investigate causal mechanisms. However, as a starting point, drawing from a schema theory for argumentation as well as a functional linguistics approach may provide insight to describe children's natural uses of these features in writing, since a particular instructional context or longitudinal design is not afforded.

    Assessing reasoning moves. The Argument Schema Theory (Reznitskaya et al., 2001; Reznitskaya & Anderson, 2002) asserts that children and adolescents develop generalized knowledge of argumentative structures through participating in group or dialogic argumentative activities. A method of testing their hypotheses has been through analyzing independent writing samples produced after group argumentation as a measure of transfer and as a proxy for individual argumentative reasoning. In the present study, 2nd and 3rd grade students were

administered a persuasive writing task to be used as an assessment measure for a study of knowledge of genre elements (Ingebrand, 2016). Although there was no group discussion or debate preceding the writing exercise, an analytic method of parsing the essays to assess the individual ideas within the essays could still be an effective strategy to investigate students' argumentative reasoning (Reznitskaya, Kuo, Glina, & Anderson, 2009). A coding scheme created for an argumentation intervention (i.e., Kuhn & Crowell, 2011) was applied in order to examine relative levels of sophistication of argumentation. For example, an individual T-unit could be coded as one of four possibilities: *non argument, own side only argument, dual perspective argument,* or *integrative perspective argument.* The non-arguments generally consisted of statements that were irrelevant, repetitive of an earlier argument, or unclear; the own side only arguments were those that simply favored the writer's position with supporting arguments. The dual perspective arguments, however, were more complex, in that they addressed the opposing position in an attempt to weaken (or show the negatives) of the opposing arguments. Finally, the integrative perspective arguments were the most sophisticated, and displayed either an acknowledgment of a negative of the writer's own position *or* acknowledged a positive of the opposing position, while still maintaining one's own position. Parsing the essay into T-units afforded scoring each T-unit as one of those four argument types. This would provide a much richer description of students' argumentative reasoning, rather than one holistic writing score that encompasses complexity of argumentation within it.

Assessing connectives. The cohesive device of connectives, and in particular those used in a coordinating rather than subordinating fashion, were selected as another aspect to investigate in this analytic approach to assessing the academic language in student argumentative writing. The knowledge-telling model (Bereiter & Scardamalia, 1987) asserts that children essentially

stack their pre-existing content and ideas during the writing process in general. This appears as an additive function for these ideas, and if assessed holistically, alternative functions for introducing these ideas may not be revealed. For example, at a more fine-grained level, other functions such as opposing ideas or signaling causal relations in the text may not appear. Because argumentative writing employs complex reasoning such as the dual perspective and integrative perspective argument types listed above, an analytic assessment approach provides a means of categorizing and tabulating numerous functions for each idea that is added to a student's text.

The present study applies Halliday and Hasan's (1976) established categories of connectives: additive, adversative, causal, and temporal. Although they suggest an additional category of continuative connectives (e.g., *well, anyway*), these were not of interest for the present study. Additive connectives, used as coordinators at the beginning of a T-unit, could include words or phrases such as: *and*, *also,* or *additionally*. Adversative connectives would include: *but* or *however*. Causal connectives could be expressed as *so* or *then*. Temporal connectives included *first* or *last*.

With the data from these cognitive and linguistic facets of academic language within students' independently written essays, there was also the affordance of exploring relations among the two. For example, instead of a broad category of *linking words and phrases*, as they are listed in the Common Core State Standards within the opinion text type of writing (CCSS, 2010), having these separate categories of functional uses can give a more detailed picture of the students' productive use of connectives, and having arguments categorized as types, affords us the ability to test for relations between connective categories and argument types.

Considering reading comprehension for argumentative reasoning

In addition to analyzing the students' uses of academic language in their persuasive essays, it is important to consider that reading may be a likely contributor to the sophistication of students' argumentative reasoning in their writing. For example, research suggests that reading and writing can be highly correlated (Mehta, Foorman, Branum-Martin, & Taylor, 2005), but the relation of these variables among different age groups remains an open question. Due to the likelihood that reading comprehension is an indicator of the sophistication of argument use in a student's essay, the approach in the current study includes a reading comprehension measure, a vocabulary measure, and a writing fluency measure, in order to explore the correlations among these variables and the outcome of sophistication of argumentation, and thus to determine whether these variables are not so highly correlated that they could remain as indicator variables in the eventual regression models. Beyond the hypothesized relation of reading comprehension with argument sophistication, I anticipate a few additional results. Given our previous findings that middle schoolers produced each argument type (Taylor, Lawrence, Connor, & Snow, 2019), I anticipate the 2nd and 3rd graders' essays to display each argument type as well, although the more complex arguments (dual perspective and integrative perspective) may show up to a lesser extent than for the 6th – 8th grade essays. For the use of connectives, I hypothesize that students will demonstrate each signaling function of connectives: additive, adversative, causal, and temporal. Similar again to our previous findings among middle schoolers, I anticipate that adversative connectives may predict overall argument sophistication in these students' essays.

Research questions

The academic language aspects of argumentative reasoning and use of connectives are assessed in a large sample of 2nd and 3rd grade students' persuasive essays. The study asks the following research questions:

1. What is the nature and variability of 2nd and 3rd grade students' uses of *argument types* (i.e., no argument, own side only, dual perspective, and integrative perspective) in their argumentative essays? Are there differences according to sociodemographic variables or grade? How much variability is accounted for within teachers?

2. Controlling for essay length, is reading comprehension associated with argument sophistication for these 2nd and 3rd grade students' essays?

3. What is the nature and variability of 2nd and 3rd grade students' uses of *connectives* (i.e., additive, adversative, causal, and temporal) in their argumentative essays? Are there differences according to sociodemographic variables or grade? How much variability is accounted for within teachers?

4. Beyond a possible relation with reading comprehension and controlling for essay length, is the use of connectives associated with argument sophistication?

Method

Sample

The data are from a larger study of elementary students' individualized literacy instruction aided by technology (see Ingebrand & Connor, 2016, for a description of the technology). Two schools participated in the study in south-central Arizona. The student participants (N=385) were nearly balanced across the two schools (52% attending one of the

schools). A majority of the students in this data set indicated Hispanic (71.95%) as their ethnicity, followed by 8.05% Black, 5.45% White, 0.78% Asian, and 13.77% of the students identified as mixed-race or unspecified. The two schools were high-poverty schools, with 73% of the students eligible for free and reduced price lunch.

Essay context

The writing prompt (see Appendix A) asked students to provide their opinion and details as to whether or not children should be able to choose their own pets. Students were given 30 minutes to draft their essay, with no assistance given on the writing or on revisions. Thus, the essay writing exercise was an on-demand writing assessment activity.

Data analysis

The essays and all student-generated assessment measures were administered during the fall of the school year.

Covariates: Essay length and student characteristics.

Essay length. Essay length was calculated as total number of T-units (Hunt, 1965), which follows the definition of an independent clause including any dependent clauses.

Grade. All students are from either second or third grade. To control for grade level in the analysis, a dummy variable was created for grade two (GRADE_2) and grade three (GRADE_3). Grade level cohort variables were created to estimate mean differences in outcome variables by grade levels.

Gender. To control for gender in the analysis, a dummy variable was created for female (FEMALE = 1) and male (FEMALE = 0).

English proficiency status. The collaborating school district provided basic information about participating language minority students, differentiating English Language Learners

(ELLs) and Fluent English Proficient students. In Arizona, in order for students' English proficiency status to be classified as English Language Learner, they need to be identified as second language learners on a home language survey, and then they take the Arizona English Language Learner Assessment (AZELLA), which assesses students' overall English proficiency. Students classified as ELLs were assigned to a classroom designated as an English Immersion classroom and thus we used the assigned classroom as the indicator of ELL status. AZELLA scores were not available. A dummy variable for ELL students was created for the analyses.

Predictors: Baseline reading and writing measures, connectives, and arguments.

Reading comprehension. The Gates-MacGinitie Reading Test (GMRT) is a standardized, group-administered reading comprehension assessment that requires students to answer questions after viewing a series of pictures or reading text (MacGinitie, MacGinitie, Maria, & Dreyer, 2006). The GMRT is vertically equated, so it allows for comparison across grades. The Level 2 and Level 3 forms, designed for use in second and third grades, respectively, were used in this study. Reliability information reported here is based on Form S, the form completed by participants in this study. Total reliability for the Level 2 form is .92 (alpha), based on the following three sub-tests: word decoding, word knowledge, and comprehension, and the comprehension sub-test has a reliability of .92. Total reliability for the Level 3 form is .96 (alpha), based on the two sub-tests of vocabulary and comprehension, and the comprehension sub-test is .93. Extended scale scores (ESSs) will be used for the present analysis, because differences in units are the same along the entire scale.

Picture vocabulary. Woodcock-Johnson-III Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001). The WJ-III Picture Vocabulary sub-test is a measure of expressive vocabulary, where students are asked to correctly identify illustrations using specified

terminology. Median reliability for this sub-test was .77 for the 7 to 19 years age range. Reliability was .77 for age 9, which was near the mean of this sample. Scores used from this assessment were the W scores, which are determined across a range of ages, and can thus be compared across the two grades.

Writing fluency. The Writing Fluency sub-test in Woodcock-Johnson-III Tests of Achievement (Woodcock, McGrew, & Mather, 2001) was used to assess writing fluency. In this test, students are provided with three words and tasked with writing grammatically correct sentences that include all three words. Reliability was .77 for age 9, which is near the mean of this study sample. Scores used from this assessment were the W scores, which are determined across a range of ages, and can thus be compared across the two grades.

Connectives. Connective types were coded according to the Halliday and Hassan's (1976) four major classes of connectives and computed as frequency counts per essay:

• additive (e.g., *furthermore*, or *in addition*);

• temporal (e.g., *then*, *after that, finally*);

• causal (e.g., *therefore*, *consequently,* or *as a result*);

• adversative (e.g., *however*, *nevertheless*, or *despite this*).

Argument types. Argument types were assessed based on Kuhn and Crowell's (2011) coding scheme. Each T-unit was assigned one of the following argument types: non-arguments were coded 0; own side only arguments were coded 1; dual perspective arguments were coded 2; and integrative perspective arguments were coded 3. Integrative perspective arguments were usually expressed as two T-units (then combined into one T-unit to assign the argument code), or occasionally as one T-unit if the student is able to fully express his or her reasoning by including a dependent clause. A trained undergraduate research assistant and I dual coded 10% of the data

(Kappa = .71). This kappa statistic is considered adequate for interrater reliability (Landis & Koch, 1977; McHugh, 2012). The remaining essays were also dual coded, and discrepancies were resolved through discussion. (See Appendix B for the coding scheme with examples.)

Outcome: Argument sophistication. The highest argument used in each essay became that essay's score for argument sophistication.

Analysis plan

Descriptive statistics were generated for the measures of essay length, argument types, and connectives. Pairwise comparison of means were conducted to explore possible differences in the observed variables in the essays by grade or subgroups of students. Intraclass correlations (rho statistic) were generated to explore the variability of arguments and connectives within teacher. Correlational analysis results informed the construction of the multiple regression analyses. Cluster robust standard errors were used in the multiple regression analyses to account for the nesting of students within classrooms. Highest argument used was the outcome variable, with reading comprehension and connectives as the independent variables of interest. Essay length and the other variables used as controls will be discussed in the results.

Results

Research Question 1. Uses of arguments

Prior to estimating the frequencies of arguments in the essays, the calculation of percent of cases (essays) containing at least one incidence of each argument type revealed the non-argument code present in 71% of cases; the own-side only argument code present in 72% of cases; the dual perspective code present in 36% of cases; and the integrative perspective code present in 13% of cases. Descriptive statistics for argument types and essay length for the sample

as a whole and disaggregated by grade are reported in Table 2.1. Across both grades, essay

length showed substantial variability; the average essay length was 6.35 T-units, with the

minimum essay length being 1 T-unit and the maximum essay length being 31 T-units. The mean

of non-argument coded T-units across all essays was 2.37, with a range of 0 to 24. The mean of

own side only arguments was 2.85 T-units, with a range of 0 to 15; the mean of dual perspective

arguments was .97 T-units, with a range of 0 to 14. Finally, the mean of integrative perspective

arguments was .15 T-units, with a range of 0 to 3. There were grade differences in the more

complex arguments, the dual perspective arguments and integrative perspective arguments. Third

grade students wrote more dual perspective arguments ($t = 4.32$; $p < .000$), as well as more

integrative perspective arguments ($t = 2.76$; $p < .006$), than did second grade students. Regarding

sociodemographic variables, there were no differences in essay length and frequencies of

argument types by FARL qualified. However, there were differences by gender and language

proficiency. Females wrote longer essays than did males ($t = 2.48$; $p < .013$), and ELLs wrote

shorter essays ($t = -3.00$; $p < .003$) and included fewer own side only arguments ($t = -3.00$; $p <$

.012) than did English only students.

Table 2.1

Means and standard deviations for essay length and argument types ($n = 385$)

Essay Features	M(SD)	2nd grade M(SD)	3rd grade M(SD)
Length			
Essay length in T-units	6.35(4.64)	5.90(4.69)	6.80(4.55)
Argument features			
No argument	2.37(3.18)	2.65(3.47)	2.09(2.85)
Own side only	2.85(3.03)	2.57(2.95)	3.13(3.09)
Dual perspective	.97(1.84)	.58(1.08)	1.37(2.31)
Integrative perspective	.15(.41)	.09(.34)	.21(.47)

For the argument sophistication measure (i.e., the highest argument used in the essay),

there was a significant grade difference; third grade students scored higher on this measure than

second grade students ($t = 4.94$; $p < .000$). There were no significant differences by sociodemographic variables of gender, FARL qualification, or language proficiency.

ICCs are shown in Table 2.2. The rho statistics were calculated in order to discover the variability within teachers, since these data were nested within teachers. Essay length had the largest amount of variability with a *rho* of .15, so 15% of the variability of essay length measured in T-units was accounted for at the teacher level. In addition, frequencies of non arguments, own side only arguments, and dual perspective arguments each had a *rho* of greater than .05 and even approaching .10, so these results informed the construction of a regression model clustered by teacher to estimate robust standard errors. The finding of a *rho* of .01 for integrative arguments and .05 for the highest argument used demonstrates that these variables were less stable within teachers.

Table 2.2

Intraclass Correlations of Arguments by Teacher ($n = 385$)

Essay Features	Rho Coefficient	Standard Error
Essay length		
Essay length in T-units	.15	.06
Argument features		
No argument	.08	.04
Own side only arguments	.09	.04
Dual perspective arguments	.07	.04
Integrative perspective arguments	.01	.02
Highest argument used	.05	.03

Research Question 2. Relation between reading comprehension and argument sophistication

This second research question asks whether reading comprehension is associated with argument sophistication in these second and third grade students' essays. Table 2.3 displays descriptive statistics for the measures of reading comprehension, picture vocabulary, and writing fluency. Table 2.4 presents pairwise correlations between each variable and argument

sophistication (coded as the highest argument used in the essay). Exploratory correlations

revealed no significant correlations among demographic variables and argument sophistication;

therefore, only covariates to be used in the regression model are presented here. The correlations

between each variable and argument sophistication were significantly greater than 0 ($p < .05$).

There were moderate and significant positive correlations between the measures of writing

fluency and reading comprehension with argument sophistication. There were weak but

significant positive correlations between picture vocabulary, third grade, and essay length with

argument sophistication. Conversely, there was a weak negative correlation between second

grade and argument sophistication. The strong positive correlation between writing fluency and

reading comprehension ($r = .6274$) suggested a high degree of collinearity, with the resultant

need to eliminate one when fitting the multiple regression model. Writing fluency was selected,

because reading comprehension was necessary to answer the research question. The correlation

results led to the conclusion to control for these variables in the subsequent regression models.

Table 2.3

Means, standard deviations, and ranges for performance on reading and writing measures ($n = 385$)[a]

Variable	M(sd)	Min.	Max.
WJ Reading Comp			
Full sample	408.09 (37.14)	276	510
2nd grade	391.35 (37.14)	276	491
3rd grade	424.91 (28.62)	343	510
WJ Picture Vocab			
Full sample	476.62 (10.81)	435	509
2nd grade	473.64 (11.27)	435	509
3rd grade	479.61 (9.46)	456	505
WJ Writing Fluency			
Full sample	476.01 (12.83)	450	515
2nd grade	470.81 (11.46)	450	498
3rd grade	481.23 (12.00)	457	515

[a]Note: WJ Picture Vocab and WJ Writing Fluency were each missing scores for 2 students, so $n = 383$ for those measures.

Table 2.4

Correlations among highest argument and covariates

	1	2	3	4	5	6	7
1. Highest argument	1						
2. Reading comp	.370***	1					
3. Picture vocab	.220***	.457***	1				
4. Writing Fluency	.345***	.627***	.415***	1			
5. 2nd Grade	-.245***	-.452***	-.276***	-.407***	1		
6. 3rd Grade	.245***	.452***	.276***	.407***	-1.000	1	
7. Total T-units	.170***	.123*	-.009	.119*	-.098	.098	1

$*p < .05$. $***p < .001$.

Table 2.5 displays results from the multiple regression model that was constructed to estimate whether there was a significant association between reading comprehension and argument sophistication. With argument sophistication as the outcome, then, reading comprehension, picture vocabulary, third grade, and essay length were entered as covariates in the model. Robust standard errors (Huber-White adjustment) were used to account for the nesting of students within teachers. Raw scores were used for the argument sophistication dependent variable, and standardized scores were used for the continuous variables of reading comprehension, picture vocabulary, and essay length. Table 2.4 displays results of this model, which explained 17% of the variance in argument sophistication in these second and third grade students' essays. Reading comprehension was positively associated with argument sophistication ($p < .001$), such that essays written by students with a reading comprehension score 1 *SD* above the mean scored .247 higher on argument sophistication. Picture vocabulary did not have a significant association with argument sophistication, but the coefficient indicates that students with a picture vocabulary score 1 *SD* above the mean scored .054 higher on argument sophistication. Grade was also not significant in the model, but the coefficient indicates that third grade students scored .136 higher on argument sophistication. Essay length as measured in total

T-units was significant and positively associated with argument sophistication, such that essays with a length 1 *SD* above the mean scored .107 higher on argument sophistication. Note that all interactions were tested and none were significant.

Table 2.5

Regression Model for Argument Sophistication: Relation with Reading Comprehension ($n = 383$)[a]

Outcome var	Ind. Variables	ß	Robust Standard Error	*p*	R^2
Highest Arg	Reading comp[b]	.247***	.051	.000	
	WJ PV[b]	.054	.044	.238	
	3rd Grade	.136	.069	.068	
	Total T-units[b]	.107*	.010	.038	
	Constant	1.396***	.056	.000	
					.17***

p < .05. ***p* < .001.
[a]Note: Due to missing data on WJ Picture Vocabulary, casewise deletion resulted in *n* = 383 for analysis.
[b]Note: These variables are standardized.

Research Question 3. Uses of connectives

Percent of cases (essays) containing at least one incidence of each connective type were computed, revealing at least one additive connective was present in 65% of cases; at least one adversative connective was present in 23% of cases; at least one causal connective was present in 33% of cases; and at least one temporal connective was present in 7% of cases.

Table 2.6 presents means, standard deviations, and ranges for each connective type for the sample as a whole and disaggregated by grade. Across both 2nd and 3rd grades, total use of connectives showed substantial variability; the average was 2.96 connectives in an essay, with a range of 0 to 21 total connectives. Additive connectives were the most prevalent, with a mean of 1.92 across the sample and a range of 0 to 16. Adversative, causal, and temporal connectives each had a mean of less than 1 per essay; however, the ranges showed considerable variability, with adversative connectives ranging from 0 to 7, causal connectives ranging from 0 to 8, and temporal connectives ranging from 0 to 7. Pairwise comparison of means revealed grade differences in additive, causal, and total amount of connectives. That is, third grade students

produced more additive connectives (t = 2.43; p < .016), as well as more causal connectives (t =

2.56; p < .011) and more total connectives (t = 2.56; p < .011) than did second grade students.

There were also gender differences. Females wrote more additive connectives (t = 2.48; p <

.014), more causal connectives (t = 4.09; p < .000), and more total connectives (t = 3.38; p <

.001) than did males. There were no differences by FARL eligibility. Regarding language

proficiency, ELLs wrote fewer causal connectives (t = -2.36; p < .019) and fewer total

connectives (t = -2.70; p < .007) than did English only students.

Table 2.6

Descriptive statistics for connectives (n = 385)

Connectives									
	Full sample			*2ⁿᵈ grade*			*3ʳᵈ grade*		
	M(SD)	Min.	Max.	M(SD)	Min.	Max.	M(SD)	Min.	Max.
Additive	1.92(2.61)	0	16	1.60(.2.45)	0	15	2.24(2.74)	0	16
Adversative	.34(.81)	0	7	.33(.75)	0	5	.35(.87)	0	7
Causal	.56(1.05)	0	8	.42(1.04)	0	8	.69(1.05)	0	6
Temporal	.14(.67)	0	7	.16(.75)	0	7	.12(.59)	0	5
Total	2.96(3.47)	0	21	2.51(3.27)	0	17	3.41(3.60)	0	21

Rho statistics for connectives are presented in Table 2.7. The variability of use of

connectives accounted for by teacher was in general lower than the variability reported for the

uses of arguments within teacher (Table 2.2 earlier reported argument rhos). Total use of

connectives had the largest variability with a *rho* of .07; however the individual connective types

each had a *rho* of less than .05. This suggests that students' uses of connective types were not

very stable within teachers.

Table 2.7

Intraclass Correlations of Connectives by Teacher (n = 385)

Essay Features	Rho Coefficient	Standard Error
Connectives		
Additive	.04	.03
Adversative	.03	.02
Causal	.04	.03
Temporal	0	0
Total	.07	.04

Research Question 4. Relation between connectives and argument sophistication

Pearson correlations (see Table 2.8) were computed for argument sophistication, reading comprehension, picture vocabulary, grade, essay length, total connectives, and each of the four connectives categories (i.e., additive, adversative, causal, and temporal). Note that there were no significant correlations among demographic variables and argument sophistication.

Table 2.8

Correlations Among Highest Argument, Covariates, and Connectives

	1	2	3	4	5	6	7	8	9	10
1. Highest argument	___									
2. Reading comp.	.370***									
3. Picture vocab.	.220***	.457***								
4. 2nd grade	-.245***	-.452***	-.276***							
5. 3rd grade	.245***	.452***	.276***	-1.000						
6. Total T-units	.169***	.123*	-.009	-.098	.098					
7. Total connect.	.225***	.136**	.022	-.130*	.130*	.739***				
8. Additive connect.	.119*	.068	.006	-.123*	.123*	.639***	.879***			
9. Advers. connect.	.284***	.051	-,040	-.014	.014	.400***	.414***	.096		
10. Causal connect.	.191***	.189***	.100	-.130*	.130*	.304***	.483***	.169***	.178***	
11. Temporal connect.	.053	.076	-.021	.030	-.030	.357***	.472***	.262***	.273***	.051

*$p < .05$. **$p < .01$. ***$p < .001$.

On the basis of the correlations results, picture vocabulary, essay length, and third grade were again used as control variables in the subsequent regression model. Writing fluency was also not selected for the same reason as with the earlier finding of the strong correlation with reading comprehension, where reading comprehension is a key independent variable for the research question. Total use of connectives was also not selected for the model, as it is redundant (i.e., each connective type adds up to the total connectives).

To answer whether connectives would explain any additional variance in argument sophistication in the essays beyond reading comprehension and the same previous controls, the

variables for each connective type (coded as frequency counts) were entered into the regression model. Robust standard errors were used to account for the nesting of students within teachers. Raw scores were used for the argument sophistication dependent variable, and standardized scores were used for the continuous variables of reading comprehension, picture vocabulary, essay length, and connectives. As displayed in Table 2.9, including the connectives variables contributed 7% of additional variance in the model ($R^2 = .24$). Reading comprehension was positively associated with argument sophistication ($p < .001$), such that students with a reading comprehension score 1 SD above the mean scored .239 higher on argument sophistication. Picture vocabulary was not significant, but the coefficient suggests that students with scores 1 SD above the mean had a .06 increase in argument sophistication. Grade was also not significant, but the coefficient suggests that third grade students had a .124 higher score on average for argument sophistication. Essay length was no longer significant, but this model reflects a negative relation, such that students with an essay length 1 SD above the mean had a .043 decrease on argument sophistication. Of the connective types, there was a significant positive association between adversative connectives and argument sophistication ($p < .01$); essays with frequencies of adversative connectives 1 SD above the mean scored .237 higher on argument sophistication. Note that there were no significant interactions between the variables in this model.

Table 2.9

Regression Model for Argument Sophistication: Contribution of Connectives (n = 385)[a]

Outcome var	Ind. Variables	ß	Robust Standard Error	p	R^2
Highest Arg	Reading comp[b]	.239***	.054	.000	
	WJ PV[b]	.060	.046	.213	
	3rd Grade	.124	.062	.063	
	Total T-units[b]	-.043	.065	.518	
	Additive[b]	.082	.041	.066	
	Adversative[b]	.237**	.063	.002	
	Causal[b]	.065	.033	.066	
	Temporal[b]	-.043	.040	.294	
	Constant	1.402***	.053	.000	
					.24***

*p < .05. **p < .01. ***p < .001.
[a]Note: Due to missing data on WJ Picture Vocabulary, casewise deletion resulted in *n* = 383 for analysis.
[b]Note: These variables are standardized.

Discussion

The primary aim of this study was to examine children's productive academic language use in the area of two domains: argumentative reasoning and use of connectives to signal specific arguments. Kuhn and Crowell's (2011) coding scheme was used to assess individual arguments in the students' essays, and Halliday and Hasan's (1976) categories for connectives were used to code students' uses of connectives. Because the argumentative reasoning coding scheme was developed for a sixth grade classroom intervention on argumentation, another purpose of the study had to do with the methodological question of how, if at all, this coding scheme would work with second and third grade students' writing. In a study of middle school persuasive writing, Taylor, Lawrence, Connor, and Snow (2019) confirmed that it is a reliable tool for assessing argument types. The present study extended these findings, as the coding scheme afforded a reliable means of assessing argument types in this sample of second and third grade persuasive essays (i.e., Kappa = .71), which is considered an adequate measure of interrater reliability (Landis & Koch, 1977; McHugh, 2012). This indicates that this analytic measurement

tool may yield some rich insights into student reasoning and writing for a wide age-range in elementary and middle school.

The text-oriented analysis using Kuhn and Crowell's (2011) coding scheme indeed helped to uncover rich descriptive data for these children's uses of specific argument types in their essays. The setting for the essays was a fall writing assessment measure for a larger study, in which students had about 30 minutes to independently write their essays. The mean essay length across both grades was 6.35 T-units. Overall, this is a fairly brief piece of writing for a 30-minute time frame, which is unsurprising for this age range, especially given that there was considerable variability on this measure of length. Yet one consideration for the brevity in general is that the writing assessment took place during the fall of the school year, so it was likely that students had not had much writing instruction or practice at that point in the school year. In addition, written argumentation is a cognitively challenging task (Ferretti & Fan, 2016) and a skill that is slow to develop (Ferretti & Graham, 2019). Descriptive findings for the argument types revealed that each type of argument appeared at least once across the corpus, and notably, the most complex argument type, integrative perspective, appeared in 13% of cases (essays). This is a more sophisticated type of argumentative reasoning, wherein the student either considers a negative aspect of their own position, or a positive aspect of the opposing position, yet maintains their original position. Results indicated that third grade students wrote significantly more of these integrative perspective arguments, along with the dual perspective arguments, which are also an example of more complex reasoning, in that they also consider the opposing position with the goal of weakening the opposing position's argument instead of simply advancing the positives of their own position. Kuhn and Crowell (2011) argue that dual perspective and integrative perspective arguments, while more complex than own side only

arguments, do not reach the level of skilled and nuanced adult argumentation, but that they are fundamental for students to develop toward that end. In the present study, we characterize dual perspective and integrative perspective arguments each as increasingly more sophisticated than own side only arguments, even referring to them as complex reasoning. Elsewhere, Kuhn (1992) describes an increasingly complex nature of epistemological beliefs theorized to underlie argumentative reasoning. In addition, Kuhn and Crowell (2011) define how each argument type is qualitatively different from the other and becomes progressively more sophisticated, even though they developed the coding scheme for sixth grade students' and not adult argumentative reasoning. Thus, we think that their coding scheme serves the purpose of evaluating elementary students' argument types in their writing quite well, in that it can uncover the relative differences in complexity in the specific arguments that students use. Finally, given that we are assessing elementary students' argumentative writing, we are not investigating disciplinary ways of arguing, such as with argumentation in science or history, where there are distinct strategies or structures that are valued within each discipline. Rather, we seek to investigate students' argumentative reasoning about a general or social topic, and the coding scheme we have adopted affords just that. Because this measurement tool lends itself to studying general argumentative writing (rather than that of a particular discipline), our finding that it can be used to reliably assess a range of arguments with younger elementary students as well as middle school students, is promising for future research as well as practice. The finding that some students produced the most complex type of argument, integrative perspective (likely without any prior instruction), suggests that these argument types could be incorporated into classroom instruction dealing with oral or written argumentation.

Regarding the relation of reading comprehension and argument sophistication, there was a significant positive association in a regression model controlling for picture vocabulary, grade, and essay length, with an overall model that explained 17% of the variance on argument sophistication in the essays. While we saw significant grade differences earlier with pairwise comparison of argument sophistication means, there was not a significant association between grade and argument sophistication in the model. Essay length, however, had a significant positive association with argument sophistication in the students' essays. Research has established a connection between reading and writing (Fitzgerald & Shanahan, 2000; Graham et al., 2018). One simple explanation of the role of reading during writing is that a writer needs to read in order to verify one's ideas as they are writing (Hayes, 1996; Kellogg, 1996). Another consideration for this relation is the possible influence of the home literacy environment, as research has shown that early home literacy practices are related to language skills and reading achievement (e.g., Bus, van IJzendorn, & Pellegrini, 1995).

Connectives in students' essays also yielded rich descriptive data with low variability within teachers. If we had access to multiple writing samples per student, we could also explore the variability of connectives within students, to shed light as to how stable of a writing feature this might be within students. Regarding the relation of connectives beyond the relation of reading comprehension, reading comprehension remained significant, and adversative connectives were also significant in a model that explained 24% of the variance in essays' argument sophistication. With connectives defined as a feature of academic language, this aligns with Crosson and Lesaux's (2013a, b) recommendations for teaching connectives as a special subset of academic vocabulary in elementary school.

Conclusion

The results of the current study should be interpreted with several limitations in mind. First, the findings for this sample of children in high poverty schools may not generalize to other samples. Importantly, there is considerable variation left unexplained in the regression models, as there was fairly low variation explaining the outcome of argumentative reasoning (17% and 24%). However, given that the outcome was at the T-unit level and a measure of reasoning as opposed to overall writing quality, it is a helpful starting point. In addition, while the coding scheme discussed above appears to capture a progression of argumentative reasoning skill, we are not able to assume that it is an exact linear relationship. In future work, implementing more complex measurement tools for general argumentative writing that incorporate overall writing structure, argument quality, and argumentative strategies may help to paint a richer picture of overall argumentative writing that incorporates the reasoning elements (for possible alternative coding schemes, see Ferretti, Lewis, & Andrews-Weckerly, 2009, and Reznitskaya et al., 2009).

Despite these limitations, further research in the areas of cognitive features and linguistic features of elementary students' argumentative writing can inform writing instruction and assessment. Multiple writing samples per student are recommended, for then this type of quantitative assessment of student writing may help to illuminate individual variation on multiple cognitive and linguistic features. Future studies should employ longitudinal designs, and classroom intervention studies targeting instruction of general argumentative writing are warranted.

References

Bailey, A. L., Butler, F. A., & Stevens, R., & Lord, C. (2007). Further specifying the language demands of school. In A. L. Bailey (Ed.), *The language demands of school: Putting academic English to the test.* New Haven, CT: Yale University Press.

Bereiter, C., & Scardamalia, M. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics, Vol. 2* (pp. 142-175). Cambridge, UK: Cambridge University Press.

Bus, A. G., van IJzendorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65,* 1-21.

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

Common Core State Standards Initiative. (2018). *Standards in your state* webpage. Retrieved from http://www.corestandards.org/standards-in-your-state/.

Crosson, A. C., & Lesaux, N. K. (2013a). Connectives: Fitting another piece of the vocabulary puzzle. *The Reading Teacher, 67*(3), 193-200. doi: 10.1002/TRTR.1197

Crosson, A. C., & Lesaux, N. K. (2013b). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading, 36*(3), 241-260. doi: 10.1111/j.1467-9817.2011.01501.x

Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English, 21*(2), 185-201.

Dobbs, C. L. (2014). Signaling organization and stance: Academic language use in middle grade persuasive writing. *Reading and Writing, 27,* 1327-1352.

Duncan, A. (2009). Teachers College, Columbia University policy address on teacher preparation. Retrieved from http://www.tc.columbia.edu/articles/2009/october/arne-duncan-full-transcript/

Ferretti, R. P., & Fan, Y. (2016). Argumentative writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 301-315). New York, NY: The Guilford Press.

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and instruction. *Reading and Writing, 32, 1345-1357.*

Ferretti, R. P., Lewis, W. E., & Andrews-Weckerly, S. (2009). Do goals affect the structure of students' argumentative writing strategies? *Journal of Educational Psychology, 101*(3), 577-589.

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

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*(2)*,* 207-241. doi:10.1016/j.cedpsych.2004.08.001

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., & Aitken, A. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88*(2), 243-284.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Longman Group Ltd.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C.
    M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual*
    differences and applications (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369-388.
    doi: 10.1177/0741088312451260

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American Psychologist,*
    41(10), 1106-1113.

Hayes, J. R., & Olinghouse, N. G. (2015). Can cognitive writing models inform the design of the
    Common Core State Standards? *The Elementary School Journal, 115*(4), 480-497.
    doi: 10.1086/681909

Hunt, K. W. (1965). A synopsis of clause-to-sentence length factors. *The English Journal, 54*(4),
    300+305-309.

Ingebrand, S. W. (2016). The development of writing skills: The use of genre-specific elements
    in second and third grade students' writing. (Unpublished doctoral dissertation). Arizona
    State University, Tempe, AZ.

Ingebrand, S. W., & Connor, C. M. (2016). Assessment-to-instruction (A2i): An online platform
    for supporting individualized early literacy instruction. In S. Crossley & D. S. McNamara
    (Eds.), *Adaptive educational technologies for literacy instruction* (pp. 33-48). New York,
    NY: Routledge.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell
    (Eds.), *The science of writing: Theories, methods, individual differences and applications*
    (pp. 57-72). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young

adolescents' thinking. *Psychological Science, 22*(4), 545-552.

doi:10.1177/0956797611402512

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical

data. *Biometrics, 33*(1), 159-174.

Llosa, L., Beck, S. W., & Zhao, C. G. (2011). An investigation of academic writing in secondary

schools to inform the development of diagnostic classroom assessments. *Assessing*

Writing, 16, 256-273.

MacGinitie, W.H., MacGinitie, R.K., Maria, K., & Dreyer, L. G. (2006). Gates-MacGinitie

reading tests (4th ed.) [Kit]. Itasca, IL: Riverside.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276-

282.

Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. P. (2005). Literacy as a

unidimensional multilevel construct: Validation, sources of influence, and implications in

a longitudinal study in grades 1 to 4. *Scientific Studies of Reading, 9*(2), 85-116.

doi: 10.1207/s1532799xssr0902_1

National Center for Educational Statistics. (2012). *The Nation's Report Card: Writing 2011*

(NCES 2012-470). Washington, D.C.: Institute of Education Sciences, U.S. Department

of Education.

National Commission on Writing. (2004). *Writing: A ticket to work…or a ticket out: A survey of*

business leaders. Retrieved from

http://www.collegeboard.com/prod_downloads/writingcom/writing-ticket-to-work.pdf

Olinghouse, N. G., Santangelo, T., & Wilson, J. (2012). Examining the validity of single-occasion, single-genre, holistically scored writing assessments. In E. Van Steendam & M. Tillema (Eds.), *Measuring writing: Recent insights into theory, methodology, and practices* (pp. 55-82). Bingley, UK: Emerald Group Publishing Limited.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2-3), 155-175. http://dx.doi.org/10.1080/0163853X.2001.9651596

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In M. Pressley & C. Block (Eds.), *Comprehension instruction: Research-based best practices*, (1st edition, pp. 319-334). New York: Guilford Press.

Reznitskaya, A., Kuo, L., Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual Differences, 19,* 219-224.

Scarcella, R. (2003). *Academic English: A conceptual framework* (Tech. Rep. No. 2003-1). Santa Barbara, CA: University of California, Linguistic Minority Research Institute.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511609664.008

Taylor, K. S., Lawrence, J. F., Connor, C. M., Snow, C. E. (2019). Cognitive and linguistic features of adolescent argumentative writing: Do connectives signal more complex reasoning? *Reading and Writing, 32,* 983-1007.

https://doi.org/10.1007/s11145-018-9898-6

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015a).

Core   academic language skills: An expanded operational construct and a novel instrument to

chart school-relevant language proficiency in preadolescent and adolescent learners.

Applied Psycholinguistics, 1-33. doi:10.1017/S014271641400006X

Uccelli, P., Dobbs, C. L., & Scott, J. (2013). Mastering academic language: Organization and

stance in the persuasive writing of high school students. *Written Communication, 30*(1),

36-62. doi: 10.1177/0741088312469013

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015b). Beyond

vocabulary: Exploring cross-disciplinary academic-language proficiency and its

association with reading comprehension. *Reading Research Quarterly, 50*(3), 337-356.

doi:10.1002/rrq.104

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson-III tests of

achievement. Itasca, IL: Riverside.

CHAPTER 3

Study 2. Fourth and fifth grade argumentative writing: An analysis of

argumentative reasoning and connectives

The development of academic language can be a particular challenge for students as they move into the upper elementary years of school and beyond (Snow & Uccelli, 2009), where they face advanced literacy tasks such as producing coherent argumentative writing (Common Core State Standards Initiative, 2010; Ferretti & Graham, 2019). Although there is a paucity of research on elementary students' use of academic language features in their argumentative writing, some recent studies beginning at sixth grade have investigated argumentative reasoning in student writing (Kuhn & Crowell, 2011), academic language forms signaling organization and stance (Dobbs, 2014), and relations among argumentative reasoning and academic language features (Taylor, Lawrence, Connor, & Snow, 2019). These studies used quantitative methods to study these particular features, which afforded precise information about specific cognitive or linguistic moves students made in their writing, as opposed to a reliance on holistic writing assessment measures that may provide limited information on a range of writing skills (Llosa, Beck, & Zhao, 2011; Olinghouse, Santangelo, & Wilson, 2012). The current study examined the academic language domains of argumentative reasoning and use of connectives as an aspect of lexical precision by assessing fourth and fifth grade persuasive writing using quantitative methods.

Academic language domains for middle grades

Drawing from a number of researchers, Snow and Uccelli (2009) present an inventory of the linguistic features within academic language. They point out that "…there is no simple definition of what academic language is" (p. 112). For example, the features of *interpersonal*

stance, *information load*, *organization of information*, *lexical choices*, and *representational congruence* are arranged with examples to show a continuum, from *more colloquial* to *more academic* (p. 119-120). They also include cognitive accomplishments of *genre mastery*, *reasoning strategies*, and *disciplinary knowledge*. In the cognitive accomplishment of command of reasoning strategies, for example, evidence of a more colloquial performance would show "basic ways of argumentation and persuasion," while a more academic performance would demonstrate "specific reasoning moves valued at school," citing Reznitskaya et al. (2001). The depiction of the academic language register on a continuum from less academic to more academic is a helpful one, given that the earliest conceptualization of Cummins (1980, 1981) was purely binary (i.e., Basic Interpersonal Communicative Skill, or BICS, and Cognitive Academic Language Proficiency, or CALP). More recently, Uccelli and colleagues (2015a) have developed a theoretical construct of cross-content area academic language specifically among grades 4-8 called *Core Academic Language Skills (CALS)*. They conceptualize six areas that "…underlie skilled comprehension of academic texts" (Uccelli et al., 2015a, pp. 1084-1085): *unpacking complex words* (morphological skills); *comprehending complex sentences* (syntactic awareness skills); *connecting ideas* (skill with using academic discourse markers); *tracking themes* (skill with anaphoric referents); *organizing argumentative texts* (skills in structuring argumentative texts); and *awareness of academic register* (production and identification skills in definitions, specifically through lexical precision and concise information packing).

This advances the operationalization of academic language for researchers and practitioners, because academic language has been mostly investigated as academic vocabulary (Lesaux et al., 2010; Snow, Lawrence, & White, 2009; Townsend & Collins, 2009; Townsend, Filippini, Collins, & Biancarosa, 2012). Because upper elementary grade students and middle

school students are still developing their academic language skills, it is beneficial to study a broader range of academic language features in students' writing.

An example of the skill with the *connecting ideas* domain of academic language in the CALS construct is that of knowledge of connectives. Since knowledge of this domain of academic language has been found to predict reading comprehension for elementary students (Crosson & Lesaux, 2013), it would also be useful to investigate whether productive use of connectives in writing is related to the reasoning in elementary students' argumentative writing (another domain of academic language mentioned above). This study draws upon the idea that students need to develop academic language skills not only in order to understand the more challenging texts they will read in upper elementary grades, but also to apply these skills to the academic texts they will be asked to produce, such as pieces of argumentative writing. Therefore, it is important to understand students' independent capabilities with academic language in their argumentative writing. Gaining insight into the ways students use these features entirely on their own can be a starting point to inform later design of possible classroom interventions of academic language for argumentative writing.

A within-student approach to analyzing essay features

In the present study, the instructional context in which the 4th and 5th grade essays were produced was within the weekly units of the Word Generation academic vocabulary curriculum. Integrated into the curriculum teaching target words consisting of general academic vocabulary were weekly culminating essays eliciting students' positions about the contestable topic they had learned about, discussed, and debated in class. For example, 4th grade students learned about the topic of bilingualism, and the culminating essay asked students to elaborate their position as to whether everyone should learn a second language. Each of the essay topics proposed similar,

clear, essentially two-sided questions. I selected essays written on three different topics within each grade level, because this would afford comparing use of academic language features within students as well as between grades or subgroups of students. These independently written essays also provided a sample to investigate specific argumentative reasoning moves.

The current study's approach is grounded in an analytic (as opposed to holistic) scoring scheme. Reznitskaya and colleagues (2006, 2009) describe how analytic scoring methods can provide more detailed information about specific argumentative moves within a piece of writing. Accordingly, T-units (Hunt, 1965) are used to represent separate idea units within each essay, to enable categorization of that idea unit as one specific argumentative reasoning move. Kuhn and Crowell's (2011) analytic coding scheme is used for this study, because it assesses levels of argumentative sophistication and incorporates argumentative moves that address the opposing argument (counterarguments). In addition, our previous study (Taylor, Lawrence, Connor, & Snow, 2019) validated it as a useful and reliable tool.

Essays are parsed into individual ideas (i.e., T-units) and then each T-unit is coded as one of four distinct argument types: non argument; own side only argument; dual perspective argument; or integrative perspective argument. A T-unit would be coded as a *non argument* if the idea was irrelevant or repetitive, for example. An *own side only* argument would be assigned if the idea simply supported the student's own position. A *dual perspective* argument would be assigned if the idea offered a negative view of the opposing position. Finally, an *integrative perspective* argument would be assigned if the idea included a negative about the favored position or a positive about the opposing position. (For examples from the data, see Appendix D.) Dual perspective arguments and integrative perspective arguments are considered to be more complex argument types, in that they address the opposing position instead of relying on own

side only arguments, which would display the *myside bias* that has been well documented in research (Ferretti & Graham, 2019). A data corpus of multiple essays per student affords investigating patterns of argument use within students as well as between subgroups of students or grades.

Similarly, students' use of the domain of lexical precision as an aspect of academic language across multiple essays provides rich and detailed information about ways students deploy these linguistic features in their writing. For example, skill with using connectives such as *because* or *nonetheless* – is one of the linguistic skills featured in the academic language inventory discussed earlier (Snow & Uccelli, 2009) as well as one of the facets of the Core Academic Language Skills construct developed by Uccelli and colleagues (2015a, b). Since the function of connectives is to "explicitly signal the connections between passages of text" (Hu & Li, 2015, p. 30), studying the ways students do this in their argumentative writing is worthwhile. In particular, the work of Halliday and Hasan (1976) is still relevant for considering major categories of connectives that signal additive, adversative, causal, and temporal relationships. Additive connectives would include such words or phrases as: *also*, or *in addition*. Examples of adversative connectives would be: *but,* or *however.* Causal connectives could be expressed as: *therefore* or *as a result.* Examples of temporal connectives include: *then*, *after that, finally* (Halliday & Hasan, 1976). Because essays are parsed into T-units for argument coding, this situates the essays for ease of tabulating categories of connectives as well. We decided to focus on coordinating uses of connectives, in order to find out just how these 4th and 5th grade students signaled the forthcoming idea (argument) in the T-unit, rather than on subordinating uses of connectives. Our chosen method of T-units segmenting included dependent clauses that may indeed encompass some uses of connectives to link a dependent clause, but those uses were

beyond the scope of the present study. As a starting point, it is worthwhile to describe these students' coordinating uses of connectives, as we can then analyze them for relations with specific arguments.

Following our previous work where each of the four argument types were found in essays written by 6th through 8th grade students (Taylor, Lawrence, Connor, & Snow, 2019), I anticipated that all four of the argument types would be present in the 4th and 5th grade essays, although perhaps to a lesser extent. I also expected that there would again be a relation between adversative connectives and argument sophistication.

Research questions

The following research questions were developed to investigate the argumentative reasoning and connectives used in the 4th and 5th grade students' persuasive essays:

1. What is the nature of and variability in 4th and 5th grade students' use of *argument types* (i.e., no argument, own side only, dual perspective, and integrative perspective) in their argumentative essays? Are there differences according to sociodemographic variables, grade, or topic? How much variability is accounted for within students, within teachers, or within topics?

2. Controlling for essay length, is reading comprehension associated with strength of argumentation?

3. What is the nature and variability of 4th and 5th grade students' uses of *connectives* (i.e., additive, adversative, causal, and temporal) in their argumentative essays? Are there differences according to sociodemographic variables, grade, or topic? How much variability is accounted for within students, within teachers, or within topic?

4. (a) Beyond a possible relation with reading comprehension and controlling for essay length, is the sophistication of connectives associated with argument sophistication? (b) Beyond a possible relation with reading comprehension and controlling for essay length, are connective types associated with argument sophistication?

Methods

Sample

The sample demographics are shown in Table 3.1. The sample was balanced by grade and had a large percentage (67%) of students who were eligible for free or reduced price lunch. There was a slightly larger percentage of females (56%) across the sample. There were low percentages of students who were classified as English language learners (ELLs) or as special education status (2% and 5%, respectively). The sample was ethnically diverse.

Table 3.1

Student Demographic Characteristics

Demographic	4th grade n (% of grade)	5th grade n (% of grade)	All students n (% of sample)
Total students	33(50)	33(50)	66(100)
Gender			
Female	16(48)	21(64)	37(56)
Male	17(52)	12(36)	29(44)
Ethnicity			
White	14(42.4)	10(30.3)	24(36.4)
Black/African American	11(33.3)	11(33.3)	22(33.3)
Hispanic/Latino/a	6(18.2)	6(18.2)	12 (18.18)
Asian	2 (6.1)	4(12.1)	6 (9.09)
Two or more races/Other	0(0)	2 (6.1)	2(3.03)
Socioeconomic status			
Eligible for free or reduced price lunch	24(73)	20(61)	44(67)
Not eligible for free or reduced price lunch	9(27)	13(39)	22(33)
English proficiency designation			
Classified as ELL	1(3)	0 (0)	1(2)
Not classified as ELL	32(97)	33(100)	65(98)
Special education status			
Classified	3(9)	0 (0)	3(5)
Not classified as	30(91)	33(100)	63(95)

Procedure

The data for this study were collected in 2013-2014 from 4th and 5th grades in K-8 urban schools across two states. These data were part of the larger impact evaluation of Word Generation conducted with students in 4th through 7th grades (LaRusso, Jones, Kim, Kim, Donovan, & Snow, 2016). Three essay topics were selected within each of the 4th and 5th grade curricula. Fourth grade topics were the following: Unit 2, Should we share responsibility for each other's behavior in school? Unit 4, Should students wear school uniforms? And Unit 5, Bilingual or English only at school? The following 5th grade topics were selected: Unit 3, Who is responsible for cleaning up Oregon's shoreline? Unit 4, Should we turn donated land into a mall or a park? And Unit 6, Should we give up privacy for security? (i.e., Should we install video cameras to monitor the hallways in school?) Students wrote essays independently, as the unit-culminating essays were essentially brief, on-demand writing prompts. There were 33 4th grade students, resulting in 99 4th grade essays, and there were 33 5th grade students, also resulting in 99 essays. Thus, the analysis sample consisted of 198 essays.

Data analysis

Students' essays were transcribed as Microsoft Word documents to facilitate the coding process for arguments and connectives. The following variables were generated for analysis.

Covariates: Essay features and student characteristics.

Essay length. Essays were segmented into T-units per Hunt (1965) to provide a measure of essay length. I segmented all essays into T-units, and then a trained undergraduate research assistant checked all essays. Discrepancies were resolved through discussion. The essay length variable was calculated as the total number of T-units per essay. Standardized scores for essay length were used in the regression model.

Topic. Each grade wrote on three separate topics. Dummy variables were created for each topic in order to examine essay features by topic and to control for topic:

• 4th: Should we share responsibility for each other's behavior in school? (RESPONSIBILITY = 1; all others = 0).

• 4th: Should students wear school uniforms? (UNIFORMS = 1; all others = 0).

• 4th: Bilingual or English only at school? (LANGUAGE = 1; all others = 0).

• 5th: Who is responsible for cleaning up Oregon's shoreline? (OREGON = 1; all others = 0).

• 5th: Should we turn donated land into a mall or a park? (LANDUSE = 1; all others = 0).

• 5th: Should we give up privacy for security? (PRIVACY = 1; all others = 0).

Grade. All students were from either 4th or 5th grade. Dummy variables were created for 4th grade (GRADE_4) and 5th grade (GRADE_5) in order to examine essay features by grade and to control for grade.

Gender. To examine essay features by gender in the analysis, a dummy variable was created for female (FEMALE = 1) and male (FEMALE = 0).

FARL. To examine essay features by socioeconomic status, a dummy variable was created as to whether the student qualifies for Free and Reduced Lunch (FARL = 1) or does not qualify (FARL = 0).

Ethnicity. Ethnicity variables of White, Black, Latino/a, Asian, and Mixed race/other were each coded as binary variables, for the regression analyses. White was used as the reference group.

Predictors: Connectives.

Connectives types. As described earlier, Halliday and Hasan's (1976) four major categories of connectives were coded by hand:

• additive (e.g., *furthermore*, or *in addition*);

• temporal (e.g., *then*, *after that, finally*);

• causal (e.g., *therefore*, *consequently,* or *as a result*);

• adversative (e.g., *however*, *nevertheless*, or *despite this*).

Connectives were coded by T-unit and then calculated as frequency counts per essay. For reliability, I trained an undergraduate research assistant, and then we dual-coded 20% of the data. We had 98% agreement on the connectives codes. I coded the remaining 80% of the data and the research assistant checked these codes. Subsequent differences were resolved through discussion. Frequency counts of connective types were standardized for the regression model.

Connectives sophistication. I created a measure of sophistication of connectives as an exploratory variable. Based on Graham, Harris, and Mason's (2005) approach to holistic scoring of writing quality, the connectives sophistication score was an essay-level measure encompassing presence of connectives, logic or purpose, clarity, and variety, while not weighing one factor more than another. (See Appendix C for the rubric.) A score of 1 signified *weak* overall use of connectives; a score of 2 signified a *developing* overall use of connectives; and a score of 3 signified a *strong* use of connectives. For reliability and to achieve a more nuanced score, the following procedure was used. After training an undergraduate research assistant in the use of the rubric, 20% of the essays were dual coded for sophistication of connectives, resulting in a .96 Kappa. The remaining 80% of the essays were then split in half, and I coded one half while the undergraduate research assistant coded the other half. We then checked each other's codes and subsequent differences were resolved through discussion. A ranking procedure per Wijekumar, Graham, Harris, Barkel, Aitken, Ray, & Houston (2018) was then applied to the essays. Here we (the undergraduate research assistant and I) worked together to order the entire

essay corpus within each score range. That is, all of the essays scored as a 1 for connectives sophistication were then ranked as either 1-, 1, or 1+, and the same for the scores of 2 and 3. These ranked scores were transposed to a range of 0-8. For example, a score of 2+ became a score of 5. Standardized scores for connectives sophistication were used in the regression model.

Reading comprehension. The fall administration of the ETS-developed, Global Integrated Scenario-based Assessment (GISA; O'Reilly, Weeks, Sabatini, Halderman, & Steinberg, 2014) was used as a measure of deep reading comprehension. The GISA is psychometrically robust and demonstrates strong concurrent validity with more conventional reading comprehension measures. The GISA assesses a range of 21st century skills that go beyond low-inference comprehension measures. Scores are reported on a common scale based on a large-scale study conducted by ETS. Standardized scores for the GISA measure were used in the regression model.

Argument types. Argument types were coded as described earlier, based on Kuhn and Crowell's (2011) coding scheme. One argument type was assigned to each T-unit as follows: non arguments were coded 0; own side only arguments were coded 1; dual perspective arguments were coded 2; and integrative perspective arguments were coded 3. Integrative perspective arguments were often expressed as 2 T-units (then coded as 1 T-unit), unless the student's reasoning was fully expressed within 1 T-unit with a dependent clause, perhaps. Argument codes were tabulated as frequency counts per essay. For reliability, a trained undergraduate research assistant and I dual coded 20% of the essays, achieving a Kappa of .83. I coded the remaining essays, followed by the research assistant checking all of the codes. Any differences were resolved through discussion. (See Appendix D for the coding scheme with examples.)

Outcome: Argument sophistication.

Argument sophistication. Raw scores for the highest argument used in each essay were used as the measure for argument sophistication.

Analytic plan

Descriptive statistics were generated for the GISA reading measure, essay length, and for essay measures analyzing arguments and connectives. Pairwise comparison of means were conducted to explore possible differences by subgroups. Correlational analysis results, along with calculations of intraclass coefficients (rho statistic), informed the construction of the multiple regression analyses. Multiple regression analyses, with cluster robust standard errors to account for the essays nested within students, were constructed to answer the research questions investigating the relations with reading comprehension, connectives sophistication, and connectives categories on argument sophistication. Control variables will be discussed in the results section.

<div align="center">

Results

</div>

Research Question 1. Uses of arguments

Percent of cases (essays) containing at least one incidence of each argument type were first computed. At least one non-argument was present in 86% of the essays across the corpus. Own side only arguments appeared at least once in 74% of the essays. The more complex arguments, dual perspective and integrative, were also present in the essays. Dual perspective arguments were present at least once in 53% of the essays, and integrative perspective arguments were present in 27% of the essays.

Table 3.2 presents means, standard deviations, and ranges for essay length and the argument types disaggregated by grade and for the sample as a whole. The mean length across the sample was 5 T-units. Essay length ranged from 1 T-unit to 14 T-units. Pairwise comparison of means revealed that 5th grade essays were significantly longer compared with 4th grade essays ($t = 2.84$; $p < .01$). The mean of non-arguments in the essays was 1.81 and ranged from 0 to 7 across all essays. The mean of own side only arguments was 1.85 and ranged from 0 to 8 across all essays. The mean of dual perspective arguments was .99 and ranged from 0 to 7, while the mean of integrative perspective arguments was .34 and ranged from 0 to 3, across the sample. There were significant grade differences for own side only arguments and integrative perspective arguments; fifth grade essays contained more own side only arguments ($t = 5.29$; $p < .001$) and more integrative perspective arguments ($t = 3.91$; $p < .001$) than did fourth grade essays. Regarding sociodemographic variables, there were gender differences as well as differences by FARL qualification in the essays. Females wrote longer essays than did males ($t = 4.39$; $p < .001$). Females also included more own side only arguments in their essays than did males ($t = 2.58$; $p < .05$). Essays written by students qualifying for FARL were shorter compared with those written by students who did not qualify ($t = -2.78$; $p < .01$). Essays written by students qualifying for FARL also contained fewer dual perspective arguments ($t = -2.36$; $p < .05$) and fewer integrative perspective arguments ($t = -1.98$; $p < .05$) than did essays written by students who did not qualify. The sample contained too few students classified as ELL or classified as special education status to warrant meaningful comparisons.

There were several significant, within-grade topic differences regarding essay length and argument types. Essays on the fourth grade Language topic (i.e., Bilingual or English only at school?) were longer ($t = 2.16$; $p < .05$), had more non-arguments ($t = 2.51$; $p < .05$), and more

integrative perspective arguments ($t = 2.73$; $p < .01$) than did essays on the fourth grade Responsibility topic (i.e., Should we share responsibility for each other's behavior in school?). Essays on the fifth grade Land Use topic (i.e., Should we turn donated land into a mall or a park?) were shorter ($t = -2.04$; $p < .05$), had fewer dual perspective arguments ($t = -2.45$; $p < .05$), and fewer integrative perspective arguments ($t = -3.78$; $p < .001$) than did essays on the fifth grade Oregon topic (i.e., Who is responsible for cleaning up Oregon's shoreline?). In addition, essays on the fifth grade Privacy topic (i.e., Should we give up privacy for security?) had more integrative perspective arguments ($t = 3.15$; $p < .01$) than did essays on the fifth grade Land Use topic (i.e., Should we turn donated land into a mall or a park?).

Table 3.2

Descriptive statistics for essay length and argument types ($n = 198$)

Essay Features	M(SD)	4th grade M(SD)	5th grade M(SD)
Length			
Essay length in T-units	5.00(2.15)	4.44(1.79)	5.56(2.34)
Argument features[a]			
No argument	1.81(1.29)	1.93(1.25)	1.70(1.33)
Own side only	1.85(1.71)	1.25(1.43)	2.45(1.75)
Dual perspective	.99(1.27)	1.09(1.18)	.89(1.35)
Integrative perspective	.34(.64)	.17(.43)	.52(.76)

[a]Argument features are frequency counts per essay

Table 3.3 presents descriptive statistics for the argument sophistication measure (i.e., the highest argument used in the essay). Pairwise comparison of means revealed a significant grade difference; fifth grade essays scored significantly higher on this measure than did 4th grade essays ($t = 2.84$; $p < .01$). There were no gender or SES differences. There were two within-grade topic differences for argument sophistication among fifth grade essays. Essays written on the fifth grade Land Use topic (i.e., Should we turn donated land into a mall or a park?) had lower argument sophistication scores than did essays written on the fifth grade Oregon topic (i.e., Who is responsible for cleaning up Oregon's shoreline?) ($t = -4.25$; $p < .001$). In addition, essays

written on the fifth grade Privacy topic (i.e., Should we give up privacy for security?) demonstrated higher argument sophistication scores than did essays written on the fifth grade Land Use topic (i.e., Should we turn donated land into a mall or a park?) ($t = 4.25$; $p < .001$).

Table 3.3

Means, standard deviations, and ranges for the highest argument used in essays ($n = 198$)

Variable	M(sd)	Min.	Max.
Highest argument			
Full sample	1.89(.84)	0	3
4th grade	1.73(.81)	0	3
5th grade	2.06(.84)	1	3

To explore the variability of argument measures and essay length due to the nested nature of these data, *rho* statistics (intraclass correlations) were computed (see Table 3.4). The presence of own side only arguments in the essays was above 10% within teachers and within topics. The presence of dual perspective arguments was above 10% within students. The presence of dual perspective arguments exceeded 10% within students. The presence of integrative perspective arguments approached or exceeded 10% within students, within teachers, and within topics. Notably, the feature of essay length was the most stable within students at 37%, but also approached 10% within teachers and within topics. These results confirmed the need to control for essay length in the regression model. Clustering the essays at the student level (i.e., cluster robust standard errors) was selected for the regression model as well, because the research questions are primarily concerned with the student level.

Table 3.4

Intraclass correlations for argument measures and essay length ($n = 198$)

Variable	Rho(se) within students	Rho(se) within teachers	Rho(se) within topics
Non argument	.07(.08)	.01(.02)	.02(.03)
Own side only	.05(.07)	.12(.08)	.11(.07)
Dual perspective	.16(.08)	.04(.04)	.02(.03)
Integrative perspective	.14(.08)	.09(.06)	.16(.09)
Highest argument	.05(.07)	.05(.05)	.13(.08)
Total T-units	.37(.08)	.08(.06)	.08(.06)

Research Question 2. Relation between reading comprehension and argument strength

This second research question asks whether reading comprehension is associated with argument sophistication in these fourth and fifth grade students' essays. Descriptives for performance on the GISA measure are presented in Table 3.5. Pairwise comparison of means revealed no significant differences for the GISA scores by grade.

Table 3.5

Means, standard deviations, and ranges for performance on the GISA reading comprehension measure ($n = 66$)

Variable	M(sd)	Min.	Max.
GISA			
Full sample	984.61 (52.73)	875	1111
4th grade	977.97 (59.39)	875	1111
5th grade	991.24 (44.42)	897	1078

Table 3.6 presents pairwise correlations between control variables and argument sophistication (coded as the highest argument used in the essay). Note that there were no significant correlations between gender and argument sophistication or FARL and argument sophistication. Numbers of students with ELL status or special education status were very small, and thus were not included in the correlations or regressions. With the exception of the white and black ethnicity variables, the correlations between each variable in the table and argument sophistication were significantly greater than 0 ($p < .05$). There were weak but significant positive correlations between the reading comprehension and argument sophistication, as well as between essay length and argument sophistication. Reading comprehension was the independent variable of interest in the subsequent model, and this correlation with length confirmed a need to control for length in the regression. Fourth grade had a weak and significant negative correlation with argument sophistication, while fifth grade had a weak and significant positive correlation with argument sophistication. The correlations among grade (and the earlier findings regarding grade as well as topic differences) revealed a need to control for these variables in the regression

67

models. However, when a dummy variable for 5th grade and dummy variables for topics were all

included in the model, one topic was omitted in the output due to collinearity. This led to the

conclusion to control for topic instead of grade in the regression models, with the fourth grade

Responsibility topic as the reference group. Since each grade wrote on different topics, this

would help to elucidate more nuanced grade differences than only using grade as a control. It

should be noted, though, that topics are likely confounded with grade. There was a weak and

significant negative correlation between Latinx students and argument sophistication, and there

were weak and significant positive correlations between Asian and mixed race/other students and

argument sophistication. Thus, the ethnicity variables were included in the subsequent regression

models, with white as the reference group.

Table 3.6

Pairwise correlations among argument sophistication, reading comprehension, essay features, and student characteristics

	1	2	3	4	5	6	7	8	9
1: Highest Argument	___								
2: Reading Comp	.178*								
3: Length	.301***	.211**							
4: 4th grade	-.199**	-.126	-.259***						
5: 5th grade	.199**	.126	.259***	-1.000					
6: White	-.017	-.013	.020	.126	-.126				
7: Black	-.038	-.207**	-.175*	.000	.000	-.535***			
8: Latinx	-.159*	.106	-.037	.000	-.000	-.356***	-.333***		
9: Asian	.208**	.256***	.295***	-.105	.105	-.239***	-.224**	-.149*	
10: Mixed Race/Other	.163*	-.061	.014	-.177*	.177*	-.134	-.125	-.083	-.056

*$p < .05$. **$p < .01$. ***$p < .001$.

Cluster robust standard errors were used in the regression, to account for the nesting of

essays within students. Raw scores were used for the argument sophistication dependent

variable, and standardized scores were used for the continuous variables of reading

comprehension and essay length. Table 3.7 displays results of this model, which explained 27%

of the variance in argument sophistication in the essays. Reading comprehension was positively associated with argument sophistication ($p < .05$), such that essays written by students with a GISA score 1 *SD* above the mean scored .108 higher on argument sophistication. Essays written on the fifth grade Oregon topic were positively associated with argument sophistication ($p < .05$), such that these essays scored .49 higher on argument sophistication in comparison with the fourth grade Responsibility topic. Essays written on the fifth grade Privacy topic were positively associated with argument sophistication ($p < .01$), such that these essays scored .521 higher on argument sophistication in comparison with the fourth grade Responsibility topic. Essay length as measured in total T-units was positively associated with argument strength, such that essays with a length 1 *SD* above the mean scored .065 higher on argument sophistication. Two of the ethnicity variables were positively associated with argument sophistication, in comparison with white as the reference group. Asian students scored .311 higher on argument sophistication, and students identifying as mixed race or other scored .720 higher on argument sophistication, in comparison with white students. There were no significant interactions between the variables in this model.

Table 3.7

Regression Model for Argument Sophistication: Relation with Reading Comprehension ($n = 198$)

Outcome var	Ind. Variables	ß	Robust Standard Error	p	R^2
Highest Arg	Reading comp[a]	.108*	.052	.041	
	Uniforms 4th	.064	.153	.676	
	Language 4th	.171	.210	.416	
	Oregon 5th	.490*	.194	.014	
	Land Use 5th	-.261	.151	.089	
	Privacy 5th	.521**	.163	.002	
	Total T-units[a]	.140*	.062	.028	
	Black	.027	.114	.813	
	Latino/a	-.290	.179	.110	
	Asian	.311*	.149	.041	
	Mixed/other	.720***	.144	.000	
	Constant	1.723***	.116	.000	
					.27***

*$p < .05$. **$p < .01$. ***$p < .001$.
[a]These are standardized variables.

Research Question 3. Uses of connectives

Percent of cases (essays) containing at least one incidence of each connective type were computed, revealing that at least one additive connective was present in 57% of cases; at least one adversative connective was present in 24% of cases; at least one causal connective was present in 47% of cases; and at least one temporal connective was present in 29% of cases.

Table 3.8 displays means, standard deviations, and ranges for each connective type for the sample disaggregated by grade and for the sample as a whole. Across all essays, the mean of total use of connectives was 2.66 and ranged from 0 to 9 connectives. Additive connectives were the most prevalent, with a mean of .99 across the sample and a range of 0 to 5. Adversative, causal, and temporal connectives each had a mean of less than 1 per essay; adversative connectives averaged .29, causal averaged .82, and temporal averaged .56. The largest ranges here were 0 to 5, and considering the mean essay length was 5 T-units, this would suggest at the maximum of 5 connectives, each T-unit in that essay would have begun with a connective to signal the forthcoming idea. Pairwise comparison of means revealed significant grade differences for each connective type, but not among total use of connectives. Fifth grade essays contained significantly more additive connectives ($t = 3.17$; $p < .01$) and more adversative connectives ($t = 3.35$; $p < .001$) than did fourth grade essays. However, fifth grade essays contained significantly fewer causal connectives ($t = -3.60$; $p < .001$) and temporal connectives ($t = -4.55$; $p < .001$) than did fourth grade essays. There was one gender difference; that is, essays produced by females contained more adversative connectives ($t = 2.43$; $p < .05$) than did essays produced by males. There were no significant differences in uses of connectives by FARL eligibility. Differences were not tested by language status or special education status, since these student characteristics represented very small percentages of the sample (2% and 5%, respectively).

There were several significant, within-grade topic differences in students' uses of connectives. Essays on the fourth grade Uniforms topic (i.e., Should students wear school uniforms?) contained more total connectives ($t = 4.47$; $p < .001$) than did the essays on the fourth grade Responsibility topic (i.e., Should we share responsibility for each other's behavior in school?). The fourth grade Uniforms topic essays also contained more causal ($t = 3.58$; $p < .001$) and more temporal connectives ($t = 5.61$; $p < .001$) than did the fourth grade Responsibility topic essays. Essays on the fourth grade Language topic (i.e., Should everyone learn a second language?) also contained more total connectives ($t = 3.93$; $p < .001$), more causal connectives ($t = 3.22$; $p < .01$), and more temporal connectives ($t = 5.21$; $p < .001$) than did the fourth grade Responsibility topic essays. Essays on the fifth grade Privacy topic (i.e., Should we give up privacy for security?) contained more additive connectives ($t = 2.16$; $p < .05$) than did essays on the fifth grade Oregon topic (i.e., Who is responsible for cleaning up Oregon's shoreline?). The fifth grade Privacy topic essays also contained more adversative connectives ($t = 3.76$; $p < .001$) than did the essays on the fifth grade Land Use topic (i.e., Should we turn donated land into a mall or a park?).

Table 3.8

Descriptive Statistics for Connectives ($n = 198$)

Connectives									
	Full sample			*4th grade*			*5th grade*		
	M(SD)	Min.	Max.	M(SD)	Min.	Max.	M(SD)	Min.	Max.
Additive	.99(1.12)	0	5	.75(.90)	0	4	1.24(1.27)	0	5
Adversative	.29(.57)	0	3	.16(.42)	0	2	.42(.66)	0	3
Causal	.82(1.10)	0	5	1.09(1.24)	0	5	.55(.86)	0	5
Temporal	.56(1.05)	0	5	.88(1.27)	0	5	.23(.62)	0	3
Total	2.66(2.19)	0	9	2.88(2.40)	0	8	2.44(1.94)	0	9

Results for the connectives sophistication measure are presented in Table 3.9. The mean score for fourth grade essays was 2.98 and scores ranged from 0 to 7. The mean of fifth grade essays was 2.71 and scores ranged from 0 to 7. These scores were not statistically significantly

different. The mean of the full sample was 2.84 and scores ranged from 0 to 7. In addition, there

were no differences on the connectives sophistication measure by gender or FARL qualification.

There were two topic differences within fourth grade, but no topic differences in fifth grade.

Specifically, essays on the fourth grade Uniforms topic (i.e., Should students wear school

uniforms?) scored higher than did the essays on the fourth grade Responsibility topic (i.e.,

Should we share responsibility for each other's behavior in school?) ($t = 3.98$; $p < .001$). Essays

on the fourth grade Language topic (i.e., Should everyone learn a second language?) also scored

higher than did the essays on the fourth grade Responsibility topic ($t = 3.64$; $p < .001$).

Table 3.9

Frequencies and percentages of connectives sophistication scores ($n = 198$)

Connectives Sophistication score	4th Grade		5th Grade	
	Frequency	Percent	Frequency	Percent
0	20	20.20	10	10.10
1	17	17.17	38	38.38
2	6	6.06	4	4.04
3	3	3.03	--	--
4	31	31.31	30	30.30
5	10	10.10	6	6.06
6	1	1.01	5	5.05
7	11	11.11	6	6.06
Total	99	100	99	100

Rho statistics for use of connectives and the connectives sophistication measure within

students, teachers, and topics are presented in Table 3.10. Additive connectives and connectives

sophistication were more stable within students ($rho > .10$), while adversative, causal, temporal,

and total uses of connectives were more stable within topic ($rho > .10$). The only connective type

that was at least 10% stable within teacher was temporal connectives.

Table 3.10

Intraclass correlations for use of connectives and connectives sophistication (n = 198)

Variable	Rho(se) within students	Rho(se) within teachers	Rho(se) within topics
Additive	.16(.08)	.02(.03)	.07(.05)
Adversative	.08(.08)	.03(.04)	.14(.08)
Causal	.05(.07)	.06(.05)	.11(.07)
Temporal	.01(.07)	.11(.07)	.23(.11)
Total connectives	.05(.07)	0	.10(.07)
Connectives sophistication	.11(.07)	0	.08(.06)

Research Question 4a: Relation of connectives sophistication and argument sophistication

This fourth research question asks whether connectives sophistication (4a) and connective types (4b) are associated with argument sophistication in the essays. First, Pearson correlations are reported (see Table 3.11) for argument sophistication, reading comprehension, essay length, grade, and measures of connectives. There was a weak but significant positive correlation between connectives sophistication and argument sophistication. Fourth grade had a weak and significant negative correlation with argument sophistication, while fifth grade had a weak and significant positive correlation with argument sophistication. As with the regression for RQ3, when a dummy variable for fifth grade as well as dummy variables for the topics were included, one topic was omitted in the output due to collinearity. This led to the conclusion to again control for topic instead of grade for RQ4, in order to discover more nuanced differences. (Yet, again, we realize the limitation that grade is likely confounded with topics.) Dummy variables for topics were included, with the fourth grade Responsibility topic as the reference group. Building on the model from Research Question 2 (relation of reading comprehension and argument sophistication), reading comprehension and essay length were included as control variables. Also revealed in Research Question 2, there were no significant correlations between gender or FARL qualification and argument sophistication, and ELL status and special education

status numbers in the sample were too small to warrant including in the correlations and regressions. Dummy variables for ethnicity (with white as the reference group) were again included as controls, and robust standard errors (Huber-White adjustment) were used to account for the nesting of essays within students.

Table 3.11

Pairwise correlations among argument sophistication, connectives measures, and control variables

	1	2	3	4	5	6	7	8	9
1. Highest Argument	____								
2. Reading Comprehension	.178*								
3. Total T-units	.301***	.211**							
4. Connectives Sophistication	.230**	.294***	.516***						
5. 4th grade	-.199**	-.126	-.259***	.082					
6. 5th grade	.199**	.126	.259***	-.082	-1.000				
7. Additive	.053	.272***	.435***	.372***	-.221**	.221**			
8. Adversative	.419***	.023	.284***	.207**	-.233**	.233**	-.078		
9. Causal	.073	.051	.295***	.632***	.249***	-.249***	.069	-.036	
10. Temporal	-.019	.144*	.216**	.574***	.309***	-.309***	-.032	-.105	.467***

*$p < .05$. **$p < .01$. ***$p < .001$.

The regression testing connectives sophistication (see Table 3.12) as the independent variable of interest produced the same amount of variance (27%) as did the model with reading comprehension as the independent variable of interest. However, the results of the coefficients were different, reflecting the inclusion of connectives sophistication to the model. There was not a significant association between connectives sophistication and argument sophistication, but the coefficient indicates that essays with a score of 1 *SD* above the mean on connectives sophistication scored .101 higher on argument sophistication. There was no longer a significant association between reading comprehension and argument sophistication, but the coefficient indicates that a score 1 *SD* higher than the mean on the GISA measure had a .084 higher score on argument sophistication. Essays on the fifth grade Oregon topic and on the fifth grade Privacy

topic (in relation to fourth grade Responsibility topic essays) had similar results as the previous model, as well as the ethnicity variables of Asian and mixed race (in relation to white). There were no significant interactions between any of the variables in this model.

Table 3.12

Regression Model for Argument Sophistication: Relation with Connectives Sophistication (n = 198)

Outcome var	Ind. Variables	ß	Robust Standard Error	p	R^2
Highest Arg	Connect Soph[a]	.101	.060	.100	
	Reading comp[a]	.084	.053	.119	
	Uniforms 4th	-.008	.148	.955	
	Language 4th	.112	.214	.601	
	Oregon 5th	.517**	.192	.009	
	Land use 5th	-.255	.151	.096	
	Privacy 5th	.504**	.167	.004	
	Total T-units[a]	.087	.063	.170	
	Black	.017	.112	.878	
	Latino/a	-.283	.175	.110	
	Asian	.331*	.146	.027	
	Mixed race	.710***	.160	.000	
	Constant	1.743***	.115	.000	
					.27***

*p < .05. **p < .01. ***p < .001.
[a]These are standardized variables.

Research Question 4b. Relation of connective types and argument strength

Next, the connective types were included in the model to explore whether connective types were associated with argument sophistication in the essays. The control variables remained in the model, and robust standard errors were used to account for the nesting of essays within students. As shown previously in Table 3.11, correlations results indicated a moderate and significant positive correlation between adversative connectives and argument sophistication. This regression model (see Table 3.13) accounted for an additional 6% of the variance in argument sophistication in the essays (R^2 = .33). There was not a significant association between connectives sophistication and argument sophistication, but the coefficient indicates that a connectives sophistication score 1 *SD* above the mean had a .098 increase on argument sophistication. Reading comprehension was not significantly associated with argument sophistication, but the coefficient indicates that a GISA reading comprehension score 1 *SD* above

the mean had a .107 increase on argument sophistication. There was a significant positive

association between the fifth grade Privacy topic and argument sophistication, such that essays

written on the Privacy topic scored .405 higher on argument sophistication in comparison with

the fourth grade Responsibility topic. There was no longer a significant association for essay

length and argument sophistication. Among the connective types, adversative connectives were

positively associated with argument sophistication, such that a score 1 *SD* above the mean on

adversative connectives frequency had a .189 increase on argument sophistication. Positive

associations remained for Asian students and students identifying as mixed race and argument

sophistication. Asian students scored .415 higher on argument sophistication in comparison with

white students, and mixed race students scored .640 higher on argument sophistication than

white students. There were no significant interactions between the variables in this model.

Table 3.13

Regression Model for Argument Sophistication: Relation with Connectives (*n* = 198)

Outcome var	Ind. Variables	ß	Robust Standard Error	*p*	R^2
Highest argument	Connectives				
	Soph[a]	.098	.089	.274	
	Reading comp[a]	.107	.055	.054	
	Uniforms 4th	.081	.169	.633	
	Language 4th	.171	.215	.430	
	Oregon 5th	.361	.190	.062	
	Land use 5th	-.185	.153	.232	
	Privacy 5th	.405*	.184	.031	
	Total T-units[a]	.076	.060	.209	
	Additive[a]	-.047	.081	.570	
	Adversative[a]	.189*	.072	.011	
	Causal[a]	.048	.061	.438	
	Temporal[a]	-.107	.071	.133	
	Black	.049	.103	.634	
	Latino/a	-.186	.168	.273	
	Asian	.415**	.154	.009	
	Mixed/other	.640***	.124	.000	
	Constant	1.715***	.119	.000	
					.33***

p < .05. **p* < .01. ***p* < .001.
[a]These are standardized variables.

Discussion

The primary aim of this study was to examine upper elementary students' productive academic language use within two domains: argumentative reasoning and use of connectives to signal specific arguments. The argumentative reasoning coding scheme (Kuhn & Crowell, 2011) proved to be a useful tool for describing students' uses of arguments, as reliability was excellent (Kappa = .83). The research questions investigating these fourth and fifth grade students' uses of arguments and connectives across multiple essays each yielded rich descriptive data. The essays were written in the context of an academic vocabulary intervention (Word Generation), which afforded multiple essays per student. Students most likely only spent about 15-20 minutes writing each essay, but these brief samples provided evidence for a number of different ways these students deployed academic language in their independently written essays. The mean essay length was only 5 T-units across the sample, yet a variety of argument types and connective types were produced in the writing, which points to the possibility that brief, on-demand writing exercises embedded into instruction on contestable topics may be a worthwhile learning activity. Descriptive findings for the argument types revealed that each type of argument appeared at least once across the essay corpus, and notably, the most complex argument type, integrative perspective, appeared in 27% of cases (essays). Dual perspective arguments appeared in 53% of cases, and these arguments are also considered more complex, in that the student addresses an alternative position on the topic. In Kuhn and Crowell's (2011) study of a sixth grade classroom intervention on argumentation over three years, the dual perspective arguments consistently showed up at about 30% for each of the three years, and the integrative perspective arguments did not appear in students' writing until the third year of the intervention. The results of the present study indicated that fifth grade students wrote significantly more of these

integrative perspective arguments, than did the fourth grade students. On average, the fifth grade students wrote essays that were about 1 T-unit longer than the fourth graders, so the finding of more integrative arguments could be due in part to the fifth grade students having written longer essays. However, we did not find any significant interactions with essay length in the study. The findings that some students in the sample produced more complex reasoning in their essays suggests that these argument types could be incorporated into classroom instruction utilizing oral or written argumentation. Another consideration regarding these grade differences is that it is likely that many of these fifth grade students had participated in the Word Generation curriculum the previous school year, if they had attended the same school in the previous year. While the study design does not afford making any causal claims regarding the influence of the Word Generation curriculum on the students' reasoning in their writing, it is still interesting to find that first of all, each type of argument was evidenced across the essay corpus, and secondly, that the fifth grade students wrote longer essays and included more of the integrative perspective arguments. The essay length finding is not a surprising finding, since the research has established grade related increases in syntactic measures of length (Crowhurst, 1980; Crowhurst & Piche, 1979; Loban, 1976). However, the finding of the integrative arguments is somewhat surprising, because researchers argue that this type of counterfactual reasoning does not appear to develop on its own (Kuhn, 1992; Kuhn & Crowell, 2011). Studies have shown, though, that pre-adolescent students are able to address the opposing position in an argument when they are encouraged to do so (Kuhn & Crowell, 2011; Kuhn & Udell, 2007). Thus, a possible interpretation aligns with the argument schema theory, in that students could have started to develop an argument schema. That is, through their participation in a dialogic approach to learning about the weekly controversial topics, they may have started to anticipate that there

would be other perspectives than their own on the topic, and begin to reason about those alternative perspectives. However, these analyses are text-oriented only. Since we do not have data on the actual contexts of implementation, this is a tentative explanation.

We also find that a measure of deep reading comprehension (i.e., Global Integrated Scenario-based Assessment, or GISA; O'Reilly, Weeks, Sabatini, Halderman, & Steinberg, 2014) was associated with argument sophistication in the essays prior to including any connectives variables in the model. This finding is not surprising, given that extensive prior research has documented relations between reading and writing (e.g., Fitzgerald & Shanahan, 2000; Graham et al., 2018). In the next model, while still controlling for topic, essay length, and ethnicity, when we included independent variables to explore the linguistic feature of connectives, reading comprehension no longer had a significant association with the outcome of argumentative reasoning. This model used a researcher-created assessment of connectives sophistication, and the overall variance in the model remained the same ($R^2 = .27$), but connectives sophistication was not significant and reading comprehension no longer had a significant association with reasoning. The connectives sophistication assessment attempted to capture variety and maturity of use of the connectives in an essay. The students' overall use of connectives may have somewhat supported their argument sophistication, but it wasn't until the final model that a clearer picture began to develop. All of the earlier controls remained, while the frequencies of connective types used in the essays were added to the model. Because adversative connectives were significantly associated with argument sophistication in the essays (and the inclusion of connective types increased the variance in the model by 6%), this suggests that an awareness of the specific purposes and functions of connectives may serve as linguistic support for such writing activities in grades four and five.

Conclusion

There are limitations to keep in mind when interpreting the results of the current study. Importantly, the sample is small and so the findings may not generalize to other populations. The cross-sectional and correlational study design does not afford developmental or causal explanations. In addition, although this study involved written essays, the outcome of interest was *reasoning* as opposed to overall quality of written argumentation; therefore, implications for instruction should be interpreted with caution. There is also considerable variation left unexplained in the regression models, as the overall variance explaining the outcome of argumentative reasoning in all models was less than 35%. However, given that the outcome was a fine-grained measure of reasoning at the T-unit level and finding that specific linguistic support helped to increase the variance, this is a helpful starting point.

Despite these limitations, further research in these areas of elementary argumentative writing could still inform writing instruction and assessment. The exact coding schemes used in the study are not likely to be readily transferable to on the ground classroom assessment for teachers; however, these types of detailed, quantitative analyses of argumentative writing are encouraged, for they could lead to insights that could later inform instructional strategies for teaching argumentation and argumentative writing. As for applications to classroom assessment tools, this study could inform aspects of classroom use of analytic rubrics for argumentative writing; that is, instead of relying on common, holistic writing rubrics in the classroom, teachers may consider a more analytic rubric that provides levels of insight for different domains. For example, rubrics that separately demonstrate levels of reasoning and linguistic skills or academic language that support argumentative writing, may help students become aware of relative strengths and weaknesses regarding reasoning and the linguistic support to signal their

arguments. In this way, the rubric itself may serve as an instructional scaffold as students learn the features argumentative writing as well as some of the functions of linguistic skills for argumentative writing. Future studies with even more writing samples per student could afford growth modeling to offer longitudinal insight. Finally, while an over-emphasis on connectives is not recommended, future iterations of the Word Generation curriculum (or other academic vocabulary programs) could embed instruction of classes of connectives as part of the academic vocabulary target words. Currently, a connective may appear as one of the five weekly target words, but it would be useful for an intervention study to explore explicit instruction of connectives within the curriculum.

References

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

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading, 36*(3), 241-260. doi: 10.1111/j.1467-9817.2011.01501.x

Crowhurst, M. (1980). Syntactic complexity in narration and argument at three grade levels. *Canadian Journal of Education, 5*(1), 6-13.

Crowhurst, M., & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English, 13*(2), 101-109.

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14,* 175-187.

Cummins, J. (1981). Age on arrival and immigrant second-language learning in Canada: A reassessment. *Applied Linguistics, 2,* 132-149.

Dobbs, C. L. (2014). Signaling organization and stance: Academic language use in middle grade persuasive writing. *Reading and Writing, 27,* 1327-1352.

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and instruction. *Reading and Writing, 32,* 1345-1357.

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

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*(2)*,* 207-241. doi:10.1016/j.cedpsych.2004.08.001

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., & Aitken, A. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88*(2), 243-284.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Longman Group Ltd.

Hu, C., & Li, Y. (2015). Discourse connectives in L1 and L2 argumentative writing. *Higher Education Studies, 5*(4), 30-41. doi:10.5539/hes.v5n4p30

Hunt, K. W. (1965). A synopsis of clause-to-sentence length factors. *The English Journal, 54*(4), 300+305-309.

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

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

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

LaRusso, M., Jones, S. M., Kim, H. Y., Kim, J., Donovan, S., & Snow, C. (2016). Impacts of a discussion-based academic language program on classroom interactions in 4th through 7th grades. SREE Conference paper.

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of

    implementation of an academic vocabulary intervention for linguistically diverse students

    in urban middle schools. *Reading Research Quarterly, 45*(2), 196-228.

Llosa, L., Beck, S. W., & Zhao, C. G. (2011). An investigation of academic writing in secondary

    schools to inform the development of diagnostic classroom assessments. *Assessing*

    Writing, 16, 256-273.

Loban, W. (1976). *Language development: Kindergarten through grade twelve* (Research Report

    No. 18). Urbana, IL: National Council of Teachers of English.

Olinghouse, N. G., Santangelo, T., & Wilson, J. (2012). Examining the validity of single-

    occasion, single-genre, holistically scored writing assessments. In E. Van Steendam & M.

    Tillema (Eds.), *Measuring writing: Recent insights into theory, methodology, and*

    practices (pp. 55-82). Bingley, UK: Emerald Group Publishing Limited.

O'Reilly, T., Weeks, J., Sabatini, J., Halderman, L., & Sternberg, J. (2014). Designing reading

    comprehension assessments for reading interventions: How a theoretically motivated

    assessment can serve as an outcome measure. *Educational Psychology Review, 26*(3),

    403-424.

Reznitskaya, A., & Anderson, R. C. (2006). Analyzing argumentation in rich, natural contexts.

    Informal Logic, 26(2), 175-198.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S.

    (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2-3),

    155-175. http://dx.doi.org/10.1080/0163853X.2001.9651596

Reznitskaya, A., Kuo, L, Glina, M., & Anderson, R. C. (2009). Measuring argumentative

    reasoning: What's behind the numbers? *Learning and Individual Differences, 19,* 219-

    224.

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language

    among urban middle school students. *Journal of Research on Educational Effectiveness,*

    2(4), 325-344. doi: 10.1080/19345740903167042

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N.

    Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge:

    Cambridge University Press. doi: 10.1017/CBO9780511609664.008

Taylor, K. S., Lawrence, J. F., Connor, C. M., & Snow, C. E. (2019). Cognitive and linguistic

    features of adolescent argumentative writing: Do connectives signal more complex

    reasoning? *Reading and Writing, 32,* 983-1007. https://doi.org/10.1007/s11145-018-

    9898-6

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners:

    An intervention study. *Reading and Writing, 22*(9), 993-1019.

Townsend, D., Filippini, A., Collins, P., & Biancarosa, G. (2012). Evidence for the importance

    of academic word knowledge for the academic achievement of diverse middle school

    students. *The Elementary School Journal, 112*(3), 497-518.

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015a).

    Core academic language skills: An expanded operational construct and a novel

    instrument to chart school-relevant language proficiency in preadolescent and adolescent

    learners. *Applied Psycholinguistics, 36,* 1077-1109. doi:10.1017/S014271641400006X

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015b). Beyond

vocabulary: Exploring cross-disciplinary academic-language proficiency and its

association with reading comprehension. *Reading Research Quarterly, 50*(3), 337-356.

doi:10.1002/rrq.104

Wijekumar, K., Graham, S., Harris, K. R., Lei, P., Barkel, A., Aitken, A., Ray, A., & Houston, J.

(2018). The roles of writing knowledge, motivation, strategic behaviors, and skills in

predicting elementary students' persuasive writing from source material. *Reading and*

Writing. https://doi.org/10.1007/s11145-018-9836-7

CHAPTER 4

Study 3. Cognitive and linguistic features of adolescent argumentative writing: Do connectives signal more complex reasoning?

The present study focuses on middle school students' argumentative reasoning and use of connectives (e.g., *although, however*) in a corpus of argumentative essays (n = 158). We explored the frequencies of specific types of arguments and the proportions of classes of connectives that the students used in their essays. The Argument Schema Theory (Reznitskaya & Anderson, 2002; Reznitskaya, Anderson, McNurlen, Nguyen-Jahiel, Archodidou, & Kim, 2001) suggests we can gain insights into the complexity of students' developing reasoning skills by examining the sophisticated use of key discourse moves in their writing. We also investigated the association between the students' uses of connectives (a discourse move) and their uses of specific arguments as well as with overall argument sophistication. By better understanding the development of the cognitive and linguistic features of academic language produced by adolescents, we can design more effective instruction and assessment of argumentative writing. In this section, we describe our conceptual framework. We first explain the academic language and literacy challenges of argumentative writing for early adolescents. We then offer a rationale for assessing complex reasoning in argumentative writing and present a range of specific reasoning features used in argumentative writing. Finally, we review major word classes of connectives as possible linguistic levers to improve students' argumentative writing.

Language and literacy challenges of adolescent argumentative writing

Among the most important 21st Century skills is the ability to take a reasoned point of view and defend it. This demands cognitively sophisticated thinking and reasoning skills, which is made more difficult when these generative ideas have to be translated into written form and

follow the specific rules of the genre (Martin, 2009; Rose, 2009; Schleppegrell, 2004). Unfortunately, too many adolescents fail to achieve even basic writing proficiency (National Center for Education Statistics, 2012), and thus are unable to generate argumentative texts that are coherent and meaningful. Indeed, adolescents' literacy achievement overall did not increase over three decades (Carnegie Council on Advancing Adolescent Literacy, 2010), and only 27% of students at both 8th grade and 12th grade scored at or above the proficient level on the most recently reported National Assessment of Educational Progress (NAEP) writing assessment (National Center for Educational Statistics, 2012). At the same time, cognitively and linguistically demanding argumentative texts represent advanced literacy tasks that students are expected to comprehend as well as replicate in writing. For example, the Common Core State Standards (Common Core State Standards Initiative, 2010) feature argumentative writing across the curriculum in grades 4 through 12. These writing outcomes are concerning, given that language and literacy skills that express such complex reasoning as considering others' viewpoints and critically examining one's own viewpoints are needed not only for college and career readiness, but also to nurture students toward contributing to a healthy society (see Duhaylongsod, Snow, Selman, & Donovan, 2015, for example). Facility with argumentation and skills of lexical precision such as using connectives to signal those arguments are challenging aspects of learning the *academic language* needed as leverage for cross-content area reading and writing (Snow & Uccelli, 2009). Upper elementary and middle school students are still developing these particular academic language skills (Uccelli et al., 2015a, b).

Research has shown that early adolescence can be a period of responsiveness to instruction on argumentative reasoning (e.g., Kuhn & Crowell, 2011) as well as argumentative writing (Andrews, Torgerson, Low, & McGuinn, 2009). In addition, early adolescence has been

emphasized as an important time period in which to study the features of academic language (Uccelli et al., 2015a, b; Uccelli & Snow, 2008), since school-based texts and types of writing call for the understanding and production of an academic register (as opposed to more colloquial language). However, many students are likely to only be beginning to develop these skills, and there is little evidence of instruction on analytic writing in middle school (Lawrence, Galloway, Yim, & Lin, 2013). Understanding the development of students' reasoning as a means of supporting students' facility with the advanced language and literacy skills needed for academic writing is an important aim for education research and practice.

Argumentative writing in applied education settings

Although the terms *argument* or *argumentation* likely conjure up images of aggressive or oppositional talk or writings, this type of discourse was not the basis or goal of our study. Rather, the nature of argumentation that we attempt to describe is in line with Andriessen's (2006) notion of *collaborative argumentation*, or *arguing to learn*, which is modeled after the type of problem solving or resolution seen within arguments occurring in science as a discipline. And while the students' individual reasoning within their argumentative essays is the focus of our study, the collaborative and dialogic instructional setting of the Word Generation academic vocabulary program also influenced our overall framework for the study.

Why are middle school students' (grades 6-8) on-demand, argumentative essays a good proxy for assessing their complex reasoning? The Argument Schema Theory (Reznitskaya & Anderson, 2002; Reznitskaya et al., 2001) and other dialogic argumentation models (Crowell & Kuhn, 2014; Kuhn & Udell, 2003) ground our approach. In these models, the focus of assessment is not on the argumentation during group or dialogic arguments, but rather on the development of an individual's argument schema *after* these interactions and represented in

writing. The current study's learning environment itself was squarely focused on regular classroom discussions and debates about weekly controversial topics, as part of a broader intervention focused on middle school students' academic vocabulary learning in an effort to bolster deep reading comprehension. This aligns well with the Argument Schema Theory and dialogic models of argumentation, because the practice of engaging in regular dialogic discussions preceded the unit-culminating essays on controversial topics, and thus we could investigate the essays for evidence of specific argumentative moves.

A further premise is that a sample that includes several essays written by each student enables the use of an analytic scheme to understand the participants' argumentative reasoning. With this methodology (Reznitskaya & Anderson, 2006; Reznitskaya, Kuo, Glina, & Anderson, 2009), each idea unit (we use T-units; Hunt, 1965) within each essay is categorized as incorporating one specific type of argument at a time. We adapted Kuhn and Crowell's (2011) coding scheme for our analytic scoring, because it assesses levels of argumentative sophistication and incorporates argumentative moves that address the opposing argument (counterarguments). We describe the coding scheme below and provide examples from our data in Appendix E. Having a fine-grained measure of argumentation at the level of the T-unit allowed us to investigate two similar outcomes. First, we could examine the relations of the classes of connectives with frequencies of specific types of arguments utilized in the essays. Second, we could use a single score of the highest level of argument type used in each essay as an outcome of argument sophistication. This allowed us to examine the relations of connectives with overall argument sophistication. Connectives were features of language use in the students' writing that signaled various cohesive functions, such as adding on information or marking opposing viewpoints.

Reasoning moves that display the relative complexity of specific arguments

Detailed analytic scoring of argumentative writing can reliably provide information about the specific argumentative moves within a piece of writing (Reznitskaya et al., 2009). The coding scheme from Kuhn and Crowell's (2011) classroom argumentation intervention among sixth graders used this type of analytic scoring. Instead of providing one score to encapsulate the overall argumentation displayed in a piece of writing, they found that the individual ideas (i.e., *T-units*, Hunt, 1965) within students' essays could be reliably coded in one of four distinct categories: non-arguments (if the idea was irrelevant or repetitive, for example); own side only arguments (if the idea simply supported their own position); dual perspective arguments (if the idea offered a negative view of the opposing position); and integrative perspective arguments (if the idea included a negative about the favored position or a positive about the opposing position). Dual perspective and integrative perspective arguments were characterized as more complex reasoning. Investigating the more fine-grained aspects of reasoning incorporated into adolescent argumentative writing can help us to better understand the nature and development of this reasoning.

Prior research documents a *myside bias* in young adolescents (as cited in Wolfe & Britt, 2008); that is, students tend to only support their own position in an argument instead of considering what the opposing arguments might be. However, young adolescents were found to be able to attend to the opposing position in an argument when explicitly asked to do so (Kuhn & Udell, 2007) and upon regular participation in classroom dialogic argumentation instruction (Kuhn & Crowell, 2011). Since middle school students have thus shown facility with deploying complex reasoning skills in their writing, and because students in the current study were involved in ongoing classroom discussion and debate activities that created the opportunity to

develop an argument schema, we expected to see evidence of the more sophisticated arguments in their unit-culminating writing samples.

Connectives as potential leverage for argumentative writing

An inventory of academic language features informed by linguistic and educational research includes the feature of lexical precision (Snow & Uccelli, 2009). For example, proficiency with understanding and choosing appropriate connectives (e.g., *because, nonetheless*) may help to equip young adolescents with the knowledge and skills they need as they encounter more challenging texts (Crosson & Lesaux, 2013a, b; Uccelli et al., 2015a, b). Connectives are cohesive devices within a text that "explicitly signal the connections between passages of text" (Hu & Li, 2015, p. 30). Connectives also serve various functions between passages of text; major categories include connectives that signal additive, temporal, causal, and adversative relationships (Halliday & Hasan, 1976). The category of additive connectives includes words or phrases such as: *furthermore*, or *in addition*. Temporal connectives include such examples as: *then*, *after that, finally*. Causal connectives could include: *therefore*, *consequently,* or *as a result.* Finally, adversative connectives could be expressed as: *however*, *nevertheless*, or *despite this* (Halliday & Hasan, 1976).

Whereas prior research has shown that knowledge of connectives is a challenge and can predict success in reading comprehension (Crosson & Lesaux, 2013b), this study extends this research by exploring whether students' productive use of connectives in academic writing is associated with more complex argumentation. Such an association would have implications for instruction in and assessment of argumentative writing. A natural language processing technology tool that explicitly examines the use of connectives in a text facilitates assessment of these connectives. Known as the Tool for the Automated Analysis of Cohesion (TAACO;

Crossley, Kyle, & McNamara, 2016), the TAACO can identify students' production of these major classes of connectives in their writing. The TAACO calculates a proportion of each type of connective used within each essay (i.e., a count of each connective type over the essay's total number of words), so the scores provide a density measure for each connective type per essay.

Research questions

The primary purpose of this study is to examine the frequencies of cognitive features (argument types) and the densities of linguistic features (connectives) in the persuasive writing of middle school students, and to explore the relations of connective uses with specific arguments and with overall strength of argumentation. We predict that adversative connectives will be a stronger predictor of the more sophisticated arguments (dual perspective and integrative perspective) than the other categories of connectives assessed (additive, causal, and temporal), because adversative connectives may appear to signal an opposing idea or argument. Thus, we also expect to see a relation between adversative connectives and the overall argument sophistication in these middle school students' essays. We addressed the following research questions in the present study:

1. To what extent do middle school students utilize arguments and non-arguments in their argumentative essays on a variety of topics? To what extent do they employ connectives in their essays?

2. Controlling for essay length and topic type, is the use of connectives associated with specific argument types in the essays?

3. Controlling for essay length and topic type, is the use of connectives associated with overall argument sophistication in middle school students' argumentative essays?

Method

Research context

The data for this study were collected from three middle schools (grades 6-8) in an urban West Coast school district during the 2011-2012 academic year. The schools were participating in the second year of a larger study (Lawrence, Francis, Paré-Blagoev, & Snow, 2017) of the adolescent literacy curriculum called Word Generation (Donovan, Snow, & Daro, 2013). The evaluation of the intervention focused on general academic vocabulary acquisition and reading comprehension; as part of the curriculum, each of the 24 weeklong units culminated with engaging students in an independently written argumentative essay. The same Word Generation curriculum was used with all of the sixth – eighth grade students within each school, so students in a range of grades responded to the same essay prompts in their workbooks.

Participants

Table 4.1 presents the frequency distribution of student demographic characteristics for the analytic sample (n = 40) disaggregated by grade and for the sample as a whole. A corpus of 150-175 total essays was desired, and students' essays (and hence their participation in the study) were included in the data corpus if there were at least 3 essays from the 4 units being analyzed. Because student demographic information was not initially known, there was not an even distribution of participants according to background characteristics.

The participant sample for this study consisted of 6 sixth graders, 16 seventh graders, and 18 eighth graders, for a total of 40 students. Participants came from diverse racial/ethnic, socioeconomic, and linguistic backgrounds. The majority of students identified themselves as Chinese (52.5%); the next-largest group comprised students who identified themselves as White (22.5%); and a minority of students identified themselves as Hispanic/Latino (7.5%), Vietnamese

(5%), Black/African American (2.5%), or Japanese (2.5%). Three students' (7.5%) ethnicities were not identified. Fifty-five percent of the participants were eligible for free and reduced price lunch.

Table 4.1

Student Demographic Characteristics

Demographic	6th grade n(% of grade)	7th grade n(% of grade)	8th grade n(% of grade)	All students n(% of sample)
Total students	6 (15)	16 (40)	18 (45)	40 (100)
Gender				
Female	2 (33.3)	2 (12.5)	10 (55.6)	14 (35)
Male	4 (66.7)	14 (87.5)	8 (44.4)	26 (65)
Ethnicity				
Chinese	4 (66.7)	8 (50)	9 (50)	21 (52.5)
White	1 (16.7)	5 (31.3)	3 (16.7)	9 (22.5)
Hispanic/Latino		1 (6.3)	2 (11.1)	3 (7.5)
Vietnamese	1 (16.7)	1 (6.3)		2 (5)
Black/African American		1 (6.3)		1 (2.5)
Japanese			1 (5.6)	1 (2.5)
Missing			3 (16.7)	3 (7.5)
English proficiency designation				
English only (EO)	3 (50)	5 (31.3)	7 (38.9)	15 (37.5)
Initially English fluent (IFEP)	1 (16.7)		3 (16.7)	4 (10)
Redesignated Eng. fluent (RFEP)	2 (33.3)	8 (50)	6 (33.3)	16 (40)
Limited English fluent (LEP)		3 (18.8)	2 (11.1)	5 (12.5)
Socioeconomic status				
Not eligible for free or reduced price lunch	2 (33.3)	7 (43.8)	9 (50)	18 (45)
Eligible for free or reduced price lunch	4 (66.7)	9 (56.3)	9 (50)	22 (55)
CST achievement level[a]				
Advanced	5 (83.3)	5 (33.3)	9 (50)	19 (48.7)
Proficient	1 (16.7)	5 (33.3)	7 (38.9)	13 (33.3)
Basic		2 (13.3)	1 (5.6)	3 (7.7)
Below basic		2 (13.3)		2 (5.1)
Far below basic		1 (6.7)	1 (5.6)	2 (5.1)

a. CST data were not available for 1 student

The sample was also linguistically diverse. The school district provided detailed information regarding the students' English language proficiency status, according to the designations used by California at the time. While states in the U.S. approach criteria for classification into and exit from English language learner status differently (National Research Council, 2011), California's recommended procedure identified monolingual students whose

home language is English as English only (EO); and for students for whom a language (or languages) other than English is spoken, a home language survey and an initial English language proficiency assessment (i.e., the California English Language Development Test, or CELDT) may identify students as Initially English Fluent Proficient (IFEP) or Limited English Proficient (LEP). LEP students whose yearly evaluation process yields an overall proficiency level adequate for reclassification are identified as Reclassified English Fluent Proficient (RFEP). Four criteria were considered for LEP students' reclassification: a score of "English proficient" on the CELDT; California Standards Test (CST) scores in English-Language Arts; teacher evaluation; and parental input (National Research Council, 2011). In our sample, 37.5% were classified as EO, 10% were IFEP, 40% were RFEP, and 12.5% were LEP. On the English-language arts standards tested in the California Standards Test (CST, the state's standardized test at the time), the majority of students in the sample scored in the *Basic*, *Proficient*, or *Advanced* levels with 7.7% of students scored in the Basic range, 33.3% in the Proficient range, and 48.7% in the Advanced range. However, the CST achievement levels do not represent an equated scale across grades and so can only be used as a within-grade measurement. In addition, it should be noted that the term "LEP" has been replaced with the more common term "English learner" or "EL."

Data corpus

The Word Generation vocabulary curriculum consisted of 24 weekly units, each of which was organized around a contestable question. Four essay topics from the second 12 weeks of the program were selected for analysis (see Table 4.2 for the list of units and topics). All students completed the four essays except two were missing the essay for Unit 20, resulting in a corpus of 158 essays. Students were familiar with the essay topics, because each topic had been discussed

and debated in different content area classrooms on the four days prior to the writing assignment. The essays were brief, on-demand writing activities; teachers did not provide feedback on the essays nor did students engage in revision activities. Each essay topic had the same instructions: *TAKE A STAND. Support your position with clear reasons and specific examples.*

Table 4.2

Topics/Essay Prompts Represented in the Data Corpus

Week number	Number of essays	Essay prompt
13	40	Is the death penalty justified?
16	40	Who is responsible for teen smoking?
18	40	Should drugs be legalized?
20	38	Who is to blame for high school dropouts?

Data analysis

Students' handwritten essays were transcribed as Microsoft Word documents to facilitate the argumentation coding process. The transcribed essays were also converted into plain text files in order to use the TAACO (Crossley et al., 2016) for the automated analysis of the linguistic features (connectives). The following measures were then generated to analyze the data.

Covariates: Essay features and student characteristics.

Essay length. Essays were segmented into T-units in order to have a basic measure of essay length. A T-unit was comprised of an independent clause or an independent clause with any dependent clauses (Hunt, 1965). Trained research assistants segmented the essays into T-units. Segmented essays were verified by the lead researcher. Essay length was calculated as the total number of T-units per essay.

Binary topics. There were two topic types for the essays: binary and open ended. The binary topics variable indicates that the topic elicited a yes or no stance (binary topics = 1) or not (binary topics = 0; i.e., the topic was open-ended and thus generated a stance that could include

multiple entities). Because there is some evidence that topic type and prompts can influence aspects of writing (Lawrence, Niiya, & Warschauer, 2015), we controlled for topic type in our analyses. The two binary essay topics related to whether the death penalty is justified and whether drugs should be legalized. The two open-ended topics related to who is responsible for teen smoking and who is to blame for high school dropouts.

Gender. To control for gender in the analysis, we included a dummy variable to indicate students who were female (female = 1) or male (female = 0).

Socioeconomic status (SES). We used students' free and reduced price lunch eligibility as a measure of students' SES. A dummy variable was created to indicate students who were eligible to receive free and reduced price lunch (SES = 1) and those who were not (SES = 0).

Predictors: Connectives.

Connectives were measured using the text analysis tool known as the Tool for the Automatic Analysis of Cohesion, version 1.0 (TAACO; Crossley et al., 2016). The TAACO batch-processes the text files and calculates a proportion of each type of connective used in each essay. In essence, it is a density measure of each particular type of connective: the frequency of each use of a connective type is the numerator, and the denominator is the number of words per essay. The connective indices included in the TAACO are theoretically and rhetorically based. As described earlier, the four connective types selected for this analysis were based on Halliday and Hasan's (1976) four major categories of connectives: additive, adversative, causal, and temporal. The following are exact titles of the indices we used from the TAACO, along with corresponding examples: "[all_additive-bk]" for additive connectives: *again, also, further, furthermore, in addition, moreover*; "[negativelogical-bk]" for adversative connectives: *alternatively, although, however, in contrast, notwithstanding that, on the other hand*;

"[all_causal-bk]" for causal connectives: *because, consequently, hence, since, therefore, thus*; and "[all_temporal-bk]" for temporal connectives: *after, before, finally, first, next, while*. We did not control for basic connectives (e.g., *for, and, nor*) so that our analyses would include all of students' functional uses of each of our connectives of interest.

Outcomes: Argument types and argument sophistication.

Argument types. Argument types were assessed in order to measure the effect of students' uses of connectives on each type of argumentative reasoning represented in the essays. We coded the argument types based on Kuhn and Crowell's (2011) coding scheme applied to sixth grade persuasive writing tasks (see Appendix E for the coding scheme and examples from our data). Each T-unit was assigned an argument code. The following four argument categories were distinguished in the present study: *non-argument* was coded "0" if the T-unit stated a position only with no support, was unclear, or was a repeat of an earlier argument; *own side only* argument was coded "1" if the T-unit offered only positive support for the favored position; *dual perspective* argument was coded "2" if the T-unit offered a negative relevant to the opposing position; and *integrative perspective* argument was coded "3" if the T-unit included a negative about the favored position or a positive of the opposing position. Integrative perspective arguments were usually expressed in two T-units, or occasionally in one T-unit if the student was able to fully express his or her reasoning by including a dependent clause. If the student took two T-units to express an integrative perspective argument, it was coded as a single unit. As previously described, dual perspective arguments and integrative perspective arguments were characterized as complex reasoning, versus own side only arguments demonstrating a more simplistic "myside" tendency and non-arguments displaying no argumentative reasoning (or no new argumentative reasoning if a repeat argument). Argument codes were calculated as

frequency counts per essay, and then used as individual continuous outcomes to explore the contributions of connectives to each type of argument or non-argument.

Trained research assistants dual coded 25% of the data in order to assess interrater reliability on the argument types. Cohen's kappa ranged from .78 to .93 across the four topics, indicating high levels of reliability. Discrepancies were subsequently resolved through discussion. The lead researcher coded the remaining essays.

Argument sophistication. This outcome also concerns the students' argumentative reasoning expressed in the essays, but the difference with this dependent variable is that as a global outcome (i.e., a single score of the highest level of argument used in the essays), we can go beyond the affordances of the previous outcomes and explore the predictability of connectives on overall argument sophistication. The goal of the argument sophistication outcome was not to assess the argumentative structure or writing quality of the entire essay, but instead to capture an overall measure of the most sophisticated level of argumentative reasoning used within each essay.

Analytic plan

Descriptive statistics were generated for the measures of essay length, arguments, and connectives. Post hoc pairwise comparisons of means (Tukey honestly significant difference tests if more than two groups) were conducted in order to test for differences in the observed variables in the essays across grade levels as well as to test for differences in calculated variables as a function of demographics. Subsequent correlational analyses, including calculations of intraclass correlation coefficients (rho statistic), informed the construction of the regression analyses. We used the Huber-White adjustment (Huber, 1967; White, 1980) in the regressions to account for the nesting of essays within students. Regression analyses with essay length and

essay topic type as essay feature covariates were conducted to explore the predictive power of connectives on each argument type. The essay feature covariates, along with the student characteristics of gender and SES as controls, were then included in a final regression analysis exploring the predictive power of connectives on the overall argument sophistication in the essays.

Results

Essay length, arguments, and connectives: Descriptive and correlational analyses

Our first research question addressed the extent to which students employed argument types and connectives across the essay corpus. For the purposes of analyzing basic presence of arguments, we calculated the percent of essays employing at least one incidence of each argument code. The non-argument code was present in 97% of essays; the own side only code appeared in 86% of essays; the dual perspective code was present in 50% of essays; and the integrative perspective code was present in 42% of essays. Table 4.3 shows the descriptive statistics for length, arguments, and connectives for essays across the entire sample and disaggregated by grade. On average, essays contained nearly ten T-units per essay (*mean* = 9.77), with a standard deviation of 5.58. Essays displayed about five no argument statements per essay (*mean* = 5.15), with a standard deviation of 4.24. Essays exhibited, on average, nearly three own side only arguments per essay (*mean* = 2.77, *sd* = 2.46), a little more than one dual perspective argument per essay (*mean* = 1.33, *sd* = 2.09), and less than one integrative perspective argument per essay (*mean* = .52, *sd* = .68). The mean proportions of connectives were low: additive connectives comprised 5% of each essay, adversative connectives comprised 1% of each essay,

causal connectives comprised 2% of each essay, and temporal connectives comprised 1% of each essay.

Table 4.3

Descriptive Statistics for Arguments and Linguistic Features per Essay ($n = 158$)

Essay Features	M(SD)	6th grade M(SD)	7th grade M(SD)	8th grade M(SD)
Global feature				
Essay length in T-units	9.77(5.58)	6.54(2.00)	10.77(5.83)	9.97(5.83)
Argument features[a]				
No argument	5.15(4.24)	3.17(1.83)	5.98(4.58)	5.1(4.34)
Own side only	2.77(2.46)	2.00(1.41)	2.81(2.51)	2.99(2.65)
Dual perspective	1.33(2.09)	.96(1.60)	1.53(1.85)	1.28(2.09)
Integrative perspective	.52(.68)	.42(.72)	.45(.65)	.61(.70)
Linguistic features[b]	.05(.02)	.04(.02)	.05(.03)	.05(.02)
Additive connectives				
Adversative connectives	.01(.01)	.01(.01)	.01(.01)	.01(.01)
Causal connectives	.02(.02)	.02(.02)	.03(.02)	.02(.01)
Temporal connectives	.01(.01)	.01(.01)	.01(.01)	.01(.01)

a. Argument features are frequency counts per essay
b. Connectives are proportions of connectives over total words

For the purpose of approximating a global measure of argument sophistication, we created a variable indicating the highest argument used in each essay. Table 4.4 shows the descriptive statistics for this measure. On average, essays exhibited a score close to a dual perspective argument as the highest argument utilized (*mean* = 2.09), with sixth grade essays scoring below the mean (1.88) and seventh and eighth grade essays scoring above the mean (2.06 and 2.19, respectively).

Table 4.4

Descriptive Statistics for Highest Argument Used ($n = 158$)

Essay Features	M (SD)	6th grade M (SD)	7th grade M (SD)	8th grade M (SD)
Essay feature				
Highest argument used	2.09 (.87)	1.88 (.85)	2.06 (.87)	2.19 (.88)

In order to examine possible differences among the observed variables of essay length, argument types, connectives, and the highest level of argument used, we conducted pairwise comparisons of means (Tukey's honestly significantly different test if more than two groups).

There were significant mean differences in essay length by grade (seventh grade essays contained more T-units than sixth grade essays, $t = 3.24$; $p < .004$, and eighth grade essays contained more T-units than sixth grade essays, $t = 2.68$; $p < .02$). There were also differences for essay length or argument types according to gender and SES. Essays written by females contained more integrative perspective arguments than males ($t = 1.95$; $p < .05$). Essays written by students eligible for free and reduced price lunch (FRL) contained fewer total T-units ($t = -4.04$; $p < .000$), fewer non-argument statements ($t = -3.31$; $p < .001$), and fewer integrative perspective arguments ($t = -2.17$; $p < .03$) than those who were not eligible. However, there were no mean differences according to language proficiency designation for essay length or argument types. For classes of connectives, the only significant difference according to our grade and sociodemographic variables was that essays written by eighth grade students contained a lower proportion of causal connectives than essays written by seventh grade students ($t = -.2.74$; $p < .007$). For the argument sophistication variable, there were no significant differences according to grade or language proficiency. However, essays written by females exhibited stronger overall argumentation than males ($t = 2.86$; $p < .005$), and essays written by students who qualified for FRL exhibited weaker overall argumentation than those who did not ($t = -2.66$; $p < .009$).

Not surprisingly, differences on essay length according to grade and student characteristics signaled the need to control for essay length in our subsequent regression analyses, in addition to the control of topic type discussed above. Due to few differences found on the argument and connectives variables of interest, we decided to only use essay length and topic type as our controls and not introduce any demographic controls to the regression analyses for our second research question, which explored the predictability of connectives on specific argument types. However, the significant differences in overall argument sophistication for

gender and SES indicated the need to control for these student characteristics in the regression for our third research question, which examined the predictability of connectives on overall argument sophistication. Thus, in light of the small sample size (< 20 students per grade) and the few differences for our variables of interest, we decided to pool the data for further analyses as opposed to incorporating multilevel models.

Correlational analyses were conducted in order to explore common variance among the measures of essay length, topic type, arguments, and connectives. As shown in Table 4.5, significant correlations among the essay length and binary topics variables confirmed the need to enter these variables as controls in our subsequent regression analyses. It is also worth noting that dual perspective arguments were significantly negatively correlated with own side only arguments, indicating that essays with more dual perspective arguments contained fewer own side only arguments. In addition, adversative connectives were significantly positively correlated with integrative perspective arguments, indicating a co-occurrence of adversative connectives with the most complex type of reasoning.

Table 4.5

Correlations Among Essay Features Measured

	1	2	3	4	5	6	7	8	9	10
1. Total T-units	1									
2. Binary topics dummy var.	.052	1								
3. Non-arguments	.889***	-.033	1							
4. OSO arguments	.444***	-.296***	.162*	1						
5. DP arguments	.312***	.577***	.154	-.356***	1					
6. IP arguments	.176*	.065	.030	.031	.048	1				
7. Additive connectives	.126	-.149	.043	.201*	-.011	.068	1			
8. Causal connectives	.030	.112	-.024	-.013	.160*	-.014	-.032	1		
9. Adversative connectives	.080	.083	.110	-.066	-.030	.290***	.293***	-.119	1	
10. Temporal connectives	.001	.130	-.033	.035	.063	-.092	.037	-.099	-.084	1

*p < .05. ***p < .001.

In order to examine the variability of essay length, arguments, and connectives within students, we calculated intraclass correlation coefficients (ICCs). Essays were nested within students, so it is expected that more variance on some variables may be explained at the student level than the essay level. Table 4.6 reports the ICCs (rho statistic), or the fraction of the variance at the student level for the variables. Not surprisingly, the essay feature of length was highly stable within students, with 67% of the variance on total number of T-units explained at the student level. Use of non-argument statements in essays was also highly stable within students, with 56% of the variance explained at the student level. The arguments and connectives were more variable within students (rhos = 0 - .25), so we decided to account for the nesting of essays within students using the Huber-White adjustment (robust standard errors) in subsequent regressions rather than a full multilevel model.

Table 4.6

Intraclass Correlations (*n* = 158)

Essay Features	Rho Coefficient	Standard Error
Global features		
Essay length in words	.75	.05
Essay length in T-units	.67	.06
Argument features		
No argument	.56	.08
Own side only arguments	.09	.08
Dual perspective arguments	0	
Integrative perspective arguments	.18	.08
Linguistic features		
Additive connectives	.25	.09
Adversative connectives	.03	.07
Causal connectives	.11	.08
Temporal connectives	.06	.07

Relations of connectives to specific arguments

Our second research question examined the contribution of connectives to specific

argument types. Using multiple linear regression analyses, we modeled the predictability of

connectives on argument types separately, controlling for essay length and topic type and using

robust standard errors (Huber-White adjustment). As shown in Table 4.7, the control predictors

of essay length and binary topics were statistically significant in relation to non-arguments, own

side only arguments, and dual perspective arguments, but not in relation to integrative

perspective arguments. Additive connectives significantly negatively predicted non-arguments (β

= -18.369; *p* = .03), with a significant overall model that explained 81% of the variance in the

essays. Connectives did not significantly predict own side only arguments, but the overall model

was significant, explaining 32% of the variance in the essays. Adversative connectives

significantly negatively predicted dual perspective arguments (β = -23.679; *p* = .05), with a

significant overall model that explained 43% of the variance in the essays. On the other hand,

adversative connectives significantly positively predicted integrative perspective arguments (β =

20.13; p = .006), with a significant overall model that explained 11% of the variance in the

essays. Causal and temporal connectives did not predict argument types.

Table 4.7

Regression Models by Argument Type: Contribution of Connectives to Arguments, Controlling for Essay Length and Binary Topics (n = 158)

Argument type	Variable	ß	Robust Standard Error	p	R^2
Non arguments	Total T-units	.686	.034	<.0001	
	Binary topics	-.809	.32	.015	
	Additive	-18.369	8.168	.03	
	Causal	-8.374	7.395	.264	
	Adversative	32.32	18.959	.096	
	Temporal	-5.108	9.38	.589	
	Constant	-.291	.431	.504	
					.81***
OSO arguments	Total T-units	.199	.04	<.0001	
	Binary topics	-1.466	.387	.001	
	Additive	13.029	8.597	.138	
	Causal	.596	8.671	.946	
	Adversative	-28.77	19.316	.144	
	Temporal	11.916	8.574	.172	
	Constant	1.05	.529	.054	
					.32***
DP arguments	Total T-units	.097	.02	<.0001	
	Binary topics	2.221	.263	<.0001	
	Additive	5.993	4.542	.195	
	Causal	7.735	5.945	.201	
	Adversative	-23.679	11.696	.05	
	Temporal	-2.717	6.868	.695	
	Constant	-.957	.314	.004	
					.43***
IP arguments	Total T-units	.019	.011	.08	
	Binary topics	.054	.118	.647	
	Additive	-.653	2.605	.804	
	Causal	.044	2.772	.987	
	Adversative	20.13	6.931	.006	
	Temporal	-4.09	3.885	.299	
	Constant	.198	.168	.246	
					.11*

*p < .05. ***p < .001.

Relations of connectives to argument sophistication

Our third research question examined whether these middle school students' uses of

connectives were related to overall argument sophistication in their essays on multiple

controversial topics. Here we were interested in the most sophisticated type of argumentative

reasoning expressed as a cognitive outcome of the writing, as opposed to a global measure of argumentative structure or writing quality. In our regression model, we entered essay length and binary topics as essay feature controls and gender and SES (i.e., eligible for FRL) as student characteristic controls. The four classes of connectives (additive, adversative, causal, and temporal) were our independent variables of interest. As displayed in Table 4.8, linear regression results indicate that these variables explain 24% of the variability in students' argument sophistication in the essays, accounting for the nesting of essays within student. Essay length was not a significant predictor. Binary topics were significant positive predictors of argument sophistication ($\beta = .514$; $p = .001$). Essays written by female students were significantly positively related to argument sophistication ($\beta = .438$; $p = .002$). As for connectives, adversative connectives significantly positively predicted argument sophistication ($\beta = 17.251$; $p < .021$).

Table 4.8

Predictors of Overall Argument Sophistication

	Overall argument sophistication			
Variable	ß	Robust Standard Error	p	R^2
Essay length	.015	.013	.239	
Binary topics	.514**	.144	.001	
Female	.438**	.129	.002	
SES	-.272	.146	.069	
Additive	-2.806	2.713	.307	
Adversative	17.251*	7.163	.021	
Causal	3.351	3.206	.302	
Temporal	-3.902	4.360	.376	
Constant	1.622***	.265	.000	
				.24***

*$p < .05$. **$p < .01$. ***$p < .001$.

Discussion

We begin with a summary of our results. In this study, a diverse sample of middle school students ($n = 40$) produced a corpus of 158 persuasive essays. Researchers coded the essays for students' uses of argument types, and the Tool for the Automated Analysis of Cohesion (TAACO; Crossley et al., 2016) calculated students' uses of connectives in the essays.

Descriptive analyses revealed that all types of arguments (i.e., non-argument, own side only, dual perspective, and integrative perspective) were attempted at least once across all essays, and complex reasoning moves were even present to a good extent. Low proportions of the classes of connectives (i.e., additive, adversative, causal, and temporal) were employed across all essays. Intraclass correlations revealed that essay length was quite stable within students, whereas there was more within student variability for uses of arguments and connectives. Further, separate regression analyses revealed significant relations among connectives and argument types in the essays: the use of additive connectives was negatively related to non-arguments, the use of adversative connectives was negatively related to dual perspective arguments, and the use of adversative connectives was positively related to integrative perspective arguments. Finally, a regression analysis testing the predictability of connectives on argument sophistication, over and above essay length, topic, gender, and SES, revealed a significant positive relation with adversative connectives.

Our results concur with Kuhn and Crowell's (2011) general finding that young adolescents are capable of producing complex argumentative moves in their writing and that these arguments can be identified and quantitatively assessed. Whereas our study did not afford a developmental design, previous developmental studies in applied education settings have claimed that young adolescents may not routinely develop the ability to attend to the opposing position in an argument, but that they are capable of doing so when prompted (Kuhn & Udell, 2007) or when provided with targeted educational interventions on argumentation (Kuhn & Crowell, 2011). The Kuhn and Crowell (2011) study followed an experimental and comparison group of sixth graders across three years, finding that in year three the experimental group produced significantly more total arguments than the comparison group on the topic of whether

teacher pay should be equal or experience-based, even after accounting for essays becoming longer over time in both conditions. In addition, integrative perspective arguments did not appear until the third year, and only among the experimental group. It is noteworthy, then, that in our corpus of 158 essays across four different topics, these middle school participants employed dual perspective reasoning and integrative perspective reasoning in a considerable number of essays. Dual perspective arguments appeared at least once in 50% of the essays, and integrative perspective arguments appeared at least once in 42% of the essays. We are not able to make claims about the mechanisms of consolidation following the Word Generation learning activities; however, we are intrigued by the evidence of individual complex argumentative reasoning, despite the brief amount of time dedicated to writing. The instructional time spent on Word Generation – including the unit-culminating on-demand essay – was only 15-20 minutes per day. Finding a range of complexity in students' written arguments suggests that students were attempting complex reasoning entirely on their own by evaluating others' and their own viewpoints, following classroom discussion and debate on the topics. Although our study did not assess the classroom discourse during the Word Generation activities, this finding supports the application of the Argument Schema Theory (Reznitskaya & Anderson, 2002; Reznitskaya et al., 2001) as a useful theoretical foundation for moving forward with this line of research which investigates arguments produced in an instructional environment emphasizing dialogic argumentation prior to written argumentation.

Beyond the basic presence of specific argumentative moves, this research explored the use of the linguistic feature of connectives in the essays, finding significant relations between classes of connectives and specific types of arguments. Few studies have examined the use of connectives in argumentative writing for young adolescents. An exception is Crowhurst (1987),

who compared sixth, tenth, and twelfth grade students' uses of connectives in an argumentative writing sample with uses of connectives in a narrative writing sample. Across both genres and all three grades, she found very low mean percentages of connectives employed: 4.4% additive, 2.0% adversative, 1.4% causal, 1.5% temporal, and 0.6% continuative. Indeed, our corpus analysis also reports quite low proportions of use within these categories of connectives, although we did not measure continuative connectives. Across the sixth, seventh, and eighth grade Word Generation essays, students wrote the following mean proportions of connectives: 5% additive, 1% adversative, 2% causal, and 1% temporal. One of the ways our results extend these findings on connectives produced in argumentative writing is that connectives emerged as significant predictors of specific arguments when the argument type was the outcome in separate regressions.

First, a negative relation between additive connectives and non-argument statements emerged, indicating that the stronger the presence of additive connectives, the fewer the non-argument statements in the essays. This is not surprising, given that the function of additive connectives is to *add on* information. Again, T-units were coded as non-arguments if they did not contain a specific type of reasoning or if they repeated an argument, so essays that used additive connectives appeared to signal a weaker presence of non-substantive arguments, which implies perhaps a more intentional use of adding on more substantive arguments of one type or another. Next, a negative relation between adversative connectives and dual perspective arguments indicated that the stronger the presence of adversative connectives, the fewer the dual perspective arguments in the essays. Alternatively, a positive relation between adversative connectives and integrative perspective arguments indicated that the stronger the presence of adversative connectives, the greater the presence of integrative perspective arguments in the

111

essays. For the two types of more complex reasoning (dual perspective and integrative arguments), adversative connectives appeared to signal the most sophisticated of these types of reasoning (integrative perspective arguments), but did not seem to signal dual perspective arguments. These middle school students were thus able to address the opposing position to their favored position on a controversial topic perhaps without employing the major category of adversative connectives. However, when it came to using language in ways that would integrate the two opposing positions, they did incorporate adversative connectives to signal this complex idea. This association between adversative connectives and more complex reasoning in a sample of middle school writers is an initial finding in research examining these two particular dimensions together in early adolescent writing. Unfortunately, we do not have measures of students' prior knowledge of connectives nor did we assess whether the students *correctly* utilized connectives in their essays, but we conjecture as a starting point that students were attempting to signal adversative functions in the language of their integrative arguments. Prior research suggests that the signaling language and the complex reasoning are likely a bidirectional relationship (Cook-Gumperz & Gumperz, 1992), so we cannot claim whether it is access to connectives that pushes students into thinking more integratively or whether the integrative thinking they are already doing is simply expressed with this language feature.

In addition, the nested nature of our essay corpus afforded calculating the intraclass correlation coefficients (ICCs) for our variables, in order to see the extent to which variables were explained simply due to the nesting of essays within students. Our results demonstrate within student variability regarding arguments and connectives (rhos = 0 - .25), while the measure of essay length was unsurprisingly rather stable within students (rho = .67). Aside from the ICC for non-arguments (rho = .56), students' uses of the arguments and connectives were

rather variable across essays. This suggests that students may be employing cognitive and linguistic features in different ways across different topics. Though the current study does not provide any definitive answers about the within student variability for arguments and connectives, one possible explanation is that the essay topic itself may matter. Students may employ these features differently, depending on aspects of the topic or prompt itself, or perhaps depending on their knowledge of the topic. Our data did not include a measure of prior knowledge, but it was hoped that repeated exposure to the topic and practice discussing the issues on each of the days leading up to the weekly essays would somewhat account for background knowledge.

Based on prior research (Lawrence et al., 2015), we controlled for the type of topic when testing the relations among arguments and connectives, as well as in the final regression testing the predictability of connectives on overall argument sophistication. Binary topics (i.e., Should drugs be legalized?, Is the death penalty justified?) were found to significantly negatively predict non-arguments and the more simplistic, own side only arguments, as well as to significantly positively predict the more complex arguments, dual perspective arguments. Topic type was not significantly related to the most sophisticated type of argument, integrative perspective (though the overall model including the topic control was significant); however, binary topics positively predicted overall argument sophistication in the final regression model. Similarly, Al-Adeimi's (2018) study of the influence of classroom discussion on fourth through seventh grade Word Generation essays found that contestable discussion questions positively predicted persuasive essay scores, whereas semi-open discussion questions negatively predicted persuasive essay scores. In the current study, then, it may be that students were more likely to communicate stronger stances when the nature of the writing prompt encouraged a binary, or yes/no response.

On the other hand, when the prompt was open-ended, students may have been more apt to simply note various own side only arguments for a range of potential stances. For example, the open-ended prompts in this study asked "who is responsible" for the issues of teen smoking or dropping out of school. Rather than delving into multiple arguments for one strong position potentially afforded by the binary prompts, the students may have explored multiple own side only reasons attributed to why the parents are responsible, the schools are responsible, *and* why the teens themselves are responsible. Thus, further investigation of topic effects and of students' language use within multiple essay topics is a fruitful area for argumentative writing research.

Limitations and future research

There are several limitations of this study to address as well as future directions for research. First, it should be noted that the findings of the present study are limited to a small sample size of students ($n = 40$), and may not generalize to other samples of middle school students. The correlational nature of these analyses also limits the utility of these findings in explicating causal mechanisms accounting for the results. Future research investigating the relation of connectives to argument types and argument sophistication in persuasive writing should also be conducted with longitudinal as well as classroom intervention methods.

Methodologically, the question may be raised about the lack of an overall measure of argumentative writing structure or writing quality. It should be noted that the analytic approach we used to assess the argument types might account for writing quality to some extent, since repetitive arguments or unclear statements were coded as non-arguments. In addition, analyzing the fine-grained cognitive and linguistic features in the essays was the intentional design of the study. It was not our goal to investigate the global argumentative structure, but rather to explore this methodology affording quantifying types of arguments students produce within a brief

academic text. Nonetheless, a previous study has investigated overall writing quality among fifth grade students' persuasive essays produced in the Word Generation program. Mancilla-Martinez (2010) incorporated a researcher-designed measure of writing quality and documented students' improvement in writing quality, particularly in the latter ten weeks of the curriculum.

The efficacy of the natural language processing tool (i.e., TAACO) for assessing students' uses of connectives must also be viewed with caution, in that it produced a simple proportion of connectives used over the total number of words in the essay. The TAACO effects calculations of connectives based on Halliday and Hasan's (1976) classic categories of additive, adversative, causal, and temporal; however, the tool does not allow us to search and further compare specific usage within categories. The TAACO was chosen as a starting point for simultaneously studying arguments and connectives. However, future studies of middle school students' uses of connectives in their academic writing should compare hand-coded measures of connective usage with automated tools such as the TAACO. Hand coding the connectives would reveal which exact connectives students utilize within each category, and whether they rely on immature or mature uses of those connectives. Knowledge of connectives beyond vocabulary breadth has been shown to uniquely predict reading comprehension (although more so for English only students than English learners) (Crosson & Lesaux, 2013b), and so we speculate that an awareness of how connectives function in multiple perspective arguments could be useful for writing instruction. However, rather than encouraging educators to simply focus on instruction of discrete words and phrases, teaching the overall functions of classes of connectives within the genre of argumentation might be more meaningful. For example, genre-based pedagogy strives to "provide learners with metalinguistic tools to recognise and use the language patterns of the texts they encounter" (Rose, 2009, p. 162). Future classroom intervention studies

testing the instruction of how connectives function within students' multiple perspective arguments could illuminate the utility of such a metalinguistic tool.

Finally, it is noteworthy that for this sample of middle school students' essays, only 24% of the variance in argument sophistication was explained by our controls and the connectives assessed. The large amount of variance still unexplained suggests that perhaps other student characteristics and other features of academic language might also contribute to students' overall argument sophistication. In particular, future research should include assessments of students' vocabulary knowledge, in order to partial out the contribution of vocabulary to students' productive uses of connectives. Despite such limitations, future studies should continue to explore the analytic measurement of argumentative and linguistic features of students' writing, as well as investigate pedagogical interventions, in order to inform young adolescents' development of the academic language needed for academic writing.

Conclusion

Given the emphasis of argumentative reasoning in current educational standards in the US and the importance of young adolescents developing these cognitively sophisticated skills for their futures, the topic we studied is salient. The findings from this study suggest that independent production of complex reasoning within the context of an adolescent academic vocabulary curriculum is attainable, and suggests that the use of adversative connectives relates to complexity of reasoning in their argumentative writing. The correlational design of this study does not enable us to determine whether the relationship of connectives and argumentation is causal, nor to make claims about the Word Generation curriculum's influence on these students' writing. However, finding evidence of complex reasoning in this diverse sample of middle

school students' essays demonstrates the merit of investigating the academic language in these repeated, on-demand writing exercises in a curriculum that engages students in classroom discussion and debate. The academic language contained within written arguments is an area where middle school students are still developing and where more research is needed.

Acknowledgments

References

Al-Adeimi, S. (2018). *Talking to learn: Investigating the relationship between classroom discussion and persuasive writing* (Unpublished doctoral dissertation). Harvard Graduate School of Education, Cambridge, MA.

Andrews, R., Torgerson, C., Low, G., & McGuinn, N. (2009). Teaching argument writing to 7- to 12-year-olds: An international review of the evidence of successful practice. *Cambridge Journal of Education, 39*(3), 291-310. doi:10.1080/03057640903103751

Andriessen, J. (2006). Arguing to learn. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences,* (1st Edition, pp. 443-460). Cambridge: Cambridge University Press.

Carnegie Council on Advancing Adolescent Literacy. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success.* New York, NY: Carnegie Corporation of New York.

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

Cook-Gumperz, J., & Gumperz, J. (1992). Changing views of language in education: The implications for literacy research. In R. Beach, M. Green, M. Kamil, & T. Shanahan (Eds.), *Multidisciplinary perspectives on literacy research*. Urbana, IL: National Council of Teachers of English.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods, 48,* 1227-1237. doi:10.3758/s13428-015-0651-7

Crosson, A. C., & Lesaux, N. K. (2013a). Connectives: Fitting another piece of the vocabulary puzzle. *The Reading Teacher, 67*(3), 193-200. doi: 10.1002/TRTR.1197

Crosson, A. C., & Lesaux, N. K. (2013b). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading, 36*(3), 241-260. doi: 10.1111/j.1467-9817.2011.01501.x

Crowell, A., & Kuhn, D. (2014). Developing dialogic argumentation skills: A 3-year intervention study. *Journal of Cognition and Development, 15*(2), 363-381.

Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English, 21,* 185-201.

Donovan, M. S., Snow, C. E., & Daro, P. (2013). The SERP approach to problem-solving research, development, and implementation. *National Society for the Study of Education Yearbook, 112*(2), 400-425.

Duhaylongsod, L., Snow, C. E., Selman, R. L., & Donovan, M. S. (2015). Toward disciplinary literacy: Dilemmas and challenges in designing history curriculum to support middle school students. *Harvard Educational Review, 85*(4), 587-608.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Longman Group Ltd.

Hu, C., & Li, Y. (2015). Discourse connectives in L1 and L2 argumentative writing. *Higher Education Studies, 5*(4), 30-41. doi:10.5539/hes.v5n4p30

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1,* 221-233. Berkeley, CA: University of California Press.

Hunt, K. W. (1965). A synopsis of clause-to-sentence length factors. *The English Journal, 54*(4), 300+305-309.

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young

    adolescents' thinking. *Psychological Science, 22*(4), 545-552.

    doi:10.1177/0956797611402512

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5),

    1245-1260.

Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking*

    and Reasoning, 13(2), 90-104. doi: 10.1080/13546780600625447

Lawrence, J. F., Francis, D., Paré-Blagoev, J., & Snow, C. E. (2017). The poor get richer:

    Heterogeneity in the efficacy of a school-level intervention for academic language.

    Journal of Research on Educational Effectiveness, 10(4), 767-793.

    doi:10.1080/19345747.2016.1237596

Lawrence, J. L., Galloway, E. P., Yim, S., & Lin, A. (2013). Learning to write in middle school?

    Insights into adolescent writers' instructional experiences across content areas. *Journal of*

    Adolescent and Adult Literacy, 57(2), 151-161. http://doi.org/10.1002/jaal.219

Lawrence, J. F., Niiya, M., & Warschauer, M. (2015). Narrative writing in digital formats:

    Interpreting the impact of audience. *Psychology of Language and Communication, 19*(3),

    201-221. http://dx.doi.org/10.1515/plc-2015-0012

Mancilla-Martinez, J. (2010). Word meanings matter: Cultivating English vocabulary knowledge

    in fifth-grade Spanish-speaking language minority learners. *TESOL Quarterly, 44*(4),

    669-699.

Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and*

    Education, 20, 10-21.

National Center for Educational Statistics. (2012). *The Nation's Report Card: Writing 2011* (NCES 2012-470). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Research Council. (2011). *Allocating federal funds for state programs for English language learners.* Panel to Review Alternative Data Sources for the Limited-English Proficiency Allocation Formula under Title III, Part A, Elementary and Secondary Education Act. Committee on National Statistics and Board on Testing and Assessment. Washington, DC: The National Academies Press.

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In M. Pressley & C. Block (Eds.), *Comprehension instruction: Research-based best practices*, (1st edition, pp. 319-334). New York: Guilford Press.

Reznitskaya, A., & Anderson, R. C. (2006). Analyzing argumentation in rich, natural contexts. *Informal Logic, 26*(2), 175-198.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2-3), 155-175. http://dx.doi.org/10.1080/0163853X.2001.9651596

Reznitskaya, A., Kuo, L, Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual Differences, 19,* 219-224.

Rose, D. (2009). Writing as linguistic mastery: The development of genre-based literacy pedagogy. In D. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *The SAGE handbook of writing development* (pp. 151-166). London: SAGE.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*.

    Mahwah, NJ: Lawrence Erlbaum Associates.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N.

    Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge:

    Cambridge University Press. doi: 10.1017/CBO9780511609664.008

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015a).

    Core academic language skills: An expanded operational construct and a novel

    instrument to chart school-relevant language proficiency in preadolescent and adolescent

    learners. *Applied Psycholinguistics, 36,* 1077-1109. doi:10.1017/S014271641400006X

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015b). Beyond

    vocabulary: Exploring cross-disciplinary academic-language proficiency and its

    association with reading comprehension. *Reading Research Quarterly, 50*(3), 337-356.

    doi:10.1002/rrq.104

Uccelli, P., & Snow, C. E. (2008). A research agenda for educational linguistics. In B. Spolsky

    & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 626-642). Malden,

    MA: Blackwell Publishing.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test

    for heteroskedasticity. *Econometrica, 48,* 817-830.

Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation.

    Thinking and Reasoning, 14(1), 1-27. doi:10.1080/13546780701527674

CHAPTER 5

Summary, Implications, and Conclusion

The primary goal of my dissertation was to understand and further explore the ways children and adolescents use certain features of academic language in their writing, using quantitative methodologies and an educational linguistics perspective. Table 5.1 presents a summary of the key findings, followed by a summary and overarching themes.

Table 5.1
Summary of Key Findings

Chapter	Key Finding
<Chapter 2> Argumentative Reasoning and Connectives in Second and Third Grade Students' Persuasive Essays	1. Percent of cases with at least 1 of each argument type: non-argument (71%); own-side only (72%); dual perspective (36%); integrative perspective (13%). 2. Percent of cases with at least 1 of each connective type: additive (65%); adversative (23%); causal (33%); temporal (7%). (Connectives hand coded as frequency counts) 3. Mean essay length across grades was 6.35 T-units & no significant grade differences. Highest argument mean(sd) score across grades = 1.47(.84). 3rd grade scored significantly higher than did 2nd grade. 4. Reading comprehension and use of adversative connectives were associated with argument sophistication in the essays.
<Chapter 3> Fourth and Fifth Grade Argumentative Writing: An Analysis of Argumentative Reasoning and Connectives	1. Percent of cases with at least 1 of each argument type: non-argument (86%); own side only (74%); dual perspective (53%); integrative perspective (27%). 2. Percent of cases with at least 1 of each connective type: additive (57%); adversative (24%); causal (47%); temporal (29%). (Connectives hand coded as frequency counts) 3. Mean essay length across grades was 5 T-units & 5th grade essays were significantly longer than were 4th grade essays. Highest argument mean(sd) score across grades = 1.89(.84). 5th grade essays scored significantly higher than did 4th grade essays. 4. Reading comprehension, certain 5th grade topics, and use of adversative connectives were associated with argument sophistication in the essays.
<Chapter 4> Cognitive and Linguistic Features of Adolescent Argumentative Writing: Do Connectives Signal More Complex Reasoning?	1. Percent of cases with at least 1 of each argument type: non-argument (97%); own side only (86%); dual perspective (50%); integrative perspective (42%). 2. Mean(sd) proportions of connectives in essays computed by TAACO: additive = .05(.02); adversative = .01(.01); causal = .02(.02); temporal =.01(.01). 3. Mean essay length across grades was 9.77 & 7th and 8th grade essays were both significantly longer than 6th grade essays. Highest argument mean(sd) score across grades = 2.09(.87). There were no significant differences on essay scores between grades.[a] 4. Binary topics, females, and use of adversative connectives were associated with argument sophistication in the essays.

[a]Note: These descriptives for the highest argument scores, though used in the regression in Chapter 4, were inadvertently omitted from the published study.

In Study 1 (Argumentative reasoning and connectives in second and third grade students' persuasive writing), I explored use of an argumentative reasoning coding scheme for writing among 2nd and 3rd grade students' essays (N=385) that was created for a middle school intervention on argumentation. I investigated argument types and use of connectives within the writing samples, and tested for relations among the essay features using a regression approach. The 2nd and 3rd grade sample of students were from high poverty schools in the Southwest US. Each type of argument, as well as each type of connective appeared across the sample, and although the essays tended to be brief, significant relations were found between connectives and argument sophistication. Specifically, adversative connectives were related to argument sophistication, controlling for essay length and grade. In addition, results indicated a significant relation between reading comprehension and argument sophistication.

In Study 2 (Fourth and fifth grade argumentative writing: An analysis of argumentative reasoning and connectives), I examined argument types and use of connectives in a sample of 66 4th and 5th grade students' essays (N=198) using descriptive analyses and a regression approach. The students were from two states in the US participating in the CCDD trial of the academic vocabulary program, Word Generation. Each argument type as well as each connective type appeared across the sample. Essays tended to be brief, but significant relations were found among connectives and argument sophistication. Reading comprehension again had a significant relation with argument sophistication; however, this relation disappeared when connectives measures were added to the model. A measure of sophistication of connectives was also not significant, but the model variance remained the same ($R^2 = .27$) as when reading comprehension was the independent variable of interest. Once connective types were added to the final model,

an additional 6% of variance was explained ($R^2 = .33$), and adversative connectives had a significant positive association with argument sophistication in the essays.

In Study 3 (Cognitive and linguistic features of adolescent argumentative writing: Do connectives signal more complex reasoning?), I examined argument types and connectives in a sample of 60 6th – 8th grade students' essays (*N*=158) using descriptive statistics and a regression approach. The students were from urban schools in the West Coast of the US participating in the efficacy trial of Word Generation. Results indicate that essays with binary topics contained stronger arguments than the essays with open-ended topics. In addition, adversative connectives had a significant positive association with argument sophistication.

Themes

Each study in this dissertation answered specific research questions regarding children's and early adolescents' academic language use in their argumentative writing. There are several overarching themes, in light of the guiding theories, that emerged across the studies.

Developmental patterns in the argumentative writing of 2nd – 8th grade students

The results from these three studies confirm that an educational linguistics-informed approach to analysis of text can yield some insightful findings regarding children's and early adolescents' patterns of academic language production in their argumentative writing. First, it must be noted that although these data are not longitudinal (a limitation to be addressed again in the conclusion), the three studies investigating students' argumentative essays spanning 2nd through 8th grades utilized the same outcome, highest level of argumentative reasoning used, and employed similar methods to answer the research questions. The text-oriented analyses uncovered uses of specific argument types at the level of the T-unit (independent clause plus any

dependent clauses, per Hunt, 1965), ranging from a non-argument, to a simplistic own-side only argument, to a more complex, dual-perspective argument that considers and attempts to weaken an alternate perspective, to the most sophisticated, integrative-perspective argument, signifying a stance that maintains one's original position yet can integrate a competing idea, which is that there may be some downsides to one's own positions, or there may be some potential benefits to the opposing viewpoint. This argumentative coding scheme originated in Kuhn and Crowell's (2011) study of a three-year, 6th grade classroom intervention on argumentation. Applying it to these three studies, this coding scheme was found to be a reliable means of assessing specific argument types in one persuasive essay topic in 2nd and 3rd grade (Chapter 2); three topics in 4th grade and three topics in 5th grade (Chapter 3); and four topics in 6th through 8th grades (Chapter 4). Because argumentative reasoning and argument text structure itself has been included in conceptions of academic language (Snow & Uccelli, 2009; Uccelli et al., 2015a, b), this measurement tool is, from the linguistic perspective, a means of capturing this facet of academic language use in student writing. The other linguistic features of texts investigated were overall length (measured in total T-units) and uses of connectives, either researcher-coded (Chapters 2 and 3), or computed by the natural language processing tool, TAACO (Chapter 4) (Crossley, Kyle, & McNamara, 2016). In each study, the connective types were based on Halliday and Hasan's (1976) seminal categories: additive, adversative, causal, and temporal. Lexical precision for the purpose of signaling ideas, such as through using connectives, is another facet of academic language (Snow & Uccelli, 2009; Uccelli et al., 2015a, b) that was investigated in detail in the present studies.

The findings indicate that overall, the students' argumentative writing systematically gains in argument sophistication, although with some surprises. For the 2nd and 3rd grade sample

(1 essay per student), the mean and standard deviation of the highest level of argumentative reasoning used across the essays was 1.47(.84), which equates to a solid, own side only argument to about half-way to the next level of a dual perspective argument. The mean and standard deviation of the highest argument used across the 4th and 5th grade sample of essays (3 essays per student) was 1.89(.84), so this mean score reflects a score closer to a dual perspective argument. Finally, the mean of the 6th through 8th grade sample of essays (4 essays per student) was 2.09(.87), which is just above a dual perspective argument score. Thus, across the three samples of essays, there is a gradual increase for the mean argument sophistication measure. It is also important to notice the progression from each grade individually, although it is not completely linear. The mean for 2nd grade was 1.26(.81), and for 3rd grade it was 1.68(.83), with 3rd grade scoring significantly higher on the highest argumentative reasoning used in the essays. This means that 2nd grade scored closer to an own side only argument, while 3rd grade scored closer to a dual perspective argument. In 4th grade, the mean score was 1.73(.81), while the 5th grade mean score was 2.06(.84). This means that 4th grade was slightly closer to a dual perspective argument score than 3rd grade, while the 5th grade mean score moved to above a dual perspective score. For the 6th through 8th grade sample, an interesting shift occurred. Although the overall means for each of the three studies increased, when we disaggregate the scores by grade, 6th grade had a lower mean score than 5th grade, and 7th grade score was equal to the 5th grade score. That is, 6th grade mean score was 1.88(.85), so approaching a dual perspective argument, and 7th grade was 2.06(.87). Finally, the 8th grade mean score was 2.19(.88), which is slightly above a dual perspective argument. There were no significant differences between the grades in the 6th through 8th grade sample.

Why did the 6$^\text{th}$ grade essays score lower, on average, than the 5$^\text{th}$ grade essays? And, why were the 7$^\text{th}$ grade essays equal, on average, to the 5$^\text{th}$ grade essays? One possible explanation is that, although both of the samples were writing in response to Word Generation writing prompts, the 5$^\text{th}$ grade students may have had an additional year of involvement in the Word Generation program, if they had attended the same school in the prior year, so these students may have had an additional year of practice with this type of writing and thinking exercise. While there were only six 6$^\text{th}$ grade students, their CST (California Standards Test) English-language arts scores demonstrated proficient or advanced scores, so lower skills is not a likely explanation. However, an alternate explanation to the 5$^\text{th}$ grade students having an additional year of Word Generation engagement, is that in Study 3 (Chapter 4), a topic effect was detected. That is, two of the writing prompts were coded as binary topics, and two of the prompts were coded as open-ended topics, with the binary topics resulting in a significantly higher amount of complex reasoning. Thus, only two of the four topics that the 6$^\text{th}$ through 8$^\text{th}$ grade sample wrote about were the binary topics, which tended to elicit stronger argumentation. Note that the 4$^\text{th}$ and 5$^\text{th}$ grade topics were already similar types of topics (i.e., binary), so they were not coded as different *types* of topics. Again, though, these two studies are completely different samples, so the comparisons between them must realize that limitation.

Regarding essay length, it was interesting that the essays across the three studies did not show a progression in length. For example, the mean essay length in the 2$^\text{nd}$ and 3$^\text{rd}$ grade sample was 6.35 T-units, but the mean length decreased in the 4$^\text{th}$ and 5$^\text{th}$ grade sample to 5.0 T-units, and then the mean increased again in the 6$^\text{th}$ through 8$^\text{th}$ grade sample to 9.77 T-units. These findings may be a function of the amount of time students were given to write. The 2$^\text{nd}$ and 3$^\text{rd}$ grade students had 30 minutes for writing, while the 4$^\text{th}$ through 8$^\text{th}$ grade students likely only

had about 15-20 minutes to write each essay. This may explain the decrease in essay length in 4th and 5th grades. It is also important to note, however, that the mean T-unit length of essays may be masking increases in essay length if measured a different way, such as total word length or clauses per T-unit. Our measure of T-units was based on Hunt's (1965) conception, which includes dependent clauses. We did not measure clauses per T-unit, which, upon such an investigation, may reveal fewer overall T-units for 4th and 5th grades, but essays comprised of longer and more complex T-units. Further study of these students' syntactic complexity in their argumentative writing is warranted.

Instructional practices that make visible argumentative reasoning for writing

Overall, the mean scores for the highest level of argumentative reasoning used in the younger children's (i.e., 2nd and 3rd grades) essays tended toward the own side only argument range, and the 4th through 8th grade essays tended toward the dual perspective range. The results for the younger children concur with findings reflecting the myside bias (Ferretti & Graham, 2019), and the upper elementary and middle school results concur with studies that show that middle school students are capable of acknowledging and addressing the opposing position in an argument with encouragement or prompting (Kuhn & Crowell, 2011; Kuhn & Udell, 2007). These findings indicate that the design of Word Generation as a dialogic curriculum – that is, discussion and debate exposing students to multiple perspectives during the week leading up to the culminating, Friday writing exercise to craft their own perspective – may be a promising avenue for students to practice exploring multiple perspectives and writing in such a way that acknowledges those alternate viewpoints. This is also confirmation of studies that dialogic instruction can be a path toward the development of students' individual reasoning (Kuhn & Crowell, 2011; Reznitskaya & Anderson, 2002; Reznitskaya et al., 2001).

There is another noteworthy finding across the studies relevant to this theme of instructional practices that make visible argumentative reasoning for writing. The fact that the most complex type of reasoning, integrative perspective, showed up at all is noteworthy. In Kuhn and Crowell's (2011) three-year intervention for teaching argumentation in 6th grade, this type of thinking only appeared in the experimental condition in the third year of the intervention. The percentages of essays employing integrative perspective in the present studies showed the following progression: 13% in the 2nd and 3rd grade sample; 27% in the 4th and 5th grade sample; and 42% in the 6th through 8th grade sample. Note that 13% in 2nd and 3rd grades would be at the student level, while 27% at 4th and 5th grades and 42% at 5th and 6th grades would be at the essay level, because students wrote three essays each or four essays each, respectively. In addition, for the control group, the percentage of participants making dual perspective arguments (also considered complex reasoning) remained steady at about 30% across all three years. In the present studies, the 2nd and 3rd grade sample demonstrated 36% of essays containing at least one dual perspective argument; 4th and 5th grade essays contained 53%; and 6th through 8th grade essays contained 50%. These findings also reveal higher percentages than for the 6th grade control group not receiving the argumentation intervention in Kuhn and Crowell (2011).

Again, there appears to be a developmental progression in the present studies with these more complex types of reasoning, and why would these types of reasoning seem to naturally appear in the absence of an argumentation intervention? That is, the 2nd and 3rd grade essays were a baseline writing measure (i.e., on-demand writing at the beginning of the year without instruction), and the 4th through 8th grade essays were also independently written in a brief, 15-20 minute writing time. A possible explanation for the younger children is that the prompt (Should children be able to choose their own pets?) was carefully selected so as to be a familiar topic that

would likely not have needed additional background knowledge. Thus, the prompt may have facilitated the children's immediate engagement with reasoning about the topic. The topic also lent itself to a binary outcome, in that children could assume the alternative was that parents should be able to choose, and either way, they were able to consider and write about the alternative perspective. A possible explanation for the $4^{th} – 8^{th}$ grade essays again falls under the interpretation that participating in regular, dialogic instruction about the topics prior to writing may have fostered a culture of reasoning in the classroom that may have transferred to the individual students' reasoning expressed in their essays.

Reading and writing activities for practicing academic language

The conceptual framing and methods of this study are consistent with Myhill's (2009) call for linguistics-oriented writing research to be cognizant of multiple theoretical perspectives, and to "align knowledge of linguistic development with cognitive and socio-cultural insights in into writing processes" (p. 17). The data in these three studies seem to fit with Kim and Schatschneider's (2017) DIEW (direct and indirect effects of component skills on writing), which expands the Simple View of Writing (Berninger, Abbott, Abbott, Graham, & Richards, 2002; Juel, Griffith, & Gough, 1986) by operationalizing the component skills of oral language as the factors of the text generation process. The parallels of their model to these studies, which integrate linguistic, sociocultural, and cognitive perspectives, are clear. First, the foundational oral language skills of vocabulary and grammatical knowledge can be seen as mapping onto students' uses of the linguistic feature of connectives that these studies investigated. Knowledge of connectives can be considered as a special subset of academic vocabulary (Crosson & Lesaux, 2013a, b), and then marshaling them for specific functions to signal ideas during text generation, are skills that seem to encompass both of the foundational oral language skills of vocabulary and

grammatical knowledge. All three studies found a significant relation between connectives and more sophisticated reasoning in the students' essays. Specifically, adversative connectives predicted more complex reasoning. This extends the findings of the DIEW model, where these foundational skills contribute to the higher-order cognitive skills, in that the model was fit for first grade students, and these data are from a range of students beyond first grade. Thus, there are now also some indications for the conceptual model of language for argumentative reasoning presented in Chapter 1. Argumentative writing, as the macro-genre or text-type, when implemented in an instructionally relevant way (i.e., carefully selected topics and quality discussion and debate occurring prior to writing), can provide a training ground of opportunities for students to practice the word-level, foundational oral language skills and, at the micro-genre/sentence level, the higher-order cognitive skills of reasoning and perspective taking. This also aligns with the sociocultural and cognitive theory of argument schema theory (Reznitskaya & Anderson, 2002; Reznitskaya et al., 2001), as mentioned above.

Results from Studies 1 and 2 underscore the reading-writing relationship (Fitzgerald & Shanahan, 2000; Graham et al., 2018). Both studies found a relation between measures of reading comprehension and the writing outcome measure of highest level of argumentative reasoning used. This is consistent with the research literature that has established that reading interventions can influence writing outcomes (Graham et al., 2018). Researchers also contend that writing predicts reading comprehension (e.g., Graham & Hebert, 2011a, b), as well as the possibility that it may be a bi-directional relationship (Graham et al., 2018). Word Generation (Study 2) is a cross-content area academic vocabulary program aimed at improving reading comprehension (Snow, Lawrence, & White, 2009; Lawrence, Crosson, Paré-Blagoev, & Snow, 2015). Programs such as this that facilitate not only word learning, but also academic language,

complex reasoning, and perspective taking, may be a promising path for post-primary language and literacy learning, in order to continue instruction in support of reading comprehension beyond 3rd grade. Study 1 did not involve an academic vocabulary or reading intervention, but the finding of reading predicting sophistication of argumentation in the 2nd and 3rd grade students' essays, extends our knowledge and confirms the hypothesis that reading comprehension would be associated with this micro-genre/sentence-level measure of reasoning in writing. Thus, further research in productive use of academic language in student writing, as well as in reading instruction that makes academic language visible for not only early elementary, but also late elementary and middle school, is warranted.

Implications

The studies in this dissertation are correlational, and as such, they do not compare teaching practices. However, these results are consistent with findings from a body of research that I have reviewed and I present here as a tentative set of suggestions for practice, along with some recommendations for research.

Table 5.2 presents implications for practice and research. For pedagogical implications, this dissertation can provide suggestions for argumentative writing instruction and fostering students' awareness and use of academic language in their writing. Across the three studies, I found evidence supporting the importance of (1) the suitability of argumentative writing activities, and (2) making visible the aspects of academic language in writing. First, regarding the suitability of argumentative writing activities, educators are encouraged to consider the nature and frequency of types of writing exercises for their students (Chapters 3 & 4). In grades 4-8, teachers can consider that extended pieces of writing utilizing every stage of the writing

process may not need to be the only aim in writing instruction; instead or in addition, they can consider brief but frequent opportunities for students to write about their perspectives on contestable topics. The purpose of these 15-20 minute independent writing exercises would be more of a thinking opportunity and a chance to consolidate their opinion on a topic the class has previously discussed or debated; that is, students have already been made aware of and interacted with multiple perspectives on the topic. In addition, regarding selection of argumentative writing topics, one consideration is that for on-demand, diagnostic tools to perhaps assess students' writing skills and perspective taking at the beginning of the year, educators are encouraged to choose familiar topics so that the children may have similar levels of prior knowledge on the topic (Chapter 2). For upper elementary and middle school, teachers can select topics with distinct perspectives (as opposed to open-ended topics) for classroom discussion and writing.

Second, educators are encouraged to make visible to their students the aspects of academic language needed for argumentative writing. For example, even as early as 2nd and 3rd grade (Chapter 2) but also throughout elementary and middle school (Chapters 3 & 4), teachers can introduce ways of reasoning about contestable topics, with the aim of showing different functions of types of arguments for a given topic (e.g., whether to support one's own stance or to weaken the other stance). Making types of reasoning visible can also serve the purpose of demonstrating how each individual argument is a writing subgoal, leading to the overall synthesis of the student's stance on a topic. At the same time, educators are encouraged to instruct students in the area of general academic vocabulary, while presenting connectives as a special subset of that vocabulary. Again, by showing students the functions of these word classes, they can encourage the use of more sophisticated academic language to their students, so they can deploy examples of precise academic language as they write argumentative essays.

Table 5.2

Implications for Practice and Research

	Implications
For Practice	1. *Suitability of argumentative writing activities* • Implement regular, brief writing exercises on contestable topics as an individual follow up to classroom discussions and debate (Be mindful that extended pieces of writing need not be the only aim in writing instruction) (Chapters 3 & 4) • Select familiar topics for diagnostic writing assessment (Chapter 2), and topics with clear potential stances, as opposed to open-ended topics (Chapters 3 & 4)
	2. *Making visible the aspects of academic language in writing* • Implement instructional activities to help students to become aware of writing for thinking subgoals, such as types and functions of reasoning (e.g., advancing your own argument or weakening the opposing argument) (Chapters 2, 3, & 4) • Integrate instruction on connectives within instruction on general academic vocabulary (Chapters 2, 3, & 4)
For Research	1. *Use of analytic scoring of features in student writing* • Essay analysis at the T-unit level to capture variability on arguments and linguistic features (Chapters 2, 3, & 4) • Natural language processing tools to compute linguistic features in essays (Chapter 4)
	2. *Use of multi-occasion assessments* • Collect multiple samples of writing for each student for more reliable and robust data (Chapters 3 & 4) • Test whether production of cognitive or linguistic features in essays differ as a function of topic or student (Chapters 3 & 4)

For research implications, I would like to highlight the following methodological suggestions: (1) use of analytic scoring of features in student writing, and (2) use of multi-occasion assessments. First, writing researchers are encouraged to use analytic scoring of features at a fine-grained level in students' writing, such as the T-unit level or other definitions of an idea unit (Chapters 2, 3, & 4). For example, for essay level or other discourse-level analyses, instead of a single score encompassing multiple aspects, segmenting the overall piece of writing into idea units can afford a more detailed analysis of individual reasoning moves within the piece of writing. This can lend insights into how students are using, in essence, sentence-level academic language features and writing subgoals within a larger context of writing, which can also capture the variability of language use within micro-genres of academic writing. By gaining insight into these idea-by-idea uses of academic language within students,

this could help serve as groundwork for developing research and instructional tools for individualizing writing instruction. In addition, writing researchers are advised to consider using natural language processing tools such as the TAACO (Chapter 4) for automatic calculation of a variety of linguistic features in students' writing. By doing so, this is another useful way to discover the variability in elementary and middle school students' use of linguistic features in their writing, which could also serve as groundwork for developing research and instructional tools for individualizing instruction.

Second, I encourage researchers to incorporate multiple occasions of writing assessments in their study designs. Collecting multiple samples of writing for each student affords a more robust and reliable means of data analysis (Chapters 3 & 4). Not only can it help to rule out a topic effect if there is only one topic, it can also potentially afford longitudinal analysis, to investigate students' growth over time. In addition, having multiple samples of writing within students affords comparison of the writing features by topic, to investigate the nature of topic effects. Within student analyses can also illuminate which features of writing vary more by student or by language feature, which can help to inform further research and instruction.

Conclusion

Taking these three studies together, there are some limitations to address. First, these data are not longitudinal, but rather three separate correlational studies. Second, the writing outcome measure is not an overall quality measure, but rather a T-unit level measure of reasoning within the writing, and so we should exercise caution when interpreting these results as linear. That is, while we should not assume that the range of argument levels – from non argument, to own side only, to dual perspective, to integrative perspective – demonstrates an exact linear progression,

we can consider that the argument coding scheme (Kuhn & Crowell, 2011) for these argument types incorporates epistemological theories which support that each type is increasingly more challenging and qualitatively different from the other (Kuhn, 1992). Thus, while we should exercise caution, using this outcome measure is a good starting point for building on this line of research that can later explore more complex measures of argumentative writing that incorporate reasoning, such as presented by Ferretti, Lewis, and Andrews-Weckerly (2009) or Reznitskaya and colleagues (2009).

Yet even considering the limitations, these studies contribute to the fields of educational linguistics and writing research and provide some directions for future research. Elementary and middle school students are faced with more demanding texts to comprehend and produce in writing, such as with argumentative writing. These studies shed light on the nature of the reasoning, as well as the linguistic features of connectives, that students independently use in their writing, and how they are related. These findings shed light on later language development from grades 2 through 8, whereas most language development studies have focused on early childhood (Myhill, 2009). The findings can also inform the development of interventions for teaching and research of academic language, reasoning, argumentative writing, and reading comprehension. For future work, further research on assessment of argumentative writing, teaching and learning of vocabulary, longitudinal studies, as well as qualitative and design-based studies to examine the take up of programs such as Word Generation, will enrich and extend our knowledge even further.

References

Berninger, V. W., Abbott, R. D., Abbott, S., Graham, S., & Richards, T. (2002). Writing and

reading: Connections between language by hand and language by eye. *Journal of*

Learning Disabilities, 35(1), 39-56.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text

cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior*

Research Methods. doi:10.3758/s13428-015-0651-7

Crosson, A. C., & Lesaux, N. K. (2013a). Connectives: Fitting another piece of the vocabulary

puzzle. *The Reading Teacher, 67*(3), 193-200. doi: 10.1002/TRTR.1197

Crosson, A. C., & Lesaux, N. K. (2013b). Does knowledge of connectives play a unique role in

the reading comprehension of English learners and English-only students? *Journal of*

Research in Reading, 36(3), 241-260. doi: 10.1111/j.1467-9817.2011.01501.x

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and

instruction. *Reading and Writing, 32, 1345-1357.*

Ferretti, R. P., Lewis, W. E., & Andrews-Weckerly, S. (2009). Do goals affect the structure of

students' argumentative writing strategies? *Journal of Educational Psychology, 101*(3),

577-589.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development.

Educational Psychologist, 35(1), 39-50.

Graham, S., & Hebert, M. (2011a). *Writing to read:* New York: Carnegie Corporation of New

York.

Graham, S., & Hebert, M. (2011b). Writing to read: A meta-analysis of the impact of writing and

writing instruction on reading. *Harvard Educational Review, 81*(4), 710-144.

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., & Aitken, A. (2018). Reading for writing:

> A meta-analysis of the impact of reading interventions on writing. *Review of Educational*

> *Research, 88*(2), 243-284.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Longman Group Ltd.

Hunt, K. W. (1965). A synopsis of clause-to-sentence length factors. *The English Journal, 54*(4),

> 300+305-309.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of

> children in first and second grade. *Journal of Educational Psychology, 78,* 243-255.

Kim, Y.-S., & Schatschneider, C. (2017). Expanding the developmental models of writing: A

> direct and indirect effects model of developmental writing (DIEW). *Journal of*

> *Educational Psychology, 109*(1), 35-50.

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

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young

> adolescents' thinking. *Psychological Science, 22*(4), 545-552.

> doi:10.1177/0956797611402512

Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking*

> *and Reasoning, 13*(2), 90-104.

Lawrence, J. F., Crosson, A. C., Paré-Blagoev, E. J., & Snow, C. E. (2015). Word Generation

> randomized trial: Discussion mediates the impact of program treatment on academic

> word learning. *American Educational Research Journal, 52*(4), 750-786.

> https://doi.org/10.3102/0002831215579485

Myhill, D. (2009). From talking to writing: Linguistic development in writing. *British Journal of Educational Psychology Monograph Series 2, No. 6 – Teaching and Learning Writing* (pp. 1-33). British Psychological Society.

Reznitskaya, A., Kuo, L., Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual Differences, 19,* 219-224.

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In C. Block & M. Pressley (Eds.), *Comprehension instruction* (pp. 319-334). New York: Guilford.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2-3), 155-175. doi: 10.1207/S15326950DP3202&3_04

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2*(4), 325-344.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In N. Torrance & D. R. Olson (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge University Press.

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015a). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics,* 1-33. doi:10.1017/S014271641400006X

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015b). Beyond

vocabulary: Exploring cross-disciplinary academic-language proficiency and its

association with reading comprehension. *Reading Research Quarterly, 50*(3), 337-356.

doi:10.1002/rrq.104

APPENDIX A

Study 1 Student Instructions Script and Writing Prompt

Student Instructions Script:

"Today you will be writing an *opinion* essay about *whether children should be able to choose their own pets (why or why not)*. Go into as much detail as you can. You should write as if you were writing an assignment for your teacher. Keep in mind everything that you know about good writing. You will have until *lunchtime*, about *30* minutes, to complete your essay. You don't have to take the whole time. Once you finish, raise your hand and I will pick up your writing, *take you back to your classroom, and give you something new to work on.* I will let you know when you have 5 minutes remaining.

You will be writing in pen. If you make a mistake, just cross it out and continue writing. If you finish early, please re-read your writing and make any necessary changes before you raise your hand to tell me you're done.

If you have any questions while writing about *whether children should be able to choose their own pets (why or why not)*, just raise your hand; you don't have to get up. Are there any questions before we start?"

*These directions will be altered where italicized to fit each prompt, specific classroom schedule, and school. The prompt will be read twice within the instructions, once at the beginning and again when asking for questions. The test proctors may also write the prompt on the board and will be allowed to repeat the prompt to a student if they specifically request it. Prior to these directions their teacher will tell the students what task(s) they will be completing after they finish the writing task. An effort will be made to ensure the tasks following the writing assessment are not significantly easier or more appealing to the students, reducing their desire to rush through the writing task.

Prompt:
*Think about whether or not children should be allowed to choose their own pet. Decide what you believe and write an essay that will convince someone to agree with you.

Note. (*) Identifies prompt adapted from Graham, Harris, & Mason (2005).

Appendix B

Coding scheme for 2nd and 3rd grade essays with examples from the data

Chosen position: YES, children should be able to choose their own pets.

Opposing position: NO, children should not be able to choose their own pets.
(Parents should choose the pets.)

Argument Type	Code
Non-argument	0
Position statement only	
Repeat of an earlier argument	
Vague or unclear	
Add-on with no reason (That's all I have to say.)	
Own Side Only (OSO) argument	1
They would take care of it because they really want that pet they choose.	
The child should make their own choices because it's their life.	
Dual Perspective (DP) argument	2
It is not fair if people's parents get to choose the pet.	
The parents might not be sure the kids want that pet.	
The kids will probably not take care of it because they didn't want it in the first place.	
Integrative Perspective (IP) argument	3
You could let someone else pick your pet (negative of own side). But I say that you should pick your own pet so that you can love and care for your pet.	

Coding scheme for 2nd and 3rd grade essays with examples from the data

Chosen position: NO, children should not be able to choose their own pets.
(Parents should choose the pets.)

Opposing position: YES, children should be able to choose their own pets.

Argument Type	Code
Non-argument	0
Position statement only	
Repeat of an earlier argument	
Vague or unclear	
Add-on with no reason (That's all I have to say.)	
Own Side Only (OSO) argument	1
Parents need to spend their money on gas, food, and other stuff for them and you.	
Moms and dads have the most money. That's why they can buy the pets.	
I should listen to my parents.	
Dual Perspective (DP) argument	2
It might cost a lot.	
Your parents might say no.	
I might pick a bad pet that can do something to me.	
Integrative Perspective (IP) argument	3
Maybe if they do have the money (positive of the other side), then you can get the pet that you want.	

Holistic Score of Connectives Sophistication

Overall considering (Not weighting one more than another):
• Presence
• Logic or Purpose
• Clarity
• Variety

1: Weak	2: Developing	3: Strong
No use of connectives Only 1 category of connectives is present Connectives are present, but uses of connectives are: • Incorrect and/or • Illogical and/or • Immature	Only 1 category of connectives is present, but there is variety within the category 2 categories of connectives present Connectives somewhat add to the logic of the arguments or the whole essay Overall use of connectives is somewhat mature	Only 2 categories of connectives present, but there is variety within each category 3+ categories of connectives present Connectives strongly add to the logic of the arguments or the whole essay Overall use of connectives is sophisticated

Appendix D

Coding scheme for 4th and 5th grade essays with examples from the data

Chosen position: YES, students should wear school uniforms.

Opposing position: NO, students should not wear uniforms.
(They should wear their own clothes.)

Argument Type	Code
Non-argument	0
Position statement only	
Repeat of an earlier argument	
Vague or unclear	
Add-on with no reason (That's all I have to say.)	
Own Side Only (OSO) argument	1
We'll be more focused on school.	
Can avoid being teased or bullied based on clothes.	
Shows team membership.	
Dual Perspective (DP) argument	2
Students might compete about clothes.	
Students would focus more on fashion than school.	
Buying your own clothes would cost too much money.	
Integrative Perspective (IP) argument	3
We always say no before we at least see what it is (negative of own side).	
We should at least see what the uniforms look like before we judge them.	

Coding scheme for 4th and 5th grade essays with examples from the data

Chosen position: NO, students should not wear uniforms.
(They should wear their own clothes.)

Opposing position: YES, students should wear school uniforms.

Argument Type	Code
Non-argument	0
Position statement only	
Repeat of an earlier argument	
Vague or unclear	
Add-on with no reason (That's all I have to say.)	
Own Side Only (OSO) argument	1
It is a right to wear what you want; gives kids freedom.	
Students should be able to express themselves.	
You'll have backup clothes if your clothes get dirty.	
Dual Perspective (DP) argument	2
Can't express yourself or wear your favorite color.	
Uniforms might cost too much.	
The material is itchy or uncomfortable.	
Integrative Perspective (IP) argument	3
If we have to wear them (positive of the other side), they should give them away for free.	

Appendix E

Coding scheme adapted from Kuhn & Crowell (2011) and verbatim sample statements from the data corpus

Type of argument	Example from data *Drugs should be legal* is the favored position	*Drugs should not be legal* is the favored position
No argument States position only with no support; unclear; repeat	I think that drugs should be legalized.	Drugs shouldn't be legalized.
Own side only Offers only positives of the favored position	If drugs were legalized, then the police will have more time paying attention to more serious crimes.	Drugs should stay illegal to prevent less people/teens from doing it.
Dual Perspective Offers negatives of the opposing position	I believe too much people are getting arrested for drugs.	First of all, legalizing them would be incompatible with saving lives.
Integrative Perspective Includes negatives of the favored position or positives of the opposing position; usually expressed in 2 T-units	Other people might say that more people will die because of drugs [negative of favored position]. But, if the government puts a tax on drugs, the economy might get better [positive of favored position].	One might say it helps lower criminal violations [positive of the opposing position], but then there would be more deaths [negative of opposing position].