



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

Karen S. Taylor¹ · Joshua F. Lawrence² · Carol M. Connor¹ · Catherine E. Snow³

© Springer Nature B.V. 2018

Abstract

The Common Core State Standards (Common Core State Standards Initiative, 2010) feature argumentative writing across the curriculum in grades 4 through 12, yet little is known about how young adolescents develop the challenging advanced language and literacy skills needed for these tasks. This study explored productive academic language use in the persuasive writing of a sample of 40 middle school students (grades 6–8) by examining the use of (1) argumentative moves that display various levels of sophistication and (2) major classes of connectives (additive, adversative, causal, and temporal) that signal different cohesive functions within a text. Essays in our analytical sample ($n=158$) were produced in the context of an academic vocabulary curriculum, Word Generation, and were transcribed, coded, and analyzed for types of arguments by researchers and undergraduate research assistants. Subsequently, connectives were calculated by the Tool for the Automated Analysis of Cohesion (TAACO; Crossley, Kyle, & McNamara, 2016). Descriptive analyses reveal that the sixth–eighth grade students in our sample deployed complex reasoning in their essays; at least one dual perspective argument was present in 50% of the essays, and at least one integrative perspective argument was present in 42% of the essays. Multivariate regression analyses (with adjusted standard errors) reveal that adversative connectives (e.g., *although*, *however*) were related to the most complex arguments (integrative perspective), controlling for essay length and topic type ($\beta=20.13$, $p=.006$), as well as to overall argument sophistication ($\beta=17.25$, $p=.02$). The results show the value of brief, curriculum-based essays for assessing students' argumentation skills.

Keywords Writing · Argumentative writing · Adolescent literacy · Academic language · Reasoning · Connectives

✉ Karen S. Taylor
taylorks@uci.edu

Extended author information available on the last page of the article

Introduction

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 et al., 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, Barr, et al., 2015; Uccelli, Galloway, et al., 2015).

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, Barr, et al., 2015; Uccelli, Galloway, et al., 2015; 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 et al., 2001; Reznitskaya & Anderson, 2002) 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 Table 9 in the "Appendix". 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, 2013b; Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015). 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, adversative, causal, and temporal relationships (Halliday & Hasan, 1976). The category of additive connectives includes words or phrases such as: *furthermore*, or *in addition*. Adversative connectives include examples such as: *however*, *nevertheless*, or *despite this*. Causal connectives could include: *therefore*, *consequently*, or *as a result*. Finally, temporal connectives could be expressed: *then*, *after that*, *finally* (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 enables assessment of these connectives. Known as the Tool for the Automated Analysis of Cohesion (TAACO; Crossley et al., 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 complex 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 this range of grades responded to the same essay prompts in their workbooks.

Participants

Table 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

Table 1 Student demographic characteristics

| Demographic | 6th grade <i>n</i> (% of grade) | 7th grade <i>n</i> (% of grade) | 8th grade <i>n</i> (% of grade) | All students <i>n</i> (% 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 (16.7) | 1 (6.3) | 2 (11.1) | 3 (7.5) |
| Vietnamese | | 1 (6.3) | 1 (5.6) | 2 (5) |
| Black/African American | | 1 (6.3) | 3 (16.7) | 1 (2.5) |
| Japanese | | | | 1 (2.5) |
| Missing | | | | 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) | 8 (50) | 3 (16.7) | 4 (10) |
| Redesignated Eng. fluent (RFEP) | 2 (33.3) | 3 (18.8) | 6 (33.3) | 16 (40) |
| Limited English fluent (LEP) | | | 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) | 1 (5.6) | 2 (5.1) |
| Far below basic | | 1 (6.7) | | 2 (5.1) |

^aCST data were not available for 1 student

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.

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) identified students as Initially Fluent English Proficient (IFEP) or Limited English Proficient (LEP). LEP students whose yearly evaluation process yielded an overall proficiency level adequate for reclassification were identified as Reclassified Fluent English 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 scoring 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 2 for the list of units and topics). All students completed the four essays except two were missing the Unit 20 essay, 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 4 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 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 the "Appendix" 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 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 non-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.

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

Table 3 Descriptive statistics for length, arguments, and connectives per essay (n = 158)

| Essay features | M (SD) | 6th grade M (SD) | 7th grade M (SD) | 8th grade M (SD) |
|----------------------------------|-------------|---------------------|---------------------|---------------------|
| Essay 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 | | | | |
| Non-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 | | | | |
| Additive connectives | .05 (.02) | .04 (.02) | .05 (.03) | .05 (.02) |
| 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) |

^aArgument features are frequency counts per essay

^bConnectives are proportions of connectives over total words

Table 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) |

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 a dual perspective argument score of “2” (1.88) and seventh and eighth grade essays scoring above the dual perspective argument score (2.06 and 2.19, respectively).

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 ($t = -.274$; $p = .007$). 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 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.

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 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

Table 5 Correlations among essay features measured

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------------------------|---------|----------|-------|----------|-------|---------|---------|-------|-------|----|
| 1. Total T-units | 1 | | | | | | | | | |
| 2. Binary topics | .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 |

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

Table 6 Intraclass correlations

| Essay features | Rho coefficient | Standard error |
|-----------------------------------|-----------------|----------------|
| Global features | | |
| Essay length in T-units | .67 | .06 |
| Argument features | | |
| Non-arguments | .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 |

nesting of essays within students using the Huber–White adjustment (robust standard errors) in subsequent regressions rather than a full multilevel model.

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 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 ($\beta = -18.369$; $p = .03$), 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 ($\beta = -23.679$; $p = .05$), 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 ($\beta = 20.13$; $p = .006$), with a significant overall model that explained 11% of the variance in the essays. Causal and temporal connectives did not predict argument types.

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

Table 7 Regression models by argument type: contribution of connectives to arguments, controlling for essay length and binary topics

| Argument type | Variable | β | Robust standard error | p | R^2 |
|---------------|---------------|---------|-----------------------|---------|--------|
| Non-arguments | Total T-units | .686 | .034 | < .0001 | .81*** |
| | 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 | |
| | | | | | |
| OSO arguments | Total T-units | .199 | .04 | < .0001 | .32*** |
| | 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 | |
| | | | | | |
| DP arguments | Total T-units | .097 | .02 | < .0001 | .43*** |
| | 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 | |
| | | | | | |
| IP arguments | Total T-units | .019 | .011 | .08 | .11* |
| | 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 | |
| | | | | | |

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

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 8, linear regression

Table 8 Predictors of overall argument sophistication

| Variable | Overall argument sophistication | | | R^2 |
|---------------|---------------------------------|-----------------------|---------|--------|
| | β | Robust standard error | p | |
| Essay length | .015 | .013 | .239 | .24*** |
| 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*** | |

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

results indicate that these variables explain 24% of the variability in students' argument sophistication in the essays, accounting for the nesting of essays within students. 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 ($\beta = 17.251$; $p = .021$).

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 3 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 min 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 et al., 2001; Reznitskaya & Anderson, 2002) 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 .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 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 intra-class 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 ($\rho = 0-.25$), 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 10 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.

Acknowledgements The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through Grant R305A090555 (Catherine Snow, Principal Investigator) to the Strategic Education Research Partnership (SERP), and through Grant R305F100026, awarded to SERP as part of the Reading for Understanding Research Initiative. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education. We thank the collaborating school districts and school personnel, as well as the participating teachers and students. We would also like to thank Breanna Briggs, Andrea Byng, Ragiah El-Shantaly, and Samaan Nur for their help and insight with applying the argument coding system to the essays. Special thanks are also due to Robert Selman and Nonie Lesaux for their support during the development of this study.

Appendix

See Table 9.

Table 9 Coding scheme adapted from Kuhn & Crowell (2011) and verbatim sample statements from the data corpus

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

References

- Al-Adeimi, S. (2018). *Talking to learn: Investigating the relationship between classroom discussion and persuasive writing* (Unpublished doctoral dissertation). Cambridge, MA: Harvard Graduate School of Education.
- 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, 291–310. <https://doi.org/10.1080/03057640903103751>.
- Andriessen, J. (2006). Arguing to learn. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (1st ed., pp. 443–460). New York, NY: 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, J. 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. <https://doi.org/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, 193–200. <https://doi.org/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, 241–260. <https://doi.org/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, 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. In B. Fishman, W. Penuel, A.-R. Allen, B. Cheng, & N. Sabelli (Eds.), *Design-based implementation research: An emerging model for transforming the relationship of research and practice*, National society for the study of education yearbook (Vol. 112, pp. 400–425).
- Duhaylonsod, 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, 587–608.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hu, C., & Li, Y. (2015). Discourse connectives in L1 and L2 argumentative writing. *Higher Education Studies*, 5(4), 30–41. <https://doi.org/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.
- Hunt, K. W. (1965). A synopsis of clause-to-sentence length factors. *The English Journal*, 54, 300+305–309.
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, 22, 545–552. <https://doi.org/10.1177/0956797611402512>.
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, 74, 1245–1260.
- Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning*, 13, 90–104. <https://doi.org/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, 767–793. <https://doi.org/10.1080/19345747.2016.1237596>.
- Lawrence, J. F., 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, 151–161. <https://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*, 201–221. <https://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*, 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)*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.
- 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, D.C.: 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 ed., pp. 319–334). New York, NY: Guilford Press.
- Reznitskaya, A., & Anderson, R. C. (2006). Analyzing argumentation in rich, natural contexts. *Informal Logic, 26*, 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*, 155–175. <https://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. <https://doi.org/10.1016/j.lindif.2008.11.001>.
- 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). New York, NY: Cambridge University Press. <https://doi.org/10.1017/cbo9780511609664.008>.
- Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015). 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. <https://doi.org/10.1017/S014271641400006X>.
- Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly, 50*, 337–356. <https://doi.org/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 & Reasoning, 14*, 1–27. <https://doi.org/10.1080/13546780701527674>.

Affiliations

Karen S. Taylor¹ · Joshua F. Lawrence² · Carol M. Connor¹ · Catherine E. Snow³

¹ School of Education, University of California, Irvine, 3200 Education, Irvine, CA 92697, USA

² Department of Education, University of Oslo, Oslo, Norway

³ Harvard Graduate School of Education, Cambridge, MA, USA