Reading and Writing
https://doi.org/10.1007/s11145-021-10140-x

A multi-dimensional examination of adolescent writing: considering the writer, genre and task demands

Penelope Collins[1] · Tamara P. Tate[1] · Joong won Lee[1] · Jenell A. Krishnan[1] · Mark Warschauer[1]

Accepted: 19 February 2021

Abstract

We examined the contributions of English proficiency, genre, and the use of textual sources to adolescent writing. The sample included 1819 native English speakers and language minority students from 127 seventh- and eighth-grade classes in an urban school district. Students were randomly assigned one of three source-based essay prompts (narrative, explanatory or argumentative) as part of the annual state assessment, and all students wrote a common, on-demand argumentative essay that did not require the use of textual sources. Overall, language minority students who were fluent English proficient wrote higher quality nonsource-based essays that contained more difficult vocabulary, were better structured, and used examples and details more effectively than native English speakers. They also outperformed native English speakers in source-based writing. Students with limited English proficiency showed weaker performance on both source-based and nonsource-based writing tasks. Differences in English proficiency held across genres. Students who wrote argumentative essays obtained higher ratings than those who wrote narrative or explanatory essays. Source-based and nonsource-based writing were moderately correlated. Regression analyses revealed that in addition to English proficiency and genre, reading comprehension's contribution to source-based writing was almost double that of nonsource-based writing. Implications of the findings for theory and practice are discussed.

Keywords Writing · Adolescent · English learners · Language minority · Middle school · Genre

We thank the participating educators and students who made this research possible. This research was supported by a grant from the US Department of Education, Award R305A150429, PI: Warschauer, Co-PIs: Collins & Farkas, "Digital Scaffolding for English Language Arts." We have no known conflict of interest to disclose.

✉ Penelope Collins
p.collins@uci.edu

[1] School of Education, University of California, Irvine, CA 92697-5500, USA

Published online: 07 March 2021

Introduction

Writing is an essential skill for all students and one that is not being mastered. According to the National Center for Educational Statistics (NCES, 2012c), only 27% of 8th-grade students performed at or above the criterion proficiency level in writing on recent national tests in the United States, 34% in reading (NCES, 2012b), and 35% in mathematics (NCES, 2012a). The challenges are even greater for children with limited proficiency in English, as only 1% of the 8th grade English learners received writing scores at or above the level considered proficient (NCES, 2012d).

The development of writing skills is challenging because writing is a complex, multidimensional task (Bazerman et al., 2017). Three dimensions critical for understanding student writing are the writers themselves, the texts they are constructing, and the writing assignment's task or constraints. These three dimensions are interrelated and are situated within a broader sociocultural context (Bazerman et al., 2017; RAND, 2002).

Understanding the writer

Learning to write requires students to use and coordinate a range of higher-order cognitive capacities (e.g., executive function, reasoning skills, and memory), motivation (such as achievement goals and beliefs about writing), and linguistic knowledge (Flower & Hayes, 1981; Graham et al., 2015). These processes come into play in each step of the writing process (Flower & Hayes, 1981), which involves: *planning*, whereby writers generate and organize their ideas; *drafting* or translating ideas into paragraphs and sentences; *revising*, whereby writers evaluate their text and make changes to improve it; and *editing*, in which writers correct spelling, grammatical and mechanical errors. Thus, learning to write effectively involves a specialized set of higher-order linguistic and cognitive practices (Graham et al., 2015).

One source of variation among individuals stems from the rapid growth of linguistically diverse students or language minority speakers in American schools. There is tremendous heterogeneity among language minority students, not just in their home languages but also in their mastery of the academic register of English used in schools. There has been growing recognition of the different profiles and needs of language minority students with limited English proficiency (LEP) and those who are relatively fluent in English (FEP) and have attended English dominant schools for at least three years (Hwang et al., 2015; Perin et al., 2017; Ragan & Lesaux, 2006). The latter is also known as Generation 1.5 (Perin et al., 2017). Although much of the literature has focused on the writing development of LEP students, a small yet growing body of research has called attention to Generation 1.5, or FEP students, as their writing performance has been found to differ from both their LEP peers and native English speakers (NES; Perin et al., 2017).

Because LEP students' English proficiency is still emerging, it is unsurprising that their compositions tend to differ from those written by their native English-speaking peers. LEP students tend to be less productive in their writing (Collins

et al., 2013; Reynolds, 2005). However, as language minority students gain English proficiency, their compositions are comparable in length to those of native English speakers (Collins et al., 2013; Grant & Ginther, 2000).

Variations in language minority students' writing go beyond productivity. LEP students rely on both their native language and their emerging English knowledge as resources in their writing (Brisk, 2012). In the upper-elementary and secondary grades, LEP students use less formal, high-frequency words, and fewer academic words than native English speakers (Collins et al., 2013; Lee, 2003). Similarly, LEP students may incorporate their home languages' syntactic structures in their writing and use less syntactically complex English structures (Brisk, 2012; Collins et al., 2013). Because LEP students' composition may use less formal writing styles or reflect the discourse structures of their home languages, their writing tends to receive lower overall quality ratings in elementary through post-secondary school (Collins et al., 2013; Hinkel, 2003; Leki et al., 2008). Further, LEP students' essays been less successful at addressing the prompt and contained fewer examples and details than NES students (Collins et al., 2013). In contrast, as language minority students gain proficiency in English, their compositions' syntactic complexity has been found to increase, even matching that of native English speakers (Collins et al., 2013; Crossley & McNamara, 2012). Similarly, their compositions receive better quality ratings as they better reflect the discourse structures and conventions of academic English (Collins et al., 2013; Leki et al., 2008). The comparable quality of FEP students' compositions to that of native English speakers has been found in the upper elementary grades (Collins et al., 2013; Leki et al., 2008) through post-secondary education (Crossley & McNamara, 2012; Leki et al., 2008). Although students' writing productivity and quality vary as a function of their English proficiency, little is known about how differences in English proficiency vary as a function of genre or the writing task.

Understanding the text

One of the challenges students face in learning to write is that writing is not an unchanging task. Instead, each act of writing is unique and informed by the students' knowledge of the world, the topic, discourse structures, and text forms appropriate for their communication goals and intended audience (Bazerman et al., 2017). A critical way that texts vary is by genre; the conventional structures used to create texts for different purposes (Biber & Conrad, 2019). The three primary genres are taught in schools—narrative, explanation, and argumentation—each has its own conventions for organizing information and using language through writing (Llosa et al., 2011; Schleppegrell, 2004). Narratives focus on the actions of characters confronting problems and retelling a sequence of events (Schleppegrell, 2004). The explanation genre is used to describe and interpret phenomena and uses logical rather than temporal organization. Finally, argumentation involves presenting a point of view and the use of claims supported by examples and evidence (Schleppegrell, 2004). Thus, to write effectively, students must learn the purposes, rhetorical structures, grammatical conventions, and word usage that characterize each genre.

Overall, students show tremendous variability in their writing performance across genres (Bouwer, Béguin, Sanders, & van den Bergh, 2015; Graham, Fitzgerald, et al., 2016a; Graham, Hebert, et al., 2016b), with student writing showing low to moderate correlations across genres (Beers & Nagy, 2011; Olinghouse & Wilson, 2013). These differences may reflect variations in students' understanding of each genre's linguistic, discourse, and structural features (Brisk, 2012; Graham, Fitzgerald, et al., 2016; Graham, Hebert, et al., 2016b). In general, students show greater knowledge of the narrative genre than the more cognitively demanding and linguistically complex explanatory or argumentative genres (Beers & Nagy, 2011; Gillespie, Ollinghouse & Graham, 2013). Further, children's understanding of genres has been found to transfer across languages (Brisk, 2012), with primary-school-aged bilingual children using the same macrostructures for narratives in both languages (Bohnacker, 2016).

Students' formal schooling and informal experiences at home shape their genre knowledge. Formal schooling shapes children's understanding of informational genres' purposes and linguistic conventions (de Oliveira & Lan, 2014). As students progress from primary through the secondary grades, their use of academic vocabulary increases to a greater extent for non-narrative than narrative writing (Durrant & Brenchley, 2019). A second, critical source of variation among the discourse and linguistic structures of narratives is culture (Clyne, 1981). School-aged, language minority children may initially use their home cultures' discourse structures in their writing and oral narratives (Danzak, 2011; Schick, & Melzi, 2010). Language minority adolescents with emerging English proficiency have used the same knowledge-telling strategies for both narratives and expository prompts (Danzak, 2011). With greater schooling and English proficiency, Generation 1.5 and FEP students' writing better reflects each genre's grammatical and discourse structures (Brisk, 2012; Leki et al., 2008). However, much of the research exploring the writing development of language minority students tend to target one genre at a time (Beck et al., 2013) or treat language minority students as a homogeneous group (Danzak, 2011). Consequently, we seek to examine whether language minority students, with varying degrees of English proficiency, show systematic differences in their writing achievement as a function of genre.

Understanding the writing task

A third factor that makes learning to write more challenging is variation in the requirements of writing assignments. For example, writing assignments may require students to rely solely on their personal experiences and background knowledge to inform their writing (nonsource-based writing) or draw their evidence and examples from textual sources (source-based writing). Source-based and nonsource-based writing share common skills, such as requiring students to construct knowledge, organize their ideas, and translate their ideas into coherent paragraphs (Flower & Hayes, 1981). However, the two tasks also have different affordances and tap different cognitive skills. Because nonsource-based writing enables students to draw from their prior knowledge and experiences for their writing, it frees them from the taxing demands of reading and representing one or more documents, and sometimes

synthesizing contradictory information across texts (Britt et al., 1999; Cumming et al., 2016). Indeed, source-based writing quality shares moderate correlations with the degree to which students engage in strategic processing while reading the texts (Anmarkrud et al., 2014). Because LEP students struggle in comprehending the source texts, they may be less effective in writing cohesive arguments and source-based essays (Cumming et al., 2016; Plakans & Gebril, 2012).

Textual sources may scaffold student writing, providing students with background knowledge for unfamiliar topics and vocabulary to enhance their writing (Gebril & Plakans, 2016). Further, textual sources may provide students with ideas for their essays and models of how to translate their ideas into words and structure their arguments (Gebril & Plakans, 2016; Ong & Zhang, 2010). Text-based sources' affordances may be particularly beneficial for LEP students, as they often experience difficulties in finding the words and grammar to express their thoughts (Beck et al., 2013). Therefore, although the cognitive demands of comprehending and representing multiple texts may differentially impede source-based writing for LEP students, those textual sources may provide scaffolding that facilitates the writing process. Thus, we seek to examine the relationships among nonsource-based writing, source-based writing, and reading skills for students with varying English proficiency levels.

The current study

In the current study, we sought to develop a multi-faceted understanding of adolescent writing across genres and across writing tasks among linguistically diverse students. In this study, we retrieved student performance scores on the annual state assessment of writing achievement, which required students to write source-based essays in one of three genres. Participating teachers also required their students to write an on-demand essay without the use of textual sources. This nonsource-based assignment, combined with the source-based writing assessment, allowed us to examine the contributions of individual differences, genre, and task demands to writing achievement. We operationalized individual differences in English proficiency, exploring the differences among native English speakers, language minority students who were proficient in English, and language minority students with limited English proficiency. Although prior research has explored the roles of English proficiency, genre, and the use of sources in writing performance, there has been little exploration of the relationships among these factors. This study focused on addressing the following research questions:

1. Does nonsource-based writing vary as a function of individual differences in English proficiency? More specifically, do the productivity, word usage, syntactic simplicity, narrativity, and overall quality of nonsource-based, argumentative essays differ for NES, FEP, and LEP students?
2. Does source-based writing for NES, FEP, and LEP students differ for narrative, explanatory, and argumentative genres?
3. What is the relationship between nonsource-based writing and source-based writing for NES, FEP, and LEP students?

Method

Study context and participants

This study's sample was recruited from a larger randomized-control study investigating the efficacy of a reading intervention, whereby texts were digitally formatted to visually highlight syntactic structures for middle school students in southern California (Tate et al., 2019). This was a convenience sample drawn from a large, diverse urban school district. While all 54 middle school English Language Arts (ELA) teachers from the larger study had been invited to participate in the current study, this sample consisted of the 45 teachers who had agreed to assign the nonsource-based writing task to their seventh- or eighth-grade ELA classes. Table 1 shows that the current subsample was similar to that of the larger study.

All of the district's middle schools shared a common curriculum, textbook series, pacing guide, and quarterly benchmark tests. It is noteworthy that the ELA curriculum placed a heavier emphasis on argumentative writing than narrative and explanatory writing. The district used two approaches to support students with limited English proficiency. First, teachers across all disciplines received training in integrated English Language Development, so that content-area teachers could provide students with explicit support in acquiring the English needed to engage with curricular content. Second, in addition to the regular ELA course, students with limited English proficiency were placed in an English Language Development course to receive additional systematic and explicit instruction in English. Each classroom had a technology cart with enough Chromebooks or iPads for each student.

Students

The 127 classrooms participating in the current study included 1819 students (709 7th grade and 1110 8th grade). 973 students were in classrooms with the reading intervention, and 846 students were from control classrooms in the original study.

Students varied in their English proficiency designations. When students initially enrolled in the school district, those whose primary language was not English were given the California English Language Development Test (CELDT) to determine their listening, reading, speaking, and writing proficiency in English. At this initial screening, students were designated as either fluent in their English proficiency (FEP) or as having limited English proficiency (LEP). For each subsequent year, LEP students were reevaluated until they were redesignated as FEP. Students were redesignated as FEP if their scores were above threshold on the CELDT, their scores on the annual English Language Arts assessment met standards or teacher recommendations based on student performance.

In total, 506 students were native English speakers (NES; 168 7th grade and 338 8th grade), while 1313 had primary languages other than English. Among the language minority students, 856 students were FEP (327 7th grade, 529 8th grade), and 457 were LEP (214 7th grade, 243 8th grade). Table 1 shows that, on average, both groups of language minority students were enrolled in this school district

Table 1 Characteristics of the analytic sample as a whole and as a function of English proficiency

	Original study	Current study sample	Native English speakers	Fluent English proficient	Limited English proficient
Grade 7 (percent)	40.5	39.0	33.2	38.2	46.8
Male (percent)	49.7	49.0	51.8	43.7	56.0
Hispanic (percent)	49.2	47.7	39.3	39.2	73.7
Asian (percent)	35.6	37.7	16.4	58.2	23.4
White (percent)	11.5	10.7	34.9	1.5	1.1
Other races (percent)	3.7	3.9	9.1	1.1	1.8
Socioeconomically Disadvantaged (percent)	74.1	74.8	53.6	78.4	91.[c]
Special education (percent)	4.4	4.1	5.7	1.5	7.2
Treatment group in original study (percent)	51.5	53.5	51.4	55.5	52.3
Number of years in US schools	8.40 (0.92)	8.41 (0.90)	8.55[a] (0.66)	8.43[b] (0.87)	8.24[c] (1.15)
Smarter balanced reading (max. 3)	2.19 (0.70)	2.22 (0.68)	2.22[a] (0.67)	2.47[b] (0.58)	1.74[c] (0.63)
N (Students)	3453	1819	506	856	457

Standard deviations are presented below in parentheses. Scores with different superscripts differed significantly

Scores with different subscripts differed at the $p < .001$ level

since kindergarten (over eight years). However, LEP students were more likely to be enrolled in the district more recently, $t(1816) = 3.71$, $p < 0.001$. Most language minority students spoke either Spanish (672 students) or Vietnamese (562 students). Other home languages spoken by ten or more students included Korean, Tagalog, Arabic, and Cantonese.

Measures

Demographic data

In addition to language designation, we collected student demographic data, including grade level, race, gender, and ethnicity, as shown in Table 1. Students eligible for free or reduced lunch were considered to be of low socioeconomic status ($N = 1361$).

Writing tasks

All participants completed two types of on-demand writing assessments. The nonsource-based writing prompt required students to write an argumentative essay without the use of texts as sources. The second required students to write an essay in response to three or four texts. Both writing assessments were administered digitally and were scored independently of the research team.

Nonsource-based writing

Revision Assistant is automated writing evaluation software, which uses machine learning algorithms to score essays based on a corpus of texts previously scored by human raters (Woods, Adamson, Miel & Mayfield, 2017). Revision Assistant hosts a library of writing prompts, organized by grade level and genre, that teachers can assign to their students for use independently of the Revision Assistant software (Turnitin.com). The 7th grade prompt for an argumentative essay required students to,

> Write a letter to your school newspaper explaining why you participate (or why you do not participate) in a specific sport, club, or other extracurricular activity at school. Use relevant examples from your life or from the experiences of friends, family members, or other people to make the reasons for your choice clear (Turnitin.com, 2018).

Students wrote their on-demand essays in a single session, using Google docs, and did not receive formative feedback. We manually corrected students' essays for spelling and capitalization errors and added missing periods to increase the accuracy of the automated scoring. We used Coh-Metrix 3.0 (Graesser et al., 2011; McNamara et al., 2014) to derive measures of essay length, word usage, syntax and narrativity. We also submitted essays to Revision Assistant (Woods et al., 2017) for machine scoring across the four dimensions of an analytic rubric addressing the clarity of the writing, the development of ideas, overall organization and language use.

Essay length

Reflected students' writing productivity. We used Coh-Metrix 3.0 to total the number of words as our index of essay length.

We operationalized writing quality in various ways, including word usage, syntactic complexity, the narrativity of the writing style, and scores evaluating the clarity, structure, and organization of the essays, as well as an overall writing quality score.

Word usage

Reflected the difficulty of the words students used in their writing. We derived two measures of word usage with Coh-Metrix 3.0. The first indicator was *Word Length,* or the mean number of letters in each word. The second indicator was the mean *Age of Acquisition*. Although the age of acquisition is often correlated with word frequency and familiarity, it is considered an index of word difficulty (Garlock et al., 2001).

Syntax usage

Operationalized in two ways using Coh-Metrix 3.0. First, *Words Before the Main Verb* was the mean number of words preceding the main verb of the main clause. This is thought to be an indicator of cognitive load, as more words before the main verb increases syntactic complexity (McNamara et al., 2014). Our second indicator was syntactic simplicity or the degree to which texts use longer sentences and complex and varied syntactic structures. To this end, we used the principal component *Syntactic Simplicity*. This variable is based on a z-score, with higher scores reflecting shorter sentences with simpler sentence structures and lower negative scores reflecting the use of longer sentences with complex syntactic structures.

Narrativity

Refers to the degree to which text reflects the discourse features of story-telling, such as the use of familiar words, assumptions of the audience's prior knowledge, and similarity to spoken conversation (Graesser et al., 2011). Texts with low narrativity scores reflect informational texts' discourse features and show the greater formality expected of written texts (Graesser et al., 2011). We obtained *Narrativity* scores using Coh-Metrix 3.0′s principal component, *Narrativity*. This variable is based on z-scores, so texts with high *Narrativity* scores were less formal and more conversational, while lower scores reflect greater formality.

Writing quality was also operationalized using the automated essay scores retrieved from Revision Assistant (turnitin.com). Revision Assistant's machine-scores are calculated based on training the system on a minimum of 300 human-scored essays for each prompt (Shermis, 2014). Thus, the scores Revision Assistant generated for each dimension of the rubric for the current prompt were based on algorithms derived from

human scores of middle-school students' essays written for this prompt. The first score, *Clarity and Focus,* was the degree to which essays presented a clear central idea early in the essay and focused on supporting that idea. The second, *Development,* was the degree to which essays contained facts, definitions, and information to support and develop the essay's central idea. *Organization* refers to the structure of the essay. It reflects the use of an engaging introduction, a strong conclusion, and transitions to support the coherence of the arguments. The fourth score, *Language and Style,* refers to using specific, engaging language and clear sentence structure to communicate ideas. Each dimension received a score ranging from 1 (limited mastery) to 4 (advanced mastery). We calculated an *Overall Quality* variable as the mean of *Clarity and Focus, Development, Organization,* and *Language and Style*. *Overall Quality* scores ranged from 1 to 4 (Cronbach's $\alpha = 0.94$).

Revision Assistant has reported reliabilities using quadratic weighted kappa, a measure of agreement for rating scales, ranging from 0.74 to 0.82 for the four scales of its middle-school argumentative prompts (Woods et al., 2017). Revision Assistant's reliability is comparable to that of human scorers and other commercial automated essay scoring software (Shermis, 2014). Further, Revision Assistant is predictive of ninth grade student performance on the Texas ELA assessment ($r = 0.74$; Mayfield, Adamson, Woods, Miel, Butler & Crivelli, 2018).

Source-based writing

The ELA performance assessment of the Smarter Balanced (n.d.) annual state test is administered online and provides a source-based writing prompt. This two-hour assessment includes three or four text- and video-based sources, such as articles from newspapers, scientific articles, and web posts, as well as comprehension questions and an essay prompt. Tools such as bilingual pop-up dictionaries and glossaries, highlighters, embedded notepads, and text-to-speech features are used to increase universal accessibility (Smarter Balanced, n.d.).

Students were randomly assigned to one of three prompts. One prompt required students to use the texts as sources for an *Explanatory Essay*. For example, after reading sources about sleep and naps, students may be asked to write an explanatory article for their school newspaper. A second prompt required students to use the sources for evidence to support their claims for an *Argumentative Essay*. The third prompt required students to write a *Narrative Essay*. Students were asked to write a fictional story for an online magazine, using the details and information from the sources to develop the characters, plot, and setting. Students' essays were scored by the Smarter Balanced consortium and received an overall writing score of 1 (below standards), 2 (near standards), or 3 (above standards). Smarter Balanced performance assessments have reported marginal reliabilities of 0.78 and 0.79 for seventh and eighth-grade students, respectively (CRESST, 2017). Student's scores, but not the essays, were retrieved from the school district.

To better understand students' profiles, we also retrieved their reading scores from the ELA Smarter Balanced (n.d.) annual assessment. The reading subtest is a computer-adapted assessment that incorporates the same universal accessibility tools as the writing subtest. For seventh- and eighth-grade students, this subtest

assesses students' skill at identifying key details, reasoning and using evidence, their analytic skills within and across texts, and their knowledge of word meanings, text structures, and language use (Smarter Balanced, n.d.). Students received an overall reading score of 1 (below standards), 2 (near standards), or 3 (above standards). The reading subtest's marginal reliabilities are 0.73 and 0.74 for seventh and eighth-grade students, respectively (CRESST, 2017).

Procedure

We retrieved students' Smarter Balanced reading and writing scores from the school district. The Smarter Balanced ELA assessment is administered online to all students approximately one month before the end of the school year. In the weeks after the state assessment had been completed, teachers assigned the nonsource-based writing prompt to their ELA classes. Students were given a single classroom period of approximately 50 min to write their essays in Google Docs, collected digitally by the research team. The university's Institutional Review Board provided ethical review and oversight.

Data analytic strategies

We calculated a series of regression analyses to address our three research questions. Our first research question explored whether students' nonsource-based writing profiles varied due to their English proficiency status. To this end, essay length, word usage, syntactic simplicity, narrativity, and quality scores were outcome variables. English proficiency was dummy coded, with NES students being the reference group. Student data were nested within classrooms by using dummy variables for all but one of the classrooms. Grade, gender, race, socioeconomic status, special education status, and enrollment in the larger study's treatment or control classes were covariates. For our second research question, we estimated the effects of English proficiency and genre (narrative, explanatory, and argumentation) and their interaction on students' source-based writing, using the same control variables detailed above and nesting student data within classroom scores. Genre was dummy-coded, with narratives as the reference level. For our third question, we calculated a series of regression models testing the contributions of English proficiency, genre, nonsource-based writing performance, and reading achievement to source-based writing achievement. Once again, we used the same control variables and nested student data within classroom.

Results

Preliminary descriptive analyses

Descriptive information, the means and standard deviations for the writing volume and writing quality of NES, FEP, and LEP students are summarized in Table 2.

Table 2 Writing features of students essays as a function of language group for nonsource-based and source-based writing

	Native English speakers	Fluent English proficient	Limited English proficient
Nonsource-based writing: revision assistant			
Writing volume Essay length (words)	295.37 (95.25)	290.44 (96.98)	257.75 (96.57)
Word use Word length	4.11 (0.28)	4.23 (0.32)	4.03 (0.26)
Age of acquisition	296.02 (32.93)	309.01 (34.99)	289.69 (32.00)
Syntax Words before main verb	3.95 (1.89)	3.85 (1.48)	3.80 (1.92)
Syntactic simplicity	−0.51 (0.73)	−0.39 (0.73)	−0.60 (0.92)
Narrativity Narrativity	1.59 (0.72)	1.44 (0.75)	1.88 (0.70)
Writing quality Clarity and focus (max. 4)	2.60 (0.81)	2.79 (0.85)	2.15 (0.75)
Development (max. 4)	2.78 (0.93)	2.66 (0.80)	2.07 (0.68)
Organization (max. 4)	2.74 (0.86)	2.89 (0.85)	2.28 (0.77)
Language and style (max. 4)	2.78 (0.93)	3.02 (0.88)	2.24 (0.80)
Overall quality (max. 4)	2.66 (0.80)	2.84 (0.77)	2.18 (0.68)
Multiple-source based writing: smarter balanced ELA performance assessment			
Overall SBAC writing (max. 3)	2.29 (0.65)	2.57 (0.54)	1.87 (0.58)

Standard deviations are presented below in parentheses

Table 3 shows correlations among essay length, syntactic simplicity, formality, and writing quality scores. The writing quality scales derived from Revision Assistant were highly correlated, with correlations ranging from 0.76 to 0.94. Essay length showed moderate to high correlations with the writing quality scores ($r=0.58$ to $r=0.67$). The two measures of word usage, word length, and age of acquisition, had moderate correlations with overall writing quality ($r=0.25$ and $r=0.33$, respectively). The two syntax use measures had little relationship with writing quality, with correlations ranging between $r=0.05$ to $r=0.18$. Narrativity had moderate negative correlations with the writing quality scores ($r=-0.29$ to $r=-0.35$), indicating that the use of more formal was associated with higher scores in clarity and focus, development, organization, and writing style.

Does nonsource-based writing performance vary as a function of individual differences in English proficiency?

Table 4 shows the results of estimating students' proficiency in English on essay length, word difficulty, syntactic simplicity, formality, and writing quality while

Table 3 Correlations among writing length, lexical difficulty, syntactic simplicity, formality and measures of writing quality for nonsource-based writing and source-based writing

	Essay length	Word length	Age of acquis	Words before main verb	Synt. simpl	Narr	Clarity and focus	Devt	Lang and style	Org	Overall quality
Essay length											
Word length	−.01										
Age of acquisition	.07**	.45***									
Words before main verb	.12***	.10***	.10***								
Syntactic simplicity	−.02	.20***	.06**	−.40***							
Narrativity	−.06**	−.69***	−.42***	−.06**	−.32***						
Clarity and focus	.63***	.22***	.30***	.16***	.03	−.33***					
Development	.62***	.23***	.30***	.17***	.01	−.32***	.87***				
Language and style	.58***	.27***	.31***	.16***	.08**	−.33***	.78***	.76***			
Organization	.63***	.20***	.30***	.17***	.06*	−.29***	.83***	.81***	.78***		
Overall quality	.67***	.25***	.33***	.18***	.05*	−.35***	.94***	.93***	.90***	.92***	
Source-based writing (SBAC)	.20***	.28***	.27***	.08**	.06*	−.23***	.38***	.38***	.44***	.38***	.43***

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

controlling for grade, race, gender, SES, special education status, and reading intervention exposure, as well as classroom fixed effects. We found that students' writing varied as a function of their proficiency in English. LEP students wrote shorter essays, ($ES = -0.23$, $p = 0.002$), that used shorter words ($ES = -0.14$, $p = 0.025$), were more narrative in style ($ES = 0.23$, $p < 0.001$), and received lower writing quality scores ($ES = -0.24$ to -0.40, $p < 0.001$) than their NES peers. In contrast, FEP students performed at least as well as the NES peers in their nonsource-based argumentative essays. FEP students' essays were of comparable length, syntactic complexity, and narrativity to NES students. However, FEP students' essays showed greater facility with language and style ($ES = 0.16$, $p = 0.009$), using longer ($ES = 0.28$, $p < 0.001$) more difficult words ($ES = 0.26$, $p < 0.001$). FEP students' essays also received higher quality ratings ($ES = 0.11$, $p = 0.018$).[1]

Does source-based writing performance vary as a function of individual differences in English proficiency?

Model 1, shown in Table 5, estimates the effects of students' proficiency in English on their scores on the Smarter Balanced performance assessment. When writing a source-based essay, FEP students outperformed their NES peers ($ES = 0.19$, $p < 0.001$), who, in turn, outperformed LEP students ($ES = -0.40$, $p < 0.001$).

The main effect of genre on students' source-based writing and its interaction with English proficiency are presented in Models 2 and 3.[2] Students who wrote argumentative essays obtained higher scores ($ES = 0.08$, $p = 0.034$) than those who wrote narrative essays. However, students writing explanatory and narrative essays received similar scores ($ES = -0.01$, *ns*). Model 3 shows that the interactions were not significant.

What is the relationship between nonsource-based writing and source-based writing for NES, FEP, and LEP students?

We calculated a series of regression analyses to examine the contributions of reading skills and nonsource-based writing to student achievement in source-based writing as a function of their English proficiency, which is shown in Table 6. Model 1 shows that with demographic controls, FEP students outperformed their NES peers ($ES = 0.19$, $p < 0.001$), who, in turn, had higher source-based writing scores than LEP students ($ES = -0.40$, $p < 0.001$). It is noteworthy that controlling for nonsource-based writing, reading achievement, and genre in Models 2–5 did not change this pattern. Even when controlling for students' reading achievement, nonsource-based writing skills, and genre, FEP students outperformed NES students ($ES = 0.17$

[1] For this and subsequent regression models, we found no significant interactions between English proficiency and race.

[2] Mean reading scores were the similar across the three writing prompts, $F(2,1792) = 0.81$, *ns*.

Table 4 Regressions predicting writing volume, lexical difficulty, syntactic simplicity, formality and overall writing quality for nonsource-based essays

	Essay length	Word length	Age of acquis	Words before main verb	Syntac. Simp	Narr	Clarity and focus	Devt	Lang and style	Org	Overall quality
FEP students	−0.03 (6.38)	0.28*** (0.02)	0.26*** (2.24)	0.10 (0.12)	0.08 (0.06)	−0.11 (0.05)	0.13* (0.06)	0.08 (0.05)	0.16** (0.06)	0.10 (0.05)	0.11* (0.05)
LEP students	−0.23** (7.07)	−0.19** (0.02)	−0.11 (2.45)	0.01 (0.12)	−0.14 (0.06)	0.30*** (0.05)	−0.34*** (0.06)	−0.37*** (0.05)	−0.40*** (0.06)	−0.32*** (0.06)	−0.38*** (0.05)
Hispanic	−0.35*** (8.90)	−0.12 (0.03)	−0.21* (3.08)	−0.19* (0.16)	−0.03 (0.07)	0.18 (0.07)	−0.36*** (0.07)	−0.35*** (0.07)	−0.34*** (0.08)	−0.44*** (0.07)	−0.40*** (0.07)
Asian	0.07 (9.15)	0.28** (0.03)	0.16 (3.27)	0.08 (0.17)	0.11 (0.08)	−0.19* (0.07)	0.17 (0.08)	0.23** (0.70)	0.26** (0.08)	0.13 (0.08)	0.21** (0.07)
Other Races	−0.20 (13.80)	0.12 (0.04)	−0.36** (4.77)	−0.16 (0.25)	0.08 (0.11)	−0.01 (0.10)	−0.17 (0.12)	−0.19 (0.11)	−0.00 (0.12)	−0.21 (0.12)	0.02 (0.10)
Grade	−0.11* (4.62)	0.34*** (0.01)	0.16** (1.60)	0.09 (0.08)	0.03 (0.04)	−0.20*** (0.04)	0.00 (0.04)	0.06 (0.04)	0.02 (0.04)	−0.01 (0.04)	0.02 (0.03)
Male	−0.30*** (4.47)	−0.18*** (0.01)	−0.30*** (1.55)	−0.05 (0.08)	−0.10* (0.05)	0.19*** (0.03)	−0.40*** (0.04)	−0.38*** (0.03)	−0.40*** (0.04)	−0.35*** (0.04)	−0.42*** (0.03)
Socio-economically disadvantaged	−0.14* (5.66)	−0.01 (0.02)	−0.01 (1.96)	−0.09 (0.10)	0.05 (0.05)	0.02 (0.09)	−0.19*** (0.05)	−0.17** (0.04)	−0.18*** (0.05)	−0.18*** (0.05)	−0.20*** (0.04)
Special education	0.02 (11.28)	−0.31** (0.02)	0.02 (3.90)	−0.19 (0.20)	−0.04 (0.09)	0.20 (0.09)	−0.22* (0.09)	−0.13 (0.08)	−0.17 (0.10)	−0.19 (0.09)	−0.19 (0.08)
Treatment	0.01 (4.47)	0.04 (0.01)	−0.01 (1.55)	0.02 (0.08)	0.06 (0.04)	−0.04 (0.03)	0.06 (0.04)	0.03 (0.03)	0.05 (0.04)	0.02 (0.04)	0.05 (0.03)
Constant	241.17*** (10.20)	4.08*** (0.03)	301.96*** (3.53)	4.17*** (0.18)	−0.47*** (0.08)	1.56*** (0.08)	3.00*** (0.09)	2.83*** (0.08)	3.12*** (0.09)	3.18*** (0.09)	3.03*** (0.08)
R^2	.07	.15	.12	.02	.02	.10	.20	.20	.24	0.18	.24

All coefficients are standardized beta weights. Classroom variables for fixed effects are not shown in the table but are included in the analyses. Standard errors are in parentheses

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

Table 5 Regressions predicting source−based writing as a function of English proficiency and genre

	Model 1	Model 2	Model 3
FEP students	0.19***	0.20***	0.20***
	(0.04)	(0.04)	(0.04)
LEP students	−0.40***	−0.41***	−0.41***
	(0.04)	(0.04)	(0.04)
Genre			
Argumentative essays		0.08*	0.09*
		(0.04)	(0.06)
Explanatory essays		−0.01	0.08
		(0.03)	(0.07)
Interactions			
FEP X argumentative			0.04
			(0.08)
FEP X explanatory			−0.09
			(0.08)
LEP X argumentative			0.05
			(0.10)
LEP X explanatory			−0.22
			(0.10)
Demographic controls			
Hispanic	−0.08	−0.06	−0.15
	(0.06)	(0.05)	(0.5)
Asian	0.18***	0.19***	0.38***
	(0.06)	(0.06)	(0.06)
Other races	0.03	0.02	0.07
	(0.08)	(0.08)	(0.08)
Grade	−0.06**	−0.17**	−0.35**
	(0.07)	(0.07)	(0.07)
Male	−0.08***	−0.05***	−0.17***
	(0.03)	(0.02)	(0.03)
Socioeconomically disadvantaged	−0.67***	−0.73***	−0.15**
	(0.03)	(0.03)	(0.03)
Special education	−0.28***	−0.25***	−0.32**
	(0.07)	(0.07)	(0.07)
Treatment	0.05	0.04	0.06
	(0.03)	(0.03)	(0.03)
Constant	2.23***	2.21***	2.50***
	(0.14)	(0.14)	(0.12)
R^2	.32	.33	.33

All coefficients are standardized beta weights. The reference group for genre was Narrative essays. Classroom variables for fixed effects are not shown in the table but are included in the analyses. Standard errors are in parentheses

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

to $ES = 0.25$), who had higher source-based writing scores than LEP students ($ES = -0.26$ to $ES = -0.40$).

Table 6 Regressions predicting contributions of reading skills and nonsource-based writing to source-based writing as a function of English proficiency

	Model 1	Model 2	Model 3	Model 4	Model 5
Language designation					
FEP students	0.19*** (0.04)	0.25*** (0.04)	0.19*** (0.04)	0.18*** (0.03)	0.17*** (0.03)
LEP students	−0.40*** (0.04)	−0.45*** (0.04)	−0.30*** (0.04)	−0.26*** (0.04)	−0.27*** (0.04)
Literacy skills					
Nonsource-based writing		0.25*** (0.06)		0.17*** (0.02)	0.17*** (0.02)
Smarter balanced reading achievement			0.41*** (0.02)	0.37*** (0.02)	0.37*** (0.02)
Essay type					
Argumentative					0.14** (0.03)
Explanatory					0.01 (0.03)
Demographic controls					
Hispanic	−0.08 (0.05)	−0.06 (0.05)	−0.06 (0.05)	0.00 (0.05)	0.01 (0.05)
Asian	0.18*** (0.06)	0.37*** (0.05)	0.29*** (0.05)	0.27*** (0.05)	0.27*** (0.05)
Other races	0.03 (0.08)	0.11 (0.08)	0.08 (0.07)	0.10 (0.07)	0.11 (0.07)
Grade	−0.06** (0.07)	−0.16*** (0.02)	−0.14*** (0.03)	−0.14*** (0.02)	−0.15*** (0.02)
Male	−0.08*** (0.03)	−0.07 (0.03)	−0.18*** (0.02)	−0.11** (0.02)	−0.11** (0.02)
Socioeconomically disadvantaged	−0.67** (0.03)	−0.12* (0.03)	−0.12* (0.03)	−0.09 (0.03)	−0.08 (0.03)
Special education	−0.28** (0.07)	−0.30** (0.06)	−0.14 (0.06)	−0.13 (0.06)	−−0.13 (0.06)
Treatment	0.05 (0.03)	0.09* (0.03)	0.07 (0.02)	0.07 (0.02)	0.06 (0.02)
Constant	2.23*** (0.14)	1.83*** (0.10)	1.51*** (0.07)	1.18*** (0.10)	1.14*** (0.08)
R^2	.32	.38	.41	.43	.44

All coefficients are standardized beta weights. Classroom variables for fixed effects are not shown in the table but are included in the analyses. Standard errors are in parentheses

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

Models 2 through 4 present the contributions of nonsource-based writing and reading achievement to source-based writing performance. In Model 2, nonsource-based writing on its own made a significant contribution to source-based writing achievement ($ES = 0.25$, $p < 0.001$), while Model 3 shows reading achievement made its own contribution to source-based writing achievement ($ES = 0.41$, $p < 0.001$).

Model 4 shows that when both nonsource-based writing and reading achievement are entered into the equation, reading achievement's contribution to source-based writing ($ES = 0.37$, $p < 0.001$) was more than twice that of nonsource-based writing ($ES = 0.17$, $p < 0.001$). These relations held, even when controlling for genre, as shown in Model 5.

Finally, Model 5 presents the contribution of genre to source-based writing when controlling for students' nonsource-based writing and reading achievement. While students who argumentative essays obtained higher scores than students receiving narrative writing prompts ($ES = 0.14$, $p = 0.005$), students writing explanatory and narrative essays showed comparable performance ($ES = 0.01$, $p = 0.78$).

Discussion

Writing is a complex task that can best be understood when adopting a multi-faceted approach, with performance influenced by variations across individuals, the genre, and the writing assignment's constraints. More specifically, we explored middle school students' writing performance with varying degrees of English proficiency across genres and for both nonsource-based and source-based writing prompts. We discuss each factor and its contributions to writing performance below.

Understanding the writer

At the heart of our model is the writer. Individual differences, particularly in their English proficiency, were robust contributors to writing performance even after controlling for student-level variables, such as gender, socioeconomic status, special education status, and reading achievement. Whereas LEP students showed the weakest performance across writing tasks, FEP students performed at least as well, but in many cases, better than NES students.

Although most language minority students in this study enrolled in the school district in kindergarten, receiving on average eight years of formal schooling in English-dominant schools, approximately 35% continued to be categorized as LEP, or long-term English learners (Estrada & Wang, 2018). Overall, LEP students showed weaker performance in both types of writing tasks than NES students. LEP students wrote shorter nonsource-based essays that were less formal and used simpler vocabulary (Collins et al., 2013; Reynolds, 2005). Although LEP students' writing has been found to use less complex syntax in past research (Collins et al., 2013; Hinkel, 2003; Kormos, 2011), we found little difference in their syntax use than their NES peers. LEP students also showed higher-order difficulties in their nonsource-based writing, showing weaker organization, elaboration and examples, and overall quality (Collins et al., 2013; Leki et al., 2008; Perin et al., 2017). These difficulties may reflect the greater cognitive demands involved in composing when processing English's linguistic features have not yet been automatized (Perin et al., 2017). Alternately, these differences may reflect diminished opportunities to learn, as long-term English learners often experience limited access to the mainstream curriculum,

instruction focused on low-level skills through remedial instruction, and linguistic isolation (Estrada & Wang, 2018). We attempted to mitigate differences in educational experiences by limiting our sample to LEP students enrolled in general ELA classes, where students must have sufficient English proficiency to access the curriculum with scaffolding by the general education teacher.

In contrast, FEP students performed as well or better than native English speakers. Like elementary-school-aged students (Carlisle, 1989; Collins et al., 2013), we found that FEP students in middle school wrote essays of comparable length to NES students. However, unlike FEP and NES elementary-school-aged students whose writing was of similar quality (Collins et al., 2013), we found that middle school-aged FEP students wrote higher quality essays than NES students, even when controlling for reading skill. FEP students' superior writing quality reflected Crossley et al. (2014) lexical style approach to successful writing. Their more robust writing performance was characterized by greater mastery of word usage, rather than differences in syntactic complexity, formality, or organization for their nonsource-based essays. Our findings extend a small yet growing body of literature examining the writing performance of Generation 1.5, which is often mixed and primarily focuses on post-secondary students (Doolan, 2013, 2017; Perin et al., 2017).

Understanding genre

A second consideration was the genre of the text, which was only possible for source-based writing. We found that students who wrote argumentative essays were more successful than those who wrote narrative or explanatory essays, despite the three groups having comparable reading comprehension scores. Although argumentative writing is thought to be more cognitively and linguistically advanced than narrative and explanatory writing (Beers & Nagy, 2011; Schleppegrell, 2004), this school districts' emphasis on argumentation as part of the ELA curriculum may have contributed to students' more robust performance in argumentative writing.

One caveat is that students were randomly assigned to write a single source-based essay using one of the three genres rather than all three genres. In contrast, past examinations of genre required students to write two or more essays in different genres (Beers & Nagy, 2011; Graham, Fitzgerald, et al., 2016a; Graham, Hebert, et al., 2016b; Olinghouse & Wilson, 2013), controlling for both prompt effects and confounds with individual differences. Despite these studies involving elementary school-aged children and nonsource-based prompts, our sample of linguistically diverse adolescents replicated variations in writing performance across genres. Further, differences across genres did not interact with language proficiency. Our findings add to a growing body of literature that suggests explicit instruction in argumentation, and other genres may be beneficial for students with varying levels of proficiency in English (O'Hallaron, 2014; Olson et al., 2015).

Understanding the writing task

The third dimension we considered was the constraints of the writing task itself, such as the requirement of using textual sources for examples and evidence. We found that students' English proficiency was an important contributor to their essays' quality both when textual sources were required, and when students used personal experience and background knowledge as their sources. However, because only the scores, and not the essays themselves, were available for the Smarter Balanced writing assessment, we could not make direct comparisons between source-based and nonsource-based writing using the same variables. Although we cannot make conclusions about relative performance across tasks, the variety of writing outcomes across the source-based and nonsource-based writing tasks yielded a common finding: FEP students wrote higher quality essays than NES students, who outperformed LEP students.

We did examine the relationships among source-based and nonsource-based writing and reading comprehension. Overall, the correlations between nonsource-based writing volume, language use, and overall quality with multiple source-based writing were moderate in size, supporting the view that both types of writing tasks share common skills (Graham et al., 2018). However, we found that reading comprehension's contribution to source-based writing was almost double that of nonsource-based writing. These findings support the view that although common skills may be involved in both types of writing tasks, these tasks' specific requirements may influence the cognitive, linguistic, and literacy skills tapped by each task (Bazerman et al., 2017). Writing assignments requiring the use of textual sources may influence students' writing performance through skill at comprehending and representing the sources and their use of the texts to scaffold word usage, the generation of ideas, and models for structuring their writing. These findings may explicate why studies have found that interventions focused on improving reading skills have led to gains in writing skills (Graham et al., 2018; Tate et al., 2019). Further, they support the pedagogical recommendation that reading and writing instruction should be integrated (Graham et al., Graham, Fitzgerald, et al., 2016).

Limitations and future directions

This study was conducted within the context of a larger, randomized control trial (Tate et al., 2019). As such, we were limited to the school district's measures as part of the annual assessments of literacy proficiency. Because of the Smarter Balanced assessment's proprietary nature, we were unable to obtain students' source-based essays, only the scores released to the district. Consequently, we could not use Coh-Metrix to derive the same language use measures to compare the two types of writing tasks directly. We encourage future research to administer both source-based and nonsource-based writing prompts, across multiple genres, to test the relations among individual differences, genre, task constraints, and contributions to writing performance.

Conclusions

Writing is a complex and multidimensional skill, and at the heart of our model is the individual writer. Language minority students are a heterogeneous group, and their English proficiency is a robust contributor to their writing skills. Whereas long-term English learners were less skilled in writing English essays, in middle school, language minority students proficient in English, or those considered Generation 1.5, wrote higher quality essays than native English speakers, even after controlling for reading skills. However, understanding writing achievement requires more than just investigating individual differences in the writer. Student achievement also varied as a function of genre, with students at all levels of English proficiency showing greater success in writing the essays in the genre that received greater instructional emphasis—argumentation. Further, although writing across tasks involves a common set of skills, it is not a monolithic skill. Instead, the writing task's nature, such as whether or not students are required to use textual sources, may influence the cognitive and linguistic skills recruited for writing.

References

Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences, 30,* 64–76. https://doi.org/10.1016/j.lindif.2013.01.007

Bazerman, C., Applebee, A. N., Berninger, V. W., Brandt, D., Matsuda, P. K., Murphy, S., et al. (2017). Taking the long view on writing development. *Research in the Teaching of English, 51,* 351–360.

Beck, S. W., Llosa, L., & Fredrick, T. (2013). The challenges of writing exposition: lessons from a atudy of ELL and non-ELL high school students. *Reading and Writing Quarterly, 29,* 358–380. https://doi.org/10.1080/10573569.2013.758938

Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: syntactic complexity and genre differentiation. *Reading and Writing: An Interdisciplinary Journal, 24,* 183–202. https://doi.org/10.1007/s11145-010-9264-9

Biber, D., & Conrad, S. (2019). *Register, genre and style*. Cambridge University Press.

Bohnacker, U. (2016). Tell me a story in English or Swedish: narrative production and comprehension in bilingual preschoolers and first graders. *Applied Psycholinguistics, 37,* 19–48. https://doi.org/10.1017/S0142716415000405

Bouwer, R., Béguin, A., Sanders, T., & van den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing, 32,* 83–100. https://doi.org/10.1177/0265532214542994

Brisk, M. E. (2012). Young bilingual writers' control of grammatical person in different genres. *Elementary School Journal, 112,* 445–468.

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: essays in honor of Tom Trabasso* (pp. 209–233). Erlbaum.

Carlisle, R. S. (1989). The writing of Anglo and Hispanic elementary school students in bilingual, submersion, and regular programs. *Studies in Second Language Acquisition, 11,* 257–280. https://doi.org/10.1017/S0272263100008135

Clyne, M. (1981). Culture and discourse structure. *Journal of Pragmatics, 5,* 61–66.

Collins, P., Hwang, J. K., Zheng, B., & Warschauer, M. (2013). Writing with laptops: a quasi-experimental study. *Writing and Pedagogy, 5*(2), 203–230. https://doi.org/10.1558/wap.v5i2.203

CRESST (2017). *Smarter balanced assessment consortium: summative technical report*. Retrieved April 20, 2020 from https://portal.smarterbalanced.org/library/en/2015-16-summative-technical-report.pdf

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: the roles of cohesion and linguistic sophistication. *Journal of Research in Reading, 35*(2), 115–135. https://doi.org/10.1111/j.1467-9817.2010.01449.x

Crossley, S. A., Roscoe, R., & McNamara, D. S. (2014). What is successful writing? An investigation into the multiple ways writers can write successful essays. *Written Communication, 31,* 184–214. https://doi.org/10.1177/0741088314526354

Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: a synthesis of recent research. *Journal of English for Academic Purposes, 23,* 47–58. https://doi.org/10.1016/j.jeap.2016.06.002

Danzak, R. L. (2011). The integration of lexical, syntactic, and discourse features in bilingual adolescents' writing: an exploratory approach. *Language, Speech, and Hearing Services in Schools, 42,* 491–505. https://doi.org/10.1044/0161-1461(2011/10-0063)

De Oliveira, L. C., & Lan, S.-W. (2014). Writing in an upper elementary classroom: a genre-based approach to teaching English language learners. *Journal of Second Language Writing, 25,* 23–39. https://doi.org/10.1016/j.jslw.2014.05.001

Doolan, S. M. (2013). Generation 1.5 writing compared to L1 and L2 writing in first-year composition. *Written Communication, 30,* 135–163. https://doi.org/10.1177/0741088313480823

Doolan, S. M. (2017). Comparing patterns of error in Generation 1.5, L1, and L2 first-year composition writing. *Journal of Second Language Writing, 35,* 1–17. https://doi.org/10.1016/j.jslw.2016.11.002

Durrant, P., & Brenchley, M. (2019). Development of vocabulary sophistication across genres in English children's writing. *Reading and Writing: An Interdisciplinary Journal, 32,* 1927–1953. https://doi.org/10.1007/s11145-018-9932-8

Estrada, P., & Wang, H. (2018). Making English learner reclassification to fluent English proficient attainable or elusive: when meeting criteria is and is not enough. *American Educational Research Journal, 55,* 207–242. https://doi.org/10.3102/0002831217733543

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365. https://doi.org/10.2307/356600

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45,* 468–492. https://doi.org/10.1006/jmla.2000.2784

Gebril, A., & Plakans, L. (2016). Source-based tasks in academic writing assessment: lexical diversity, textual borrowing and proficiency. *Journal of English for Academic Purposes, 24,* 78–88. https://doi.org/10.1016/j.jeap.2016.10.001

Gillespie, A., Olinghouse, N. G., & Graham, S. (2013). Fifth-grade students' knowledge about writing process and writing genres. *The Elementary School Journal, 113,* 565–588. https://doi.org/10.1086/669938

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: providing multilevel analyses of text characteristics. *Educational Researcher, 40,* 223–234. https://doi.org/10.3102/0013189X11413260

Graham, S., Harris, K. R., & Santangelo, T. (2015). Research-based writing practices and the common core: meta-analysis and meta-synthesis. *The Elementary School Journal, 115*(4), 498–522. https://doi.org/10.1086/681964

Graham, S., Fitzgerald, J., Friedrich, L. D., Greene, K., Kim, J. S., & Olson, C. B. (2016). *Teaching secondary students to write effectively* (IES Practice Guide No. NCEE 2017-4002; p. 97). Washington, D.C.: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences.

Graham, S., Hebert, M., Paige Sandbank, M., & Harris, K. R. (2016). Assessing the writing achievement of young struggling writers: application of generalizability theory. *Learning Disability Quarterly, 39,* 72–82. https://doi.org/10.1177/0731948714555019

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., et al. (2018). Reading for writing: a meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88,* 243–284. https://doi.org/10.3102/0034654317746927

Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing, 9*(2), 123–145. https://doi.org/10.1016/S1060-3743(00)00019-9

Hinkel, E. (2003). Simplicity without elegance: features of sentences in L1 and L2 academic texts. *TESOL Quarterly, 37*(2), 275. https://doi.org/10.2307/3588505

Hwang, J. K., Lawrence, J. F., Mo, E., & Snow, C. E. (2015). Differential effects of a systematic vocabulary intervention on adolescent language minority students with varying levels of English proficiency. *International Journal of Bilingualism, 19*(3), 314–332. https://doi.org/10.1177/1367006914521698

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20*(2), 148–161. https://doi.org/10.1016/j.jslw.2011.02.001

Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System, 31*(4), 537–561. https://doi.org/10.1016/j.system.2003.02.004

Leki, I., Cumming, A. H., & Silva, T. (2008). *A synthesis of research on second language writing in English*. Taylor and Francis.

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. https://doi.org/10.1016/j.asw.2011.07.001

Mayfield, E., Adamson, D., Woods, B., Miel, S., Butler, S., & Crivelli, J. (2018). Beyond automated essay scoring: Forecasting and improving outcomes in middle and high school writing. In: *Companion proceedings of the 8th international conference on learning analytics and knowledge* (pp. 1–8).

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

O'Hallaron, C. L. (2014). Supporting fifth-grade ELL's argumentative writing development. *Written Communication, 3,* 304–331. https://doi.org/10.1177/0741088314536524

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing: An Interdisciplinary Journal, 26,* 45–65. https://doi.org/10.1007/s11145-012-9392-5

Olson, C. B., Scarcella, R., & Matuchniak, T. (2015). English learners, writing, and the Common Core. *Elementary School Journal, 115,* 570–592. https://doi.org/10.1086/681235

Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing, 19,* 218–233. https://doi.org/10.1016/j.jslw.2010.10.003

Online Practice Test Scoring Guides and Directions for Administration (DFAs). (2020). Retrieved 23 Dec 2020. https://www.caaspp.org/ta-resources/practice-training.html

Perin, D., De La Paz, S., Piantedosi, K. W., & Peercy, M. M. (2017). The writing of language minority students: a literature review on its relation to oral proficiency. *Reading and Writing Quarterly, 33,* 465–483. https://doi.org/10.1080/10573569.2016.1247399

Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing, 17,* 18–34. https://doi.org/10.1016/j.asw.2011.09.002

Ragan, A., & Lesaux, N. (2006). Federal, state, and district level English language learner program entry and exit requirements: effects on the education of language minority learners. *Education Policy Analysis Archives, 14,* 20. https://doi.org/10.14507/epaa.v14n20.2006

RAND. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. RAND.

Reynolds, D. W. (2005). Linguistic correlates of second language literacy development: evidence from middle-grade learner essays. *Journal of Second Language Writing, 14*(1), 19–45. https://doi.org/10.1016/j.jslw.2004.09.001

Schick, A., & Melzi, G. (2010). The development of children's oral narratives across contexts. *Early Education and Development, 21,* 293–317. https://doi.org/10.1080/10409281003680578

Schleppegrell, M. (2004). *The language of schooling: a functional linguistics perspective*. Lawrence Erlbaum Associates.

Shermis, M. D. (2014). State-of-the-art automated essay scoring: competition, results, and future directions from a United States demonstration. *Assessing Writing, 20,* 53–76. https://doi.org/10.1016/j.asw.2013.04.001

Tate, T. P., Collins, P., Xu, Y., Yau, J. C., Krishnan, J., Prado, Y., et al. (2019). Visual-syntactic text format: improving adolescent literacy. *Scientific Studies of Reading, 23,* 287–304. https://doi.org/10.1080/10888438.2018.1561700

Woods, B., Adamson, D., Miel, S., & Mayfield, E. (2017). Formative essay feedback using predictive scoring models. In: *Proceedings of the International Conference on Knowledge Discovery and Data Mining* (pp. 2071–2080). ACM