

Who Gets to Read What? Tracking, Instructional Practices, and Text Complexity for Middle School Struggling Readers

Laura Northrop

Cleveland State University, Ohio, USA

Sean Kelly

*University of Pittsburgh, Pennsylvania,
USA*

ABSTRACT

In this study, the authors used the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 to examine the instructional time allocation and instructional practices in eighth-grade English language arts classes for struggling readers, as measured by track level. The authors also analyze the titles and text complexity of the last three books assigned to students. The authors found that track level continues to be a significant predictor of what happens instructionally in the classroom. Struggling readers placed in low-track classes spent a larger portion of class time on skills and strategy instruction, completing worksheets, watching videos, and reading aloud than students in grade-level classes. Students in high-track classes spent more time on literature analysis, comprehension instruction, and group projects and were more frequently assigned homework than students in grade-level classes. Although there was considerable overlap in the text complexity and the text titles of books assigned at each track level, students in low-track classes read less challenging texts than students in grade-level or above-grade-level classes. Regression models controlling for a variety of student, teacher, and school variables, including student achievement, show that these adjustments in class time allocation, instructional practices, and text complexity go above and beyond what would be expected based on student achievement alone.

Educational policymakers and practitioners continue to be concerned about the low reading achievement of middle school students. The most recent National Assessment of Educational Progress (NAEP) data show that only 36% of eighth-grade students were reading at or above proficiency and that 24% could not even comprehend at a basic level (National Center for Education Statistics [NCES], 2017). Although much attention has been focused on early struggling readers (Foorman et al., 2016; Shanahan et al., 2010; Snow, Burns, & Griffin, 1998), there has been less agreement and fewer empirical studies on how to best instruct struggling middle school students (Allington, 2007). Whereas middle school proponents have advocated for diversified, individualized curriculum that meets students at their learning level (Association for Middle Level Education, 2010), in practice, the structural organization of middle schools, coupled with the curricular demands, often leads to the use of tracking as a way to sort students by achievement level, a class-level individualization of the curriculum that may actually have unintended negative consequences for struggling readers.

The purpose of this study was to investigate how eighth-grade teachers vary their instruction and curriculum in response to track

level, impacting struggling readers. As a widespread organizational practice (Domina, Hanselman, Hwang, & McEachin, 2016; Mickelson, 2015), tracking impacts instruction, learning, and educational inequality, although the effects of ability grouping in elementary school diverge from the effects of between-class tracking in middle and secondary schools. In elementary school, particularly with respect to reading instruction, ability grouping is widely considered an effective practice, and all students, including low- and high-achieving students, appear to benefit on average from differentiated instruction (Hong & Hong, 2009; Slavin, 1987; Steenbergen-Hu, Makel, & Olszewski-Kubilius, 2016). For example, in the effective comprehensive school reform model *Success for All*, students are moved out of their regular age-graded class to receive a block of reading instruction targeted to their performance level (Borman et al., 2005).

In contrast, tracking between classrooms at the middle school level tends to increase inequality in secondary schools without a pronounced effect on mean levels of achievement growth (Carbonaro, 2005; Gamoran, 1987; Montt, 2011; Shavit & Müller, 2000; Van de Werfhorst & Mijs, 2010). Building on Allington's (2007, 2011) conceptual focus on incremental challenge in reading development, this study was motivated by the concern that tracking has particularly strong effects on opportunity to learn for struggling readers. In middle school, one organizational model is to have struggling readers attend both a regular English language arts (ELA) class, in which the main focus is the required state standards in ELA, and a reading intervention class, in which the main focus is typically on reading skills and strategies (Northrop, 2018). Yet, to address reading challenges that remain by middle school, it is important that struggling readers are well served not only in their intervention class but also in their regular instructional period (Biancarosa & Snow, 2006). Prior research has shown that in ELA, tracking in secondary school leads to differences in instructional practices and course content between the varying track levels, a phenomenon well summarized by Oakes (2005) and in additional studies subsequently discussed.

Are the instructional differences across tracks often found in the literature, and which we investigated further in this study, simply appropriate adjustments to meet students' learning needs? Overall, we hypothesized that teachers may overadjust the curriculum in tracked classrooms. In a study of the effect of tracking on teacher expectations, Kelly and Carbonaro (2012) conceptualized this overadjustment as a social categorization effect; tracking provides a powerful institutional status label that may influence teachers' approach to instruction (see also Pallas, Entwistle, Alexander, & Stluka, 1994). This theory is rooted in basic sociopsychological theories of

social perception, cognition, and particularly confirmation bias (Nickerson, 1998). Categorization is a basic mental process affecting the inferences made about others, and track labels may provide teachers with a prototypical abstraction or schema of student ability (Fiske & Taylor, 1991; Hamilton, 1981).

For example, in their study of secondary English instruction, Caughlan and Kelly (2004) argued that track labels may reinforce and activate the cultural models (Holland, Skinner, Lachicotte, & Cain, 1998) that secondary school teachers hold of high- and low-track students, which in turn can result in instruction that is less engaging and coherent for low-track students. Likewise, interviews with 25 sixth-grade teachers from one school district in Texas found similar results and showed that the teachers had lowered expectations, assigned less work in class and less homework, and adjusted their assignments to make the work less demanding for students in their non-honors classes (Worthy, 2010).

Importantly, although we anticipated some overadjustment of the curriculum in response to tracking, this is not a wholesale characterization of teacher beliefs and perceptions of student achievement as flawed. Indeed, much research has pointed instead to the general accuracy of teacher perceptions (Jussim & Harber, 2005). Rather, here we highlight the relation between track level, as a particularly strong institutional status label (Karlson, 2015; Legette, 2018; Van Houtte, Demanet, & Stevens, 2013), and specific dimensions of instruction, such as course materials and instructional practices.

In this study, we used the final wave of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) to examine how teachers varied their instructional time allocation and instructional practices in eighth-grade ELA classes. Although this research question would be best fit by a multilevel analysis approach, where the nesting of students in classrooms/teachers was explicitly modeled, by eighth grade, the ECLS-K had an insufficient sample to run multilevel models. However, the unique features of this data set—namely, that it provides nationally representative data on instructional practices, class allocation, and most importantly, text titles assigned to students to read—made it an important data set to explore. In particular, we provide new evidence on the text titles and text complexity that form the basis of students' reading and writing experiences. Although prior research has provided some insight into the cognitive demand of texts used in different track levels from smaller scale observational research (subsequently discussed), this is the first study to provide nationally representative estimates of text complexity differences by track level.

Disparate access to rigorous curriculum and instruction for struggling readers is especially important with the recent focus of educational policy on increasing

academic rigor in K–12 schools. Although dissatisfaction with the unchallenging nature of U.S. education is nothing new (e.g., National Commission on Excellence in Education, 1983), the past decade has seen policymakers once again focus on the rigor of curriculum, whether through state and/or locally created programs to increase algebra enrollment in middle school (Domina, McEachin, Penner & Penner, 2015; Dougherty, Goodman, Hill, Litke, & Page, 2015), revisions to state standards (Lauen & Gaddis, 2016), or the adoption of the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSSO], 2010a). In particular, the Common Core attempts to increase the rigor of middle school ELA curriculum by requiring students to read longer, more challenging texts at earlier grade levels (Hiebert & Mesmer, 2013). This academic intensification can be particularly challenging for struggling readers, making it important to reexamine the types of texts that students read and to consider how access to appropriately challenging texts may differ by track level in middle school.

Struggling Readers in Middle School

The NAEP data, which comprise both public and private school data, show that no fewer than one fourth and as many as two thirds of eighth-grade students were struggling in reading. What is less clear from the NAEP data are what aspects of the reading experience middle school students struggled with and how schools and teachers should best address those issues.

Longitudinal and Cross-Sectional Quantitative Findings

Many studies have examined reading growth over time, investigating why some students succeed and others fail in reading by the time they get to middle school. In particular, research has identified certain subgroups of students that tend to lag behind in reading: students from lower socioeconomic households (Luyten & Bruggencate, 2011; McCoach, O'Connell, Reis, & Levitt, 2006; Morgan, Farkas, & Hibel, 2008; Morgan, Farkas, & Wu, 2011), students who are speech and language impaired (Farkas, 2011), students with other learning disabilities (Schulte, Stevens, Elliott, Tindal, & Nese, 2016; Wei, Blackorby, & Schiller, 2011), and language-minority students with limited English proficiency (Kieffer, 2008, 2011; Nakamoto, Lindsey, & Manis, 2007). Other longitudinal research has focused less on child characteristics and more on the contributions of specific components of literacy to reading comprehension. Nakamoto et al., in a study of English

learners (ELs) from first to sixth grade, found that for the older students, comprehension achievement lagged behind decoding skills and that oral language was a significant predictor of comprehension in sixth grade. Likewise, Catts, Hogan, and Adlof (2005) followed students from second to eighth grade, examining the impact of word recognition and listening comprehension on comprehension. The researchers found that word recognition was more predictive of comprehension in second grade, whereas listening comprehension was more predictive of comprehension in eighth grade, suggesting that vocabulary and language comprehension play an important role in comprehension of older students.

In addition to longitudinal studies, several cross-sectional studies using data from middle school and early high school students have begun to build profiles of adolescent struggling readers and, like the longitudinal studies, have pointed first and foremost to the challenge of meaning making. Using data from 18,000 fourth-grade students, Valencia and Buly (2004) identified six profiles of struggling readers. Four of the six groups had difficulty with comprehension: automatic word callers (18%), struggling word callers (15%), slow word callers (17%), and disabled readers (9%). A similar study using latent class analysis, focusing on students entering ninth grade, classified approximately 60% of students as struggling comprehenders or low-average comprehenders (Brasseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011); furthermore, examining the underlying component skills of these readers revealed that four of the five classes struggled with comprehension-based components such as weak language comprehension and weak reading comprehension.

Qualitative Findings

Qualitative studies of reader profiles also have suggested that older struggling readers have difficulty with comprehension. Rogers et al. (2006) differentiated six types of readers in grades 4–8, three of which had comprehension-based issues, including students who could read fluently but had little to no comprehension, students who focused on literal information and comprehended sentence by sentence, and students who got the gist of the text but had difficulty with specific details. Finally, Dennis (2009), in a small case study of her own students, found that the majority of her sixth-grade students had issues with comprehension, not decoding. Thus, older struggling readers have a variety of issues, but by middle school, poor comprehension is one of the most pressing problems.

Typical Instructional Approaches

The varying nature of student instructional needs in middle school makes it difficult to summarize effective

practices in teaching; however, recommendations for effective practices in middle school ELA classes include access to diverse literature, opportunity for discussion, and engagement with higher order thinking skills; integration of reading and writing; and explicit vocabulary and comprehension instruction (Association for Middle Level Education, 2010; Biancarosa & Snow, 2006; Graham et al., 2016; Kamil et al., 2008). In addition to focusing on instructional practices and curricular materials, middle schools should think about the learning environment across class periods, creating extended time for reading, reading in content areas, and a relevant curriculum focused on student inquiry (Association for Middle Level Education, 2010; Biancarosa & Snow, 2006; Kamil et al., 2008). Thus, regardless of track levels, students should have access to an ELA curriculum that creates robust learning experiences around a diverse set of texts.

However, within this framework for general effective practices in middle school ELA, students in the low-track classrooms should have access to the appropriate instructional practices for struggling readers described next, which are based on individual strengths and weaknesses in reading, as well as the general patterns of need from the reader profiles reviewed previously. We hope that struggling readers will have access to this type of instruction not only in their supplemental intervention instruction but also in their regular ELA class. Effective instruction needs to be carefully matched to the areas of literacy the students struggle with and aligned with the classroom curriculum (Wonder-McDowell, Reutzel, & Smith, 2011). This includes learning in rich contexts with ample time to read during the school day (Allington, 2013; Risko & Walker-Dalhouse, 2015), especially with texts aligned to the curriculum that are at the students' appropriate reading levels (Allington, 2007, 2011). Finally, effective practices for struggling readers include access to strategic tutoring and individualized instructional interventions provided by trained specialists (Association for Middle Level Education, 2010; Biancarosa & Snow, 2006; Kamil et al., 2008), although these types of intervention are largely provided outside of the regular ELA class (Northrop, 2018) and thus were not systematically investigated in the ECLS-K.

In particular, middle school struggling readers benefit from targeted comprehension instruction (Edmonds et al., 2009), which includes a focus on identifying narrative and expository text structures, word meanings, and prior knowledge (Faggella-Luby & Deshler, 2008). Comprehension instruction should focus on explicit instruction on reading strategies (Allington, 2011; Edmonds et al., 2009; Risko & Walker-Dalhouse, 2015), an instructional practice that is particularly important for students with learning disabilities (Faggella-Luby &

Deshler, 2008). Additionally, struggling readers benefit from comprehension instruction that focuses on higher order thinking skills (Risko & Walker-Dalhouse, 2015; Swanson et al., 2017; Taylor, Pearson, Peterson, & Rodriguez, 2003; Vaughn et al., 2013). Focusing on higher order thinking skills includes more use of higher level questioning and less use of routine practice of skills (Sun, Anderson, Lin, & Morris, 2015; Taylor et al., 2003).

Given the wide range of instructional practices available to teachers to use with middle school students, particularly struggling readers, in this article, we use new data to examine common ways that schools and teachers respond to struggling readers in eighth grade: tracking students into classes by achievement level, adjusting instructional practices, adjusting text complexity, or some combination of all three.

Tracking and Instruction in ELA

In examining overall learning conditions in schools, research has suggested that there are more extensive differences in teacher instruction within than between schools (Kelly, 2010; Kelly & Majerus, 2011; Northrop & Kelly, 2018; Raudenbush, Rowan, & Cheong, 1993). A major source of this variation within schools is the practice of tracking. A robust body of research has suggested that tracking at the middle school level affects instructional focus, instructional practices, and course materials in ELA classes. This is particularly salient for struggling readers, who often find themselves in classrooms emphasizing instruction in decoding skills and with low reading volume (Allington, 2011).

Although overall class time is often the same within each school, teachers make choices about how to allocate their instructional time in each class period. Teachers of high-track classes spend more time studying literature and working on analytical writing, whereas low-track classes have more emphasis on grammar and narrative writing (Gamoran & Carbonaro, 2002; Oakes, 2005). Students in high-track classes are given more uninterrupted, sustained time to read in class (Worthy, 2010). In addition, the instruction in low-track classes is often more fragmented and less coherent than in high-track classes (Caughlan & Kelly, 2004; Page, 1991), with students working on isolated, basic skills (Worthy, 2010). Teachers in low-track classes spend more time on non-content-related material (Gamoran, Nystrand, Berends, & LePore, 1995) and on classroom management (Oakes, 2005).

Beyond basic time allocation, differences in instructional orientation and practices are evident in different track levels. Teachers in high-track classes spend more time on both teacher- and student-led discussion, ask

more authentic questions, allow more permeable textual discussion, and ask students to provide rigorous analysis of literature (Caughlan & Kelly, 2004; Gamoran & Carbonaro, 2002; Gamoran et al., 1995; Gritter, 2012; Worthy, 2010). In contrast, instructional practices noted in low-track classes include reading aloud, acting out characters from a book, using workbook, less permeable textual discussion and more initiation-response-evaluation talk, and completing fill-in-the-blank worksheets (Caughlan & Kelly, 2004; Gamoran, 1993; Gritter, 2012; Hodge, 2015). In addition to using different instructional practices, teachers in low-track classrooms also provide less instructional support to their students, which includes supporting content understanding through teacher depth of understanding, background knowledge, and ability to communicate the concepts to students; providing opportunities in class to use higher order thinking skills through problem solving; and quality of teacher feedback (Donaldson, LeChasseur, & Mayer, 2017). These differences in instructional practices reflect differences in the level of cognitive demand, with students in high-track classes receiving more rigorous and challenging assignments.

Likewise, differences in course materials are seen between track levels. High-track classes read longer and more difficult texts, including more classic texts (Gamoran, 1993; Worthy, 2010), whereas low-track classes are more likely to read young adult novels or use reading kits (Oakes, 2005). Additionally, students in low-track classes are often given reading assignments that are decontextualized (Worthy, 2010). As a result, students in high-track classes are exposed to high-status knowledge (Oakes, 2005), whereas students in low-track classes have fewer opportunities to encounter content that many would consider a prerequisite for attending college.

Given the achievement-driven nature of tracking, it is likely that struggling readers are more often placed in lower track classrooms, making studying the instructional practices in lower track classrooms crucial in understanding the experiences of middle school struggling readers. The functional logic of tracking entails tailoring materials and assignments to match tracked-students' current learning needs, so some adjustment of instructional approaches based on student academic readiness is desirable within the different track levels (Hallinan, 1994). However, it is important to note that the literature suggests that the instructional techniques used in lower track classrooms (less reading, less discussion, and more workbooks) are, at first glance, mismatched with the instructional practices that are recommended as most beneficial to struggling readers (more reading, more discussion, and less skills practice).

In a summative sense, it is difficult to know for sure, without robust observational data on classroom

practices and diagnostic student data, if tracked instructional practices are actually misaligned to the needs of students. However, in studies that controlled for prior student achievement, these differences in instructional focus, instructional practices, and course content were partly responsible for some of the differences in achievement growth observed between high- and low-track classrooms (Carbonaro, 2005; Gamoran, 1987), although highly detailed measures of instruction are not always available in large-scale databases. One study that examined use of classroom time in middle school ELA and teachers' value-added ratings found that teachers with a low value-added rating spent more instructional time on reading instruction, whereas teachers with a high value-added rating spent more instructional time focusing on content domains (Grossman, Loeb, Cohen, & Wyckoff, 2013). Carbonaro and Gamoran (2002), using a large-scale data set, found that instruction focused on literature study and analytical writing was positively associated with reading achievement, whereas instruction focused on grammar was negatively associated with reading achievement.

Text Complexity

One curricular aspect of tracking may be beneficial for struggling readers: access to high-quality texts at the students' individual reading levels (Allington, 2007, 2011, 2013). One way to study whether students are being assigned different-level texts based on individual reading level is to look at the text complexity of the books read in class. For any given text, complexity may be described in terms of a summary quantitative measure or by a qualitative appraisal of the structure, levels of meaning, knowledge demanded by the text, and so forth and may entail additional consideration of the reader (e.g., prior knowledge) or task components (NGA Center & CCSSO, 2010b).

In this study, we focused on quantitative text leveling, as it is widely used across elementary and middle schools both diagnostically and to set learning goals (as in the Common Core). Diagnostically, assessments such as the Measures of Academic Progress and STAR Literacy provide teachers with recommended text complexity levels for each student. Likewise, reading programs such as Accelerated Reader ask students to choose a book within a specific text level to read and are used in both regular classroom instruction and intervention in middle school curriculum (Slavin, Cheung, Groff, & Lake 2008).

Despite the extensive research that has gone into developing measures of text complexity, and their widespread use in schools and reform efforts, as far as we are aware, no prior research has, on a large scale, specifically

examined the text complexity of books read at different track levels. Although based on prior research (e.g., Gamoran, 1993), we expected to find that students in low-track classes, and thus many struggling readers, were likely to be reading texts with substantially less complexity than their counterparts in high-track classes.

Research Questions

In this study, we used the nationally representative ECLS-K data to examine measures of instruction in eighth-grade ELA, including basic time use and teachers' use of different types of instructional practices. In addition, we analyzed a core element of ELA instruction: the titles, authors, and text complexity of the last three books assigned to eighth-grade students. We used track level as a way to measure the exposure of struggling readers to various types of instructional practices and texts. Specifically, three questions guided our research:

1. How do English teachers allocate class time at each track level?
2. What instructional practices are used at each track level?
3. What texts and text complexity are used at each track level?

Data and Method

Sample

ECLS-K provides a large-scale, nationally representative data set from a NCES program of three longitudinal studies that followed students from kindergarten through eighth grade. In the present study, we focused on the final year, 2006–2007, when the students were in eighth grade. Our analytic sample comprised 6,700 students¹: 690 in below-grade-level classes, 4,750 in grade-level classes, and 1,260 in above-grade-level/honors classes. The levels of the classes were taken from the completed teacher questionnaires. Teachers were asked “Which of the following best describes this student’s English class?” and were given the choices “Instruction for students performing below grade level in reading,” “Regular,” and “Honors, Enrichment, or Gifted & Talented.” (Westat, 2007, p. 9).² The students in this sample came from 2,130 different schools (1,770 public and 360 private), with 3,230 teachers.

Outcome Variables

Collectively, the outcome measures studied here provide a general outline of instruction in eighth-grade classrooms, as well as more specific instructional

practices, including an especially precise student-level measure, the texts assigned to the students to read, which are coded for complexity.

Class Time Allocation

To get an overall sense of how instruction was allotted to middle school students, teachers were asked to indicate what percentage of class time they allocated to four areas: reading skills and strategies, appreciation and analysis of literature, student writing, and other topics. We created a continuous variable, representing the percentage of class time allocated to each area.

Instructional Practices

In addition to the overall big picture of what was going on in their classroom, teachers also provided information about 12 specific instructional practices that they used in the classroom (e.g., students revise a report or paper that they wrote). Here, teachers indicated how frequently they used each practice: never, once or twice a month, once or twice a week, or every day. These 12 instructional practices covered a wide range of literacy components, and the Appendix presents the complete list of instructional practices.

The guiding rationale for our measurement approach was that these measures of instruction are specific and noninterchangeable. That is, although they may share similar instructional properties (e.g., doing a read-aloud and watching a movie are both relatively passive activities compared with revising a report), each instructional activity represents specific, sometimes mutually exclusive practices. From a time use perspective, doing a read-aloud may mean that instructional activity is chosen instead of watching a movie. Prior studies involving ECLS-K data used exploratory factor analysis to identify and capture the underlying instructional constructs measured by the survey items at the earlier grade levels (Guarino, Hamilton, Lockwood, & Rathbun, 2006; Morgan, Farkas, & Maczuga, 2015). However, such an exploratory factor analysis was not appropriate in our analysis. These instructional measures generally have high item uniqueness when the covariance structure is examined (e.g., .85 for watching a movie). The item uniqueness, combined with the clear specificity of many of these time use items, led us to investigate the majority of instructional practices as individual dependent variables rather than identifying latent factors and allowing for measurement error.

However, five of the instructional practice measures pertain specifically to reading comprehension practices: students talking with one another about something that they read, writing about something that they read, making generalizations and drawing inferences

based on what they read, describing the style or structure of the text that they read, and making predictions about what they read. As with the other items, the reading items represent specific, sometimes mutually exclusive practices. As such, these items were again not good candidates for a modeling approach that identifies a reflective latent variable. We were, however, interested in analyzing collective differences in comprehension instruction across tracks. Thus, we used Stata's factor command with the principal-component factor option to create a formative principal component r score from a single-factor measurement model where the communalities were set to 1 (for a discussion of this model, see Fabrigar, Wegener, MacCallum, & Strahan, 1999; Velicer & Jackson, 1990). Factor loadings for items in comprehension instruction ranged from 0.687 to 0.886 with a median of 0.764, with low uniqueness values (ranging from .216 to .528 with a median of 0.437). The factor loadings in our principal component analysis compare favorably with results reported for factor scores in other *Reading Research Quarterly* articles (see, e.g., Kim & Phillips, 2014; Puranik & Lonigan, 2014; Reed, Petscher, & Truckenmiller, 2016), although these models entail different assumptions. Comprehension instruction has a Cronbach's alpha of .85.

Tables 1 and 2 report the adjusted differences across tracks for the seven statistically significant instructional practices, including reading comprehension instruction. Additionally, we report full regression models for two practices identified in the literature as varying by track level: "work in a reading workbook or on a worksheet" and "have assigned homework" (Gamoran, 1993; Gritter, 2012; Oakes, 2005). The coefficients for low- and high-track variables for all 12 instructional practices, including the individual items that make up the reading comprehension factor score, can be found in the Appendix.

Text Assignment and Complexity

Finally, teachers recorded the titles and authors of the last three books the student read for an assignment. For those students with track-level information, teachers recorded 16,390 fields of title text, but these were delivered by NCES precisely as written, and extensive cleaning was required. Typical errors included misspellings and abbreviations; added spaces, hyphens, or other punctuation; missing or incorrect articles (e.g., *the, an*); inclusion of the author in the title field (there was a separate field for the author); and use of informal titles (e.g., "Diary of Anne Frank"). In all, 3,890 fields were replaced to prepare data for merging with Lexile codes.³ Three hundred sixty fields were coded as various forms of "student choice," and another 30 could not be coded at all (e.g., the respondent wrote "an award

winning book"). Cleaning yielded 1,480 unique titles or text categories (e.g., student choice of a historical fiction book). We linked each title, when possible, to its Lexile level. Of the 16,390 fields, 14,230 (89%) could be assigned a Lexile. For each student, we created a mean Lexile level based on the last three books assigned by the teacher. In all cases, a Lexile level could be assigned for at least one of the three books for each student in the analytic sample of 6,700 students.

Lexile levels, a quantitative estimate of how easy or challenging a book is, are generated by an algorithm measuring syntactic and semantic features of the text, coupled with an aggregate measure of reading comprehension on the text, and rank texts on an interval scale (Cunningham & Mesmer, 2014). For example, *Green Eggs and Ham* by Dr. Seuss has a Lexile level of 30, whereas *Animal Farm* by George Orwell has a Lexile level of 1370. Although there are a variety of approaches to identify the complexity levels of texts quantitatively, we focused on the Lexile framework for this paper for several reasons. First, the Common Core uses Lexile levels to set reading levels for each grade band (NGA Center & CCSSO, 2010b). Additionally, Lexiles are not tied to a specific commercial reading program, such as Accelerated Reader, Fountas & Pinnell Literacy, and Developmental Reading Assessment, each of which has its own unique leveling system and is used in conjunction with its core reading program materials. Finally, Lexile levels are widely available; for example, many books print the Lexile level on the back cover, and many computerized assessments, such as the Measures of Academic Progress, provide recommended student Lexile levels. Although there are limitations to using Lexiles to measure text complexity, the limitations are similar to those of other quantitative measures of text complexity.⁴

Predictor Variables

Covariates included student, teacher, class, and school variables. Student variables were gender (male or female), race/ethnicity (white, black, Hispanic, Asian, or other), age, household socioeconomic status (a continuous variable using the ELCS-K-created scale that combined household income, occupation, and parental education), a dichotomous variable for whether or not the student has a learning disability (the reference group is no learning disability), a dichotomous variable for whether or not the student has a reading tutor (the reference group is no reading tutor), a dichotomous variable for student educational expectations (which captured whether or not the student thought he or she would earn a bachelor's degree; the reference group is earning less than a bachelor's degree), and an identical variable capturing parental educational expectations for their child

TABLE 1
Descriptive Statistics for the Analytic Sample

Variable	Low-track <i>M</i> (<i>SD</i>)	Grade-level track <i>M</i> (<i>SD</i>)	High-track <i>M</i> (<i>SD</i>)	F-test comparing means
<i>Student variables</i>				
Male	.62 (.49)	.52 (.50)	.42 (.50)	***
White	.47 (.50)	.61 (.47)	.66 (.47)	***
Black	.23 (.38)	.17 (.30)	.09 (.24)	***
Hispanic	.25 (.44)	.16 (.35)	.17 (.36)	***
Asian	.02 (.19)	.03 (.21)	.06 (.30)	***
Other race	.05 (.25)	.04 (.22)	.03 (.17)	**
Age	14.39 (0.55)	14.29 (0.44)	14.29 (0.42)	***
Socioeconomic status	-0.55 (0.70)	-0.05 (0.78)	0.37 (0.76)	***
Grade 5 reading score	122 (26.28)	151 (23.26)	171 (17.90)	***
Grade 8 reading score	140 (29.65)	169 (24.74)	189 (16.13)	***
Learning disability	.23 (.42)	.10 (.28)	.04 (.16)	***
Reading tutor	.31 (.46)	.11 (.30)	.03 (.18)	***
Educational expectations	.63 (.47)	.84 (.33)	.96 (.19)	***
Parental educational expectations	.54 (.50)	.74 (.41)	.93 (.26)	***
Teacher-rated ability	2.26 (0.88)	3.00 (0.94)	3.86 (0.81)	***
<i>Teacher variables</i>				
Age	42.90 (11.75)	42.06 (11.95)	43.88 (12.22)	***
White	.84 (.33)	.90 (.28)	.91 (.28)	**
Highest degree	.54 (.50)	.51 (.50)	.53 (.50)	**
Years taught	13.31 (9.35)	13.14 (9.93)	15.00 (10.17)	***
Certification	.82 (.39)	.84 (.38)	.88 (.32)	***
<i>Class and school variables</i>				
Class size	18.84 (7.49)	23.95 (6.04)	25.35 (6.22)	***
Percentage of the class a minority	.49 (.33)	.33 (.27)	.30 (.25)	***
Class time (hours/week)	6.00 (1.97)	5.55 (1.83)	5.41 (1.79)	***
Public school	.97 (.15)	.88 (.41)	.91 (.35)	***
School socioeconomic status	-0.40 (0.56)	-0.04 (0.61)	0.22 (0.62)	***

(continued)

(the reference group is earning less than a bachelor's degree). Students' prior achievement was controlled in two ways, with fifth-grade reading scores and an eighth-grade teacher rating of academic ability. Fifth grade was the most recent year prior to eighth grade that reading scores from the ECLS-K end-of-year assessment were collected, and preliminary models showed that these scores are associated with the allocation of instructional time, instructional practices, and text assignment.

Reading achievement in the ECLS-K was measured using item response theory and was recalibrated after every round of assessment to allow placement on a continuous achievement scale (Tourangeau, Nord, Lê, Sorongon, & Najarian, 2009). For the teacher rating of academic ability, eighth-grade English teachers rated students in three areas: writing achievement, oral achievement, and behavior. Writing achievement included tasks such as organizing ideas, gathering information, writing

TABLE 1
Descriptive Statistics for the Analytic Sample (continued)

Variable	Low-track <i>M</i> (<i>SD</i>)	Grade-level track <i>M</i> (<i>SD</i>)	High-track <i>M</i> (<i>SD</i>)	F-test comparing means
<i>Instructional variables</i>				
Percentage of class time on skills and strategy	39.27 (15.67)	28.17 (13.36)	23.83 (12.66)	***
Percentage of class time on literature analysis	21.80 (11.30)	28.59 (11.62)	32.41 (12.34)	***
Percentage of class time on writing	30.73 (12.70)	31.85 (11.92)	32.77 (11.71)	*
Percentage of class time on other tasks	8.20 (10.03)	11.37 (11.49)	10.99 (10.28)	***
Workbooks or worksheets	2.78 (0.95)	2.44 (0.96)	2.26 (0.97)	***
Comprehension instruction	3.04 (0.55)	3.08 (0.54)	3.21 (0.53)	***
Homework	3.02 (1.00)	3.35 (0.81)	3.51 (0.66)	***
Group activity or project	2.21 (0.70)	2.19 (0.70)	2.28 (0.64)	***
Watch video	1.82 (0.71)	1.76 (0.63)	1.73 (0.63)	**
Read aloud	3.26 (0.82)	2.98 (0.88)	2.88 (0.87)	***
Lexile level of assigned books	813 (146)	843 (128)	892 (138)	***

Note. Means are for the weighted sample, and standard deviations are for the unweighted sample. All variables in the table are binary except for age, socioeconomic status, grades 5 and 8 reading scores, teacher-rated ability, teacher age, years taught, class size, class time, and school socioeconomic status.

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

TABLE 2
Difference in Instruction by Track Level Relative to the Regular (Grade-Level) Track

Variable	Low track		High track	
	Raw difference	Adjusted difference	Raw difference	Adjusted difference
Percentage of class time on skills and strategy	11	7***	-4	-2**
Percentage of class time on literature analysis	-7	-5***	4	2**
Percentage of class time on writing	-1	-1	1	1
Percentage of class time on other tasks	-3	-2*	0	-1
Workbooks or worksheets	0.39	0.28***	-0.19	-0.09
Comprehension instruction	-0.04	-0.09*	0.13	0.13***
Homework	-0.24	-0.07	0.20	0.11*
Group activity or project	0.03	0.01	0.20	0.17***
Watch video	0.12	0.12*	0.02	-0.01
Read aloud	0.30	0.16*	-0.12	-0.06
Lexile level of assigned books	-30	-15	49	36***

Note. Adjusted differences are produced with regression models. Control variables include student, teacher, classroom, and school characteristics (see Table 3 for full regression models).

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

in a variety of genres, using the stylistic and rhetorical aspects of writing, and using grammar. Oral achievement included analytic thinking, creative thinking, and use of grammar. Behavior included whether or not the student was attentive or disruptive in class, as well as homework completion, and was included because paying attention in class and completing homework are important components of academic success (Bodovski & Youn, 2011; Jacob & Parkinson, 2015; Li-Grining, Votruba-Drzal, Maldonado-Carreño, & Haas, 2010). These three components were combined into an aggregate measure of overall academic achievement (Cronbach's $\alpha = .92$).

Teacher variables were age, race/ethnicity (a dichotomous variable indicating whether the teacher was white; the reference group is white), a dichotomous variable indicating whether or not the teacher had a master's degree (the reference group is no master's degree), the number of years taught, and a dichotomous variable indicating whether or not the teacher had regular certification (the reference group is no or nonregular certification).

Class variables were class size, the percentage of minority students in the class, and class time, measured in hours per week. School variables consisted of whether the school was public or private and school-level socioeconomic status, which was aggregated from student-level socioeconomic status.

Data Analyses

We present results by first reporting descriptive statistics for instructional time allocation, instructional practices, and text complexity across track levels. Second, we used linear regression models with robust variance estimators (Stata's svy commands), which adjust for the clustering of students in schools, to investigate whether track was predictive of these variables. In addition to the correction for clustering (using the primary sampling unit identifiers), weights and Stata identifiers were applied to control for the survey design (see Tourangeau et al., 2009). We conducted regression at the student level, which is reflective of how many students experienced these classroom conditions. The student-level model is also consistent with the measurement structure of the text complexity measure and instructional practices; assigned texts and instructional practices were reported for individual students.

In preliminary analyses, we ran school fixed-effects models to isolate within-school variation in tracking and text complexity. The point estimates are similar for the high-track effect and about half the magnitude for the low-track effect. However, because of the small number of students in each school ($n = 3.3$), the within-school coefficients are estimated with poor precision. Additionally, only 28% of the schools had students in two or more

distinct class periods, limiting the possibility for distinguishing between classroom- and school-level variance parameters. Relatedly, because of these limitations, we did not utilize class or school means-as-outcomes models (Raudenbush & Bryk, 2002), which assume sufficient within-group sample sizes for reliable aggregation (Snijders & Bosker, 1999). Thus, we focused on the results of the ordinary least squares regression models here.

We used multiple imputation, using the chained equation approach and creating five sets of imputations, to handle missing data for independent variables (Acock, 2016; Raghunathan, Lepkowski, Van Hoewyk, & Solenberger, 2001). All independent variables except gender had missing values; however, 14 of the 24 variables had less than 1% missing data. The variables with the highest percentage of missing data were student expectations of educational outcomes (15%), whether or not the student had a learning disability (11%), and household socioeconomic status (9%). Student demographic variables are often imputed in secondary analyses of ECLS-K and NCES data (e.g., Morgan et al., 2015), as these variables cannot be assumed to be missing at random. Outcome variables were not imputed.

Finally, in addition to our main multivariate regression results, we also conducted a segregation analysis to describe the overall multinomial dispersion in texts attributable to track level. The segregation analysis relied on a separate, collapsed data file of texts rather than students, where counts of track identities were stored for each text (e.g., approximately 30 low-track, 220 regular-track, and 70 high-track students read *The Call of the Wild* by Jack London). Then, we summarized the dispersion of student groups across texts, treating each text as a nesting unit. Substantively, this captured whether or not different groups of students were assigned similar or dissimilar books to read. We report two measures of segregation: the information theory segregation index (h) and the index of dissimilarity (d). The information theory index is computed as 1 minus the ratio of the average diversity of individual/specific texts to the total set/population of texts in the database. If all texts have the same track-level composition as the population, segregation will be minimal, and the information theory index (h) will be 0. If many texts have substantial overrepresentations of students from a given track level, then h will be large, approaching 1. The information theory index is now a commonly used measure of segregation, but we also report the index of dissimilarity (d), a more traditional measure of segregation (developed in the 1940s to study residential segregation; see Duncan & Duncan, 1955), which also ranges from 0 to 1 and is highly correlated with h (for a discussion of these measures, see Reardon & Yun, 2003).

Results

Table 1 presents descriptive statistics for the analytic sample. Additionally, Table 1 presents an analysis of variance of group means, suggesting that for all variables, there are statistically significant differences between track levels. Achievement scores from grades 5 and 8 suggest that tracking is often used to organize instruction for struggling readers. Although teachers were not asked directly if the student participating in the ECLS-K project was a struggling reader, students placed into the below-grade-level track in eighth grade had fifth-grade end-of-year achievement scores in reading (the closest school year to the start of eighth grade measured in the study) that were approximately 30 points below the grade-level track and nearly 50 points below the above-grade-level track (122 vs. 151 vs. 171). Thus, it seems likely that the eighth-grade track placement largely reflected reading achievement from prior grades and that the students placed in the low-track classes were often struggling readers.

Class Time Allocation

Table 1 also shows descriptive statistics for instructional time use. Teachers of low-track classes reported allocating more class time to skills and strategy instruction (approximately 39% of class time) than did teachers of grade-level (28%) and high-track classes (24%). When compared with grade-level classes, these changes in instructional time allocation are a medium to large effect size for the low-track class (Cohen's $d = 0.76$) and a small effect size for the high-track class (Cohen's $d = 0.33$). Likewise, teachers of low-track classes reported allocating less class time to literature instruction (approximately 22% of class time) than did teachers of grade-level (29%) and high-track classes (32%), a medium effect size for the low-track class (Cohen's $d = 0.59$) and a small effect size for the high-track class (Cohen's $d = 0.32$).

Table 2 compares the difference in means by track levels reported in Table 1, along with the adjusted differences (regression coefficients) from the regression models presented in Table 3. This allows ease of comparison between the differences in means in the raw data and the adjusted differences in means after controlling for student, teacher, and school characteristics. Instructional practices denoted with an asterisk(s) in Table 2 indicate that the adjusted difference is statistically significant. Controlling for student achievement and other student, teacher, and school characteristics, we found both an increase in time allocated to skills and strategy instruction in the low-track classes ($\beta = 7.11, p < .001$) and a decrease in time allocated to literature instruction ($\beta = -4.56, p < .001$), as compared

with the grade-level track. Students in low-track classes got an estimated 7% more time allocated to skills and strategy instruction and 5% less time allocated to literature instruction. Likewise, high-track students spent less time on skills and strategy ($\beta = -2.05, p < .01$) and more time on literature instruction ($\beta = 2.04, p < .01$), as compared with the grade-level track, although the adjusted differences are small. There were small differences in the percentages of class time allocated to writing across track level (ranging from about 31% to 33% of class time; see Table 1), but these did not persist in the regression models.

Instructional Practices

Similar to instructional time allocation, teachers indicated using different instructional practices more or less frequently across track levels. Following the same pattern as instructional class time allocation, descriptive statistics in Table 1 show that teachers in low-track classes more frequently used workbooks or worksheets than did teachers of grade-level or high-track classes, a small effect size for both low-track classes (Cohen's $d = 0.41$) and high-track classes (Cohen's $d = 0.20$), as compared with grade-level classes. Teachers in high-track classes more frequently used comprehension-based instruction than did teachers at the other track levels, a small effect size for high-track classes (Cohen's $d = 0.24$) and no effect size for low-track classes (Cohen's $d = 0.07$). Additionally, teachers less frequently assigned homework to students in low-track classes than did teachers at the other track levels, a small effect size for both low-track (Cohen's $d = 0.27$) and high-track classes (Cohen's $d = 0.27$).

Regression models show that track level is associated with an increase in using workbooks or worksheets at the low-track level ($\beta = 0.28, p < .001$), controlling for achievement and other background variables. Track level is also associated with both a decrease in use of comprehension-based instruction at the low-track level ($\beta = -0.09, p < .05$) and an increase in use at the high-track level ($\beta = 0.13, p < .001$). Track level is associated with an increase in the frequency of homework for students in the high-track classes ($\beta = 0.11, p < .05$).

Additionally, in exploring the individual instructional practices, we found three other instructional practices that varied by track level significantly (see Table 2 for adjusted differences and the Appendix for regression coefficients). Students in high-track classes spent more time doing group activities and projects ($\beta = 0.17, p < .001$), whereas students in low-track classes spent more time watching videos ($\beta = 0.12, p < .05$) and reading aloud ($\beta = 0.16, p < .05$), as compared with students in grade-level classes. Overall, the adjusted differences in Table 2 show that track level contributed to the

TABLE 3
Regression Models for Time Allocation, Instructional Practices, and Lexile Levels

Variable	Skills and strategy time allocation	Literature analysis time allocation	Writing time allocation	Other tasks time allocation	Workbooks or worksheets	Comprehension instruction	Homework	Lexile level
<i>Student variables</i>								
Low track	7.11*** (1.38)	-4.56*** (0.86)	-0.96 (0.88)	-1.58* (0.73)	0.28*** (0.07)	-0.09* (0.04)	-0.07 (0.06)	-14.88 (10.63)
High track	-2.05** (0.76)	2.04** (0.79)	0.86 (0.76)	-0.83 (0.62)	-0.09 (0.06)	0.13*** (0.03)	0.11* (0.04)	36.43*** (8.42)
Male	-0.25 (0.49)	-0.42 (0.43)	0.18 (0.55)	0.49 (0.42)	0.01 (0.04)	-0.03 (0.02)	0.00 (0.03)	1.93 (4.72)
Black	0.26 (1.30)	0.23 (1.18)	-0.81 (1.11)	0.29 (0.74)	0.05 (0.07)	0.01 (0.05)	0.06 (0.07)	-14.84 (9.95)
Hispanic	-0.39 (0.76)	0.75 (0.72)	-0.40 (0.63)	-0.01 (0.66)	-0.06 (0.07)	0.01 (0.03)	0.07 (0.06)	-5.89 (8.77)
Asian	-1.53 (0.93)	1.04 (0.96)	0.75 (1.09)	-0.28 (0.74)	-0.08 (0.09)	0.01 (0.04)	0.13* (0.06)	-17.65 (10.37)
Other race	0.42 (1.99)	-0.65 (1.17)	0.82 (1.51)	-0.60 (0.99)	0.00 (0.07)	-0.02 (0.06)	-0.07 (0.11)	9.98 (14.79)
Age	0.40 (0.59)	0.25 (0.53)	-0.38 (0.61)	-0.27 (0.47)	-0.01 (0.05)	0.03 (0.02)	-0.13** (0.04)	-5.69 (5.98)
Grade 5 reading score	-0.01 (0.02)	0.00 (0.01)	0.00 (0.01)	0.01 (0.01)	-0.01** (0.00)	0.00 (0.00)	0.00 (0.00)	0.29 (0.17)
Learning disability	1.63 (1.05)	-0.27 (1.09)	-0.71 (0.93)	-0.65 (0.74)	0.07 (0.06)	-0.05 (0.04)	0.06 (0.06)	15.22 (11.89)
Reading tutor	0.39 (1.03)	-0.29 (0.88)	0.35 (0.90)	-0.44 (0.70)	-0.09 (0.05)	0.09* (0.03)	0.03 (0.06)	-21.66* (10.21)
Socioeconomic status	-0.17 (0.60)	0.06 (0.47)	0.22 (0.43)	-0.11 (0.46)	-0.03 (0.03)	0.03 (0.02)	0.01 (0.04)	2.31 (5.54)
Educational expectations	0.28 (1.07)	0.53 (0.71)	-1.21 (0.96)	0.39 (0.64)	0.02 (0.07)	0.03 (0.03)	0.12* (0.06)	-3.07 (10.04)
Parent educational expectations	-0.58 (0.79)	1.38* (0.69)	-0.60 (0.82)	-0.20 (0.67)	0.01 (0.05)	0.01 (0.03)	0.08 (0.05)	-0.25 (7.65)
Teacher-rated ability	-0.54 (0.37)	0.85** (0.33)	-0.23 (0.38)	-0.08 (0.28)	0.02 (0.03)	0.03 (0.02)	0.01 (0.02)	1.04 (3.64)

(continued)

TABLE 3
Regression Models for Time Allocation, Instructional Practices, and Lexile Levels (continued)

Variable	Skills and strategy time allocation	Literature analysis time allocation	Writing time allocation	Other tasks time allocation	Workbooks or worksheets	Comprehension instruction	Homework	Lexile level
<i>Teacher variables</i>								
Age	-0.08 (0.05)	0.04 (0.06)	0.00 (0.05)	0.05 (0.03)	0.00 (0.00)	0.00* (0.00)	0.00 (0.00)	0.51 (0.55)
White	-2.58* (1.10)	3.75*** (1.03)	-0.27 (1.19)	-0.88 (0.80)	-0.29*** (0.08)	0.09 (0.06)	-0.14* (0.06)	-3.70 (14.18)
Highest degree	-2.17* (0.80)	1.40 (0.66)*	0.91 (0.59)	-0.14 (0.52)	-0.13** (0.05)	0.04 (0.03)	0.05 (0.04)	1.01 (6.88)
Years taught	0.02 (0.06)	0.03 (0.06)	-0.06 (0.05)	0.02 (0.04)	0.00 (0.00)	0.00 (0.00)	0.01** (0.00)	0.11 (0.81)
Certification	0.94 (0.90)	-1.14 (0.91)	0.38 (0.74)	-0.18 (0.72)	0.12 (0.07)	0.00 (0.04)	-0.01 (0.07)	0.97 (10.10)
Class size	-0.05 (0.06)	0.03 (0.05)	-0.03 (0.06)	0.05 (0.04)	0.00 (0.00)	0.00 (0.00)	0.01* (0.00)	0.84 (0.61)
Percentage of the class a minority	5.66*** (1.65)	-1.14 (1.39)	-1.77 (1.23)	-2.67** (1.07)	0.02 (0.11)	0.17** (0.06)	0.09 (0.10)	37.52 (14.38)
Class time (hours/week)	0.65** (0.21)	-0.18 (0.17)	-0.54* (0.25)	0.07 (0.15)	0.04* (0.02)	0.05*** (0.01)	0.02 (0.01)	2.36** (2.02)
Public school	3.28*** (0.96)	-0.05 (1.01)	2.52* (1.00)	-5.76*** (1.04)	-0.11 (0.09)	0.03 (0.05)	-0.38*** (0.06)	-16.16 (13.58)
School socioeconomic status	-3.06*** (0.68)	1.70** (0.62)	1.17 (0.61)	0.19 (0.54)	-0.04 (0.06)	-0.03 (0.03)	0.09 (0.05)	12.39 (7.07)
Constant	24.25 (8.54)	17.13 (8.70)	42.49 (9.39)	16.05 (7.62)	3.24 (0.76)	2.41 (0.34)	4.77 (0.63)	833.02 (95.62)

Note. Standard errors are in parentheses.

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

frequency of instructional practices even after controlling for covariates; however, the differences are quite small (a standard deviation difference of 0.10–0.20 between the low and high tracks).

Text Assignment and Complexity

How did the assignment of texts differ across track levels in eighth grade? Table 4 shows the 20 most frequently assigned books in each of the three track levels and the percentage of students at each track level assigned that book. Recall that more than 1,400 different titles were provided by ECLS-K teachers; eighth-grade teachers selected texts from an extensive set of possible titles. It is important to note that not all teachers were allowed to select the books read in their classrooms; many teachers were implementing curriculum that was set at the school or district level. However, it is important to document differences in text titles and complexity, regardless of whether the decision was at the teacher, school, or district level. The top 20 books in Table 4 constitute 37% of assignments for grade-level students and 28.4% and 43.9% for low- and high-track students, respectively.

Turning to differences by track, first, there were similarities in text assignment across all three track levels. Eight of the titles were frequently assigned, regardless of track level: *The Outsiders* by S.E. Hinton, *The Giver* by Lois Lowry, *The Diary of a Young Girl* by Anne Frank, *Roll of Thunder, Hear My Cry* by Mildred D. Taylor, *The Call of the Wild* by Jack London, *A Christmas Carol* by Charles Dickens, *Night* by Elie Wiesel, and *Flowers for Algernon* by Daniel Keyes. These texts represent a broad selection of classics (what we might term *new classics*) and Newbery Medal winners. Finally, although student choice was only a small part of the overall sample (approximately 2%), student choice was used in all three track levels.

One aspect that is notable about the books assigned to eighth graders across all tracks is the extent to which they continue to represent the typical canon. In a separate paper, we examined the sociodemographic diversity of authors and protagonists in these texts and found that this group of texts predominantly feature male protagonists (70%) and white protagonists (80%), with white males accounting for 52% of all protagonists (Northrop, Borsheim-Black, & Kelly, 2018). Although we found some evidence of text-student matching (e.g., a black student assigned a book with a black protagonist), we found the diversity in the texts to be lacking overall.

However, examining the remaining titles showed that there were more nuanced differences in text assignment across track levels. Students in low-track classes tended to read more popular fiction, such as

Holes by Louis Sachar, *Tears of a Tiger* by Sharon M. Draper, and *Dragonwings* by Laurence Yep, whereas students in the high-track classes read more classics, such as a variety of plays by William Shakespeare, *Animal Farm* by George Orwell, and *Lord of the Flies* by William Golding. Even when students read the same author, such as John Steinbeck, there were differences by track level, with students in the low track reading *The Pearl*, and students in the high track reading both *The Pearl* and *Of Mice and Men*.

Examining measures of segregation to summarize the overall multinomial dispersion in texts attributable to track differences, we found a value of .21 for the information theory index (h) and a value of .35 for the index of dissimilarity. This revealed what might be summarized as moderate segregation of texts by track level (for guidelines for assessing strength of segregation statistics, see Reardon & Yun, 2003). This middle range value summarizes the segregation across the 1,400+ texts, where some texts were much more commonly read in a given track and others common across tracks. For example, *Animal Farm* was read by 240 regular- and high-track students but fewer than 10 low-track students. In contrast, *Dragonwings* was read more evenly, by about 20 low-, 20 regular-, and 10 high-track students. Of course, *moderate* is still a qualitative label applied to a quantitative measure and only so descriptive, but these quantitative measures provide a benchmark for future research.

Corresponding with the differences in text assignment, we found important differences in text complexity across tracks. Table 1 shows that the average Lexile level for the low-track class is 813, compared with 843 for the grade-level track and 892 for the high track. When compared with the grade-level classes, there is a small effect size in the Lexile difference score for both the low-track class (Cohen's $d = 0.22$) and the high-track class (Cohen's $d = 0.37$). To provide a benchmark for understanding these differences, we calculated the 95%-trimmed max-min Lexile range for teachers who reported more than one book title, which yielded a trimmed mean range in Lexile scores of 151. Thus, the low-track/high-track gap of 79 points is equivalent to a little under half of the typical range of text complexity encountered within classes. Evaluated in a number of ways, then, there appears to be substantial overlap in the Lexiles read across different track levels, even as the means differ considerably. However, given that Lexile levels are not assigned to nonnarrative texts, such as Shakespeare's plays, and are missing for some of the older classics, such as collections of short stories by O. Henry and Edgar Allan Poe, it is likely that these are conservative estimates of differences in text complexity across tracks and that students in the high-track classes were actually reading texts at an even higher level than reported.

TABLE 4

The 20 Most Frequently Taught Books in Each Track Level, by Percentage of Students in the Track

Rank	Low track		Grade-level track		High track	
1	<i>The Outsiders</i> by S.E. Hinton	14%	<i>The Outsiders</i> by S.E. Hinton	18%	<i>To Kill a Mockingbird</i> by Harper Lee	14%
2	<i>The Giver</i> by Lois Lowry	8%	<i>The Giver</i> by Lois Lowry	14%	<i>The Diary of a Young Girl</i> by Anne Frank	12%
3	<i>The Diary of a Young Girl</i> by Anne Frank	7%	<i>The Diary of a Young Girl</i> by Anne Frank	9%	<i>The Giver</i> by Lois Lowry	12%
4	<i>Holes</i> by Louis Sachar	6%	<i>To Kill a Mockingbird</i> by Harper Lee	9%	<i>The Outsiders</i> by S.E. Hinton	12%
5	<i>Roll of Thunder, Hear My Cry</i> by Mildred D. Taylor	4%	<i>Roll of Thunder, Hear My Cry</i> by Mildred D. Taylor	6%	<i>Animal Farm</i> by George Orwell	10%
6	<i>The Call of the Wild</i> by Jack London	4%	<i>Night</i> by Elie Wiesel	5%	<i>Night</i> by Elie Wiesel	10%
7	<i>The Pigman</i> by Paul Zindel	3%	<i>The Pearl</i> by John Steinbeck	5%	<i>The Pearl</i> by John Steinbeck	9%
8	<i>Nothing but the Truth</i> by Avi	3%	<i>The Call of the Wild</i> by Jack London	5%	<i>Romeo and Juliet</i> by William Shakespeare	8%
9	<i>The Cay</i> by Theodore Taylor	3%	<i>Flowers for Algernon</i> by Daniel Keyes	5%	<i>A Midsummer Night's Dream</i> by William Shakespeare	6%
10	<i>A Christmas Carol</i> by Charles Dickens	3%	<i>The Pigman</i> by Paul Zindel	4%	<i>The Call of the Wild</i> by Jack London	6%
11	<i>The Watsons Go to Birmingham—1963</i> by Christopher Paul Curtis	3%	<i>The Diary of Anne Frank</i> (play) by Frances Goodrich and Albert Hackett	4%	<i>Fahrenheit 451</i> by Ray Bradbury	5%
12	<i>Night</i> by Elie Wiesel	3%	<i>A Christmas Carol</i> by Charles Dickens	4%	<i>Flowers for Algernon</i> by Daniel Keyes	4%
13	Student choice	3%	<i>Romeo and Juliet</i> by William Shakespeare	4%	<i>The Adventures of Tom Sawyer</i> by Mark Twain	4%
14	<i>The Pearl</i> by John Steinbeck	3%	<i>Across Five Aprils</i> by Irene Hunt	3%	<i>Lord of the Flies</i> by William Golding	3%
15	<i>Touching Spirit Bear</i> by Ben Mikaelson	3%	<i>The Adventures of Tom Sawyer</i> by Mark Twain	3%	<i>A Christmas Carol</i> by Charles Dickens	3%
16	<i>Bridge to Terabithia</i> by Katherine Paterson	3%	<i>My Brother Sam Is Dead</i> by James Lincoln Collier and Christopher Collier	3%	<i>Johnny Tremain</i> by Esther Forbes	3%
17	<i>Flowers for Algernon</i> by Daniel Keyes	3%	Student choice	3%	<i>Roll of Thunder, Hear My Cry</i> by Mildred D. Taylor	3%
18	<i>Tears of a Tiger</i> by Sharon M. Draper	3%	<i>Nothing but the Truth</i> by Avi	3%	<i>And Then There Were None</i> by Agatha Christie	3%
19	<i>The Diary of Anne Frank</i> (play) by Frances Goodrich and Albert Hackett	1%	<i>Johnny Tremain</i> by Esther Forbes	3%	<i>Of Mice and Men</i> by John Steinbeck	3%
20	<i>Dragonwings</i> by Laurence Yep	1%	<i>Animal Farm</i> by George Orwell	3%	<i>Across Five Aprils</i> by Irene Hunt	3%

Table 2 reproduces these raw differences, along with regression-based estimates of track effects on text complexity. After controlling for student achievement and other characteristics, the decrease in Lexile level for low-track students is a non-statistically significant 15

points, a small amount even if we could be confident in that point estimate. In contrast, the increase in Lexile level for high-track students is a substantial 36 points ($\beta = 36.43$, $p < .001$). As a benchmark for considering the effect size of this difference, we calculated the

95%-trimmed teacher-level standard deviation in mean Lexile scores, which was 120. Thus, high-track students encountered texts about one third of a standard deviation higher in complexity than regular-track students. This is consistent with a social categorization hypothesis, wherein teachers assign more difficult texts based on track labels, and beliefs concerning student achievement that exaggerate true differences across tracks; our analysis here suggested that teachers assigned more difficult texts to students in high-track classes, even after controlling for student ability.

Other Instructional Factors

In addition to our focus on class time allocation, instructional practices, and text complexity, our data show other important differences in the type of educational experiences that struggling readers had when placed in low-track classes. Students in low-track classes were more likely to be placed in smaller classes (19 students per class vs. 24 or 25 students for the grade-level and above-grade-level classes), a medium to large effect size (Cohen's $d = 0.75$). The low-track students spent slightly more time in class each week than the other tracks (six hours per week, compared with approximately 5.5 hours per week for the grade-level and above-grade-level classes), a medium effect size (Cohen's $d = 0.49$). Despite these structural advantages, low-track classes were less likely to be taught by a certified teacher (82% compared with 84% for grade-level classes and 88% for above-grade-level classes) with fewer years of experience (13 years compared with 15 years for the above-grade-level track), although there were no statistically significant differences for either certification or years of experience.

Additionally, according to the results of the parent survey, students placed in low-track classes were more likely to have a reading tutor (31%, compared with 11% of students in grade-level classes and 3% of students in above-grade-level classes). However, the survey did not ask parents to indicate when or where the tutoring took place. Thus, we do not know whether the tutor was provided by the school during school hours or was engaged by the parent for additional instruction beyond the regular school day.

Table 1 also shows that the average rate of growth between the end of fifth grade and the end of eighth grade for students across all three tracks was the same, 18 points. However, the possibility of intrinsic associations between initial status and growth, including test scaling, makes it difficult to know whether the 18-point gain for low-track students was qualitatively the same as that of high-track students (Kelly & Ye, 2017; Najarian, Pollack, & Sorongon, 2009). A thorough exploration of how tracking and instructional practices are related to

the growth of reading skills in middle school is beyond the scope of this study, as the ECLS-K did not collect data when the students were in sixth or seventh grade. Yet, these basic results suggest that overall track effects on learning were not strongly negative and that tracking likely benefited students in some ways. For example, teachers may have adjusted instruction to meet the individual needs of the students in their class in ways both revealed and not revealed in the data.

Discussion

Data from the eighth-grade ECLS-K reveal a complex picture of the classroom instruction and classroom materials that struggling readers are exposed to in middle school. In this study, we found that students in low-track classes spent more class time on skills and strategy instruction and less class time on literature analysis, more frequently completed worksheets or workbooks, were less frequently assigned homework, and read less challenging texts than their counterparts in grade-level and high-track classrooms.

Class Time Allocation and Instructional Practices

Although we could not determine whether the instructional practices implemented by eighth-grade teachers matched the instructional needs of the individual students in these data, we analyzed differences in instruction adjusting for detailed measures of student achievement and mapped the findings from the ECLS-K data to what is known about effective instruction in general for middle school readers. Middle school struggling readers need access to instruction that includes explicit strategy instruction (Allington, 2011; Edmonds et al., 2009; Faggella-Luby & Deshler, 2008; Risko & Walker-Dalhouse, 2015), and at first glance, the ECLS-K data appear to support the idea that struggling readers were receiving this type of instruction: Teachers in low-track classes spent more time on reading skills and strategy instruction than did teachers of grade-level and above-grade-level classes. However, when coupled with the data on frequency of instructional practices, the picture became murkier. Teachers of low-track classes more frequently used workbooks or worksheets than would have been anticipated based on student achievement levels alone, and reported less frequent use of comprehension-based practices, such as making predictions, making generalizations, and describing the style or structure of what was read. However, it is precisely these comprehension strategies—making predictions, drawing inferences, and so forth—that some research has recommended for all students, including struggling readers (Harvey & Goudvis, 2007; Kamil et al., 2008).

Middle school struggling readers need ample time in school to read and discuss what they are reading (Allington, 2013; Risko & Walker-Dalhouse, 2015). Here, the data paint a discouraging picture about the opportunity for struggling readers to engage in literature-based discussion: Students in low-track classes spent less time on literature analysis than did their counterparts in grade-level or above-grade-level classes. Additionally, deconstructing the comprehension instructional scale showed that teachers reported using the instructional techniques of talking or writing about what was read less frequently in low-track classrooms.

Although our study provided descriptive statistics about the class time allocation and the types of instructional practices that eighth-grade teachers implemented, without more detailed, qualitative information about the specific worksheets, strategies, and instructional practices used in the classroom, it is difficult to determine whether these instructional choices were appropriate for the students in these classrooms. It is possible that the differences in instructional choices among the tracks reflected the needs of the students, such as explicit instruction in strategy knowledge (Faggella-Luby & Wardwell, 2011; Wilson, Faggella-Luby, & Wei, 2013). The fact that students across all three track levels gained the same amount of points on the ECLS-K assessment between the end of fifth grade and the end of eighth grade suggests that teachers were matching instructional practices appropriately overall. However, if teachers were using less frequent discussion of literature and more skill-and-drill instruction, that could be reflective of the low expectations typical of low-track classes (Lewis & Diamond, 2015), thus depriving students in low-track courses of valuable opportunities to learn.

Text Assignment and Text Complexity

In addition to time throughout the school devoted to reading, Allington (2007, 2011) argued that middle school struggling readers need access to materials at their reading level that they can understand. Here, the data support the idea that struggling readers were, at least, being given texts at a slightly lower Lexile level than those in grade-level and above-grade-level classes, although the difference in mean Lexile level is small and may not be meaningful in terms of the practical application of matching students to books at their reading level. Additionally, we were somewhat surprised to find such substantial overlap in the course texts read by students across track levels. This finding is in direct contrast to prior research suggesting that students in low-track classes did not have access to the same content knowledge as students in high-track classes, with

high-track students receiving a privileged, culturally elite curriculum (Oakes, 2005). But in these data, in many cases, students in different track levels were reading the same books, such as *The Outsiders*, *The Diary of a Young Girl*, and *Roll of Thunder, Hear My Cry*. Although there were differences, it does not seem that low-track students were categorically denied access to what might be considered a typical college-preparatory curriculum, although sometimes they may have been assigned texts at levels they may have had difficulty in reading independently.

Currently, the Common Core is inadvertently compounding confusion about what texts are appropriate for struggling readers. Specifically, the Common Core has significantly increased the level of text complexity required at each grade level (NGA Center & CCSSO, 2010b), but the standards' authors may not have recognized what a substantial departure from current practice that standard represents. The average Lexile level of the texts read by students in this study was 850, below the Lexile range of 955 to 1155 recommended by the Common Core. Furthermore, this Lexile range is for students in grades 6–8, suggesting that by eighth grade, students should be reading much closer to the end target than the beginning. Although this datum was collected in 2006–2007, and thus before the implementation of the Common Core, it shows that teachers will have to significantly increase the level of text complexity across all track levels if students are to meet the end-of-year standard for text complexity, which may be a challenging task. Examination of historical growth curves suggested that only students in the top quartile will achieve the Lexile levels recommended by the Common Core (Williamson, Fitzgerald, & Stenner, 2014), thus exacerbating the problem and placing more burden on teachers to try to raise rigor across all levels, including for students in low-track classrooms. Although it is important that struggling readers have access to texts at both their independent reading levels and at grade level, it is possible that since the implementation of the Common Core, struggling readers now have even less access to texts with lower text complexity than they did when this study was conducted.

Finally, adjusted (regression-based) differences showed that with respect to text complexity, social categorization effects primarily led teachers to adjust text complexity relative to high-track students rather than generating differences between low- and regular-track students. Note that in this analysis, we could not differentiate between high-track teachers adjusting complexity upward or low- and regular-track teachers adjusting complexity downward, or some combination of the two. Rather, the findings are broadly indicative of a nontrivial amount of overcompensation for variability in student achievement. This overcompensation is associated

with students in some low- and regular-track classrooms being given instruction and texts that are not as rigorous as the students could likely be successful with. Indeed, this social categorization effect is the opposite of, or works against, the genuine, effective practices of individualization of learning opportunities envisioned by the Association for Middle Level Education (2010).

Definitively characterizing appropriate instruction for struggling middle school readers is a complex challenge; by eighth grade, it is difficult to balance the instructional needs of struggling readers with the rigorous nature of content in a middle school ELA classroom. For example, the increased focus on teaching skills and strategies found in the low-track classes in this study may be in direct conflict with the Common Core requirement to focus on literature analysis. One possible solution to increasing the instructional quality in low-track classrooms is to focus on quality of teaching, choice of instructional practices, and use of course materials in those classrooms beyond increased use of worksheets, reading aloud, and watching videos. High-quality teaching in low-track classrooms can be produced by setting high expectations, providing a rigorous curriculum, and making sure that experienced teachers teach across track levels (Gamoran, 1993, 2004).

Instructional practices specific to ELA that correlate with student achievement include explicit strategy instruction, intellectual rigor, and use of guided practice (Grossman et al., 2013). Rigor of instruction can be increased by careful attention to the tasks given to students in low-track classrooms. Indeed, item-level analysis of the eighth-grade NAEP data showed that comprehension is determined by a combination of text level and task format (Valencia, Wixson, & Pearson, 2014), suggesting that one way to increase rigor of low-track classes may actually be to maintain (relatively) lower Lexile-level books but pair those texts with more rigorous tasks. It is important for teachers to consider how the text, reader, and assignment fit together and how instruction with texts that are too easy or too challenging may impact student growth.

Limitations

An important limitation of this study was that by eighth grade, as sampled students were followed longitudinally from kindergarten, they were only minimally nested within classrooms/teachers, and we were thus unable to use multilevel models to explicitly model instruction at the class level where tracking occurs. Although the teacher surveys were completed at the individual student level, it is likely that the instruction and text assignments were often given to multiple students in the class. Although we corrected standard errors for the clustering that occurred, more robustly nested data would allow

for fuller consideration of instructional variation within and across classrooms. Although we recognize this data limitation, because this is the first time that nationally representative data on text assignments were collected, we thought it important to pair these data with text complexity codes to provide an initial set of findings on how text assignment and text complexity varied among track levels.

Finally, although the ECLS-K data provide a detailed snapshot of eighth-grade ELA instruction, they are still limited in comparison with what might be learned about the quality of instruction from rich observational data. Overall, the instructional practices self-reported by the ECLS-K teachers captured primarily quantity, not quality, of use. Data from classroom observations can provide information on both the quality of instruction and the match between instructional practice and student need. Additionally, the book titles provided in the ECLS-K data reflect only the middle part of the school year (as the majority of teachers completed the survey in the late winter or early spring) and thus do not provide as robust a picture of the curriculum as would sampling books read from the beginning, middle, and end of the school year. Furthermore, the ECLS-K data offer information on the texts that students were reading in only one of their core courses; to get a full picture of the texts that middle school struggling readers were exposed to throughout the academic year, it would be beneficial to have data from other courses, such as mathematics, science, and social studies.

Likewise, although we could identify the particular books that students were assigned, we could not identify the types of tasks and assignments that teachers paired with the texts, which likely also contributed to student achievement. For example, even when struggling readers read the same texts as high-track students, the low-track students may have been completing skill-and-drill worksheets about the text while high-track students were writing analytical essays. Similarly, although time allocation to student writing was similar across track levels, we could not identify what types of writing assignments the students engaged in and whether there was a discrepancy in the cognitive demands of the writing assignments.

Another limitation in the ECLS-K data is that the eighth-grade teachers did not report on the use of small-group interventions (such as in Tier 2 or 3 of Response to Intervention frameworks) or on the intensity of instruction and intervention, an important component of providing instruction to struggling readers (Faggella-Luby & Wardwell, 2011). Future research, including qualitative classroom studies and new technologies for efficient classroom observations (Kelly, Olney, Donnelly, Nystrand, & D'Mello, 2018; Kersting, Sherin, & Stigler, 2014), can help untangle some of these issues.

Conclusion

In this study, we used the final wave of the ECLS-K to examine how teachers varied their instructional time allocation and instructional practices in eighth-grade ELA classes. We found that student achievement, associated here with track level, continued to be a significant predictor of what happens instructionally in the classroom. Students in low-track classes spent a larger portion of class time on skills and strategy instruction, as opposed to literature analysis and comprehension instruction, and read less challenging texts than their counterparts in grade-level and high-track courses. Further, our estimates showed that track level was associated with teacher adjustments in instructional practices and course content that went above and beyond what would be expected based on observed levels of student achievement.

NOTES

We thank Byeong-Young Cho for his thoughtful comments on our work.

¹ Observations are rounded to the nearest 10, per NCES guidelines for restricted-use data

² Although transcript-based measures of course taking are common in the study of mathematics (e.g., Riegle-Crumb & Grodsky, 2010), we favored a teacher report measure in this study for several reasons. First, English course titles may not correspond as consistently with content and rigor as titles do in mathematics. Second, because it is the teacher's own understanding of the level of the course that is most likely to influence instructional practices, we favored a sociopsychological rather than structural track location measure (see Gamoran & Mare, 1989). For both of these reasons, teacher-reported track levels may be more informative than course titles from the official school transcript for analyses such as these and have been used successfully in prior studies using large-scale data sets (e.g., Carbonaro, 2005; Condron, 2008). Finally, the measure employed here is consistent with our theoretical framework, emphasizing teachers' response to an institutional status label; teachers were asked about the course itself not the students. Ideally, we would be able to trace out the relation among official track labels, as indicated by transcript data, teachers' understanding of the track levels, and instructional practices, but unfortunately, transcript data are not available for the ECLS-K.

³ The cleaning required many hundreds of replace statements, as the errors were often unique. For example, there were 10 unique misspellings of *Fahrenheit 451* (by Ray Bradbury).

⁴ For a discussion of the limitations of quantitative text complexity, see Hiebert and Mesmer (2013). The biggest limitation of quantitative text complexity is that it does not account for reader variables, including interest and engagement. Other problems with quantitative text complexity measures include that content area texts often have inflated readability levels because of repetition of specialized vocabulary (a small group of high-frequency words can skew readability levels) and average sentence length not being strong predictors of readability.

REFERENCES

Acock, A. (2016). *A gentle introduction to Stata* (5th ed.). College Station, TX: Stata.

Allington, R.L. (2007). Intervention all day long: New hope for struggling readers. *Voices From the Middle*, 14(4), 7–14.

Allington, R.L. (2011). Reading intervention in the middle grades. *Voices From the Middle*, 19(2), 10–16.

Allington, R.L. (2013). What really matters when working with struggling readers. *The Reading Teacher*, 66(7), 520–530. <https://doi.org/10.1002/TRTR.1154>

Association for Middle Level Education (2010). *This we believe: Keys to educating young adolescents*. Westerville, OH: Author.

Biancarosa, G., & Snow, C.E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Bodovski, K., & Youn, M.J. (2011). The long term effects of early acquired skills and behaviors on young children's achievement in literacy and mathematics. *Journal of Early Childhood Research*, 9(1), 4–19. <https://doi.org/10.1177/1476718X10366727>

Borman, G.D., Slavin, R.E., Cheung, A.C.K., Chamberlain, A.M., Madden, N.A., & Chambers, B. (2005). The national randomized field trial of Success for All: Second-year outcomes. *American Educational Research Journal*, 42(4), 673–696. <https://doi.org/10.3102/00028312042004673>

Brasseur-Hock, I.F., Hock, M.F., Kieffer, M.J., Biancarosa, G., & Deshler, D.D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, 21(4), 438–452. <https://doi.org/10.1016/j.lindif.2011.01.008>

Carbonaro, W. (2005). Tracking, students' effort, and academic achievement. *Sociology of Education*, 78(1), 27–49. <https://doi.org/10.1177/003804070507800102>

Carbonaro, W.J., & Gamoran, A. (2002). The production of achievement inequality in high school English. *American Educational Research Journal*, 39(4), 801–827. <https://doi.org/10.3102/0002839004801>

Catts, H.W., Hogan, T.P., & Adlof, S.M. (2005). Developmental changes in reading and reading disabilities. In H.W. Catts & A.G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Mahwah, NJ: Erlbaum.

Caughlan, S., & Kelly, S. (2004). Bridging methodological gaps: Instructional and institutional effects of tracking in two English classes. *Research in the Teaching of English*, 39(1), 20–60.

Condron, D.J. (2008). An early start: Skill grouping and unequal reading gains in the elementary years. *The Sociological Quarterly*, 49(2), 363–394. <https://doi.org/10.1111/j.1533-8525.2008.00119.x>

Cunningham, J.W., & Mesmer, H.A. (2014). Quantitative measurement of text difficulty: What's the use? *The Elementary School Journal*, 115(2), 255–269. <https://doi.org/10.1086/678292>

Dennis, D.V. (2009). "I'm not stupid": How assessment drives (in) appropriate reading instruction. *Journal of Adolescent & Adult Literacy*, 53(4), 283–290. <https://doi.org/10.1598/JAAL.53.4.2>

Domina, T., Hanselman, P., Hwang, N., & McEachin, A. (2016). Detracking and tracking up: Mathematics course placements in California middle schools, 2003–2013. *American Educational Research Journal*, 53(4), 1229–1266. <https://doi.org/10.3102/0002831216650405>

Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming high and falling short: California's eighth-grade Algebra-for-all effort. *Educational Evaluation and Policy Analysis*, 37(3), 275–295. <https://doi.org/10.3102/0162373714543685>

Donaldson, M.L., LeChasseur, K., & Mayer, A. (2017). Tracking instructional quality across secondary mathematics and English language arts classes. *Journal of Educational Change*, 18(2), 183–207. <https://doi.org/10.1007/s10833-015-9269-x>

Dougherty, S.M., Goodman, J.S., Hill, D.V., Litke, E.G., & Page, L.C. (2015). Middle school math acceleration and equitable access to eighth-grade algebra: Evidence from the Wake County Public School System. *Educational Evaluation and Policy Analysis*, 37(1, Suppl.), 80S–101S. <https://doi.org/10.3102/0162373715576076>

Duncan, O.D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2), 210–217. <https://doi.org/10.2307/2088328>

Edmonds, M.S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K.K., & Schnakenberg, J.W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, 79(1), 262–300. <https://doi.org/10.3102/0034654308325998>

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>

Faggella-Luby, M.N., & Deshler, D.D. (2008). Reading comprehension in adolescents with LD: What we know; what we need to learn. *Learning Disabilities Research & Practice*, 23(2), 70–78. <https://doi.org/10.1111/j.1540-5826.2008.00265.x>

Faggella-Luby, M., & Wardwell, M. (2011). RTI in a middle school: Findings and practical implications of a Tier 2 reading comprehension study. *Learning Disability Quarterly*, 34(1), 35–49. <https://doi.org/10.1177/073194871103400103>

Farkas, G. (2011). Middle and high school skills, behaviors, attitudes, and curriculum enrollment, and their consequences. In G.J. Duncan & R.J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 71–89). New York, NY: Russell Sage Foundation.

Fiske, S.T., & Taylor, S.E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill.

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C.A., Dimino, J., ... Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE 2016-4008). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education*, 60(3), 135–155. <https://doi.org/10.2307/2112271>

Gamoran, A. (1993). Alternative uses of ability grouping in secondary schools: Can we bring high-quality instruction to low-ability classes? *American Journal of Education*, 102(1), 1–22. <https://doi.org/10.1086/444056>

Gamoran, A. (2004). Classroom organization and instructional quality. In H.J. Walberg, A.J. Reynolds, & M.C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 141–155). Greenwich, CT: Information Age.

Gamoran, A., & Carbonaro, W.J. (2002). High school English: A national portrait. *High School Journal*, 86(2), 1–13. <https://doi.org/10.1353/hsj.2002.0021>

Gamoran, A., & Mare, R.D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology*, 94(5), 1146–1183. <https://doi.org/10.1086/229114>

Gamoran, A., Nystrand, M., Berends, M., & LePore, P.C. (1995). An organizational analysis of the effects of ability grouping. *American Educational Research Journal*, 32(4), 687–715. <https://doi.org/10.3102/00028312032004687>

Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., ... Smither Wulsen, C. (2016). *Teaching secondary students to write effectively* (NCEE 2017-4002). Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gritter, K. (2012). Permeable textual discussion in tracked language arts classrooms. *Research in the Teaching of English*, 46(3), 232–259.

Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practices in middle school English language arts and teachers' value-added scores. *American Journal of Education*, 119(3), 445–470. <https://doi.org/10.1086/669901>

Guarino, C.M., Hamilton, L.S., Lockwood, J.R., & Rathbun, A.H. (2006). *Teacher qualifications, instructional practices, and reading and mathematics gains of kindergartners* (NCES 2006-031). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Hallinan, M.T. (1994). Tracking: From theory to practice. *Sociology of Education*, 67(2), 79–84. <https://doi.org/10.2307/2112697>

Hamilton, D.L. (1981). Stereotyping and intergroup behavior: Some thoughts on the cognitive approach. In D.L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 333–353). Hillsdale, NJ: Erlbaum.

Harvey, S., & Goudvis, A. (2007). *Strategies that work: Teaching comprehension for understanding and engagement* (2nd ed.). Portland, ME: Stenhouse.

Hiebert, E.H., & Mesmer, H.A. (2013). Upping the ante of text complexity in the Common Core State Standards: Examining its potential impact on young readers. *Educational Researcher*, 42(1), 44–51. <https://doi.org/10.3102/0013189X12459802>

Hodge, E.M. (2015). *Rigor for all? The enduring tension between standardization and differentiation in the era of the Common Core State Standards* (Doctoral dissertation). Retrieved from https://etda.libraries.psu.edu/files/final_submissions/10539

Holland, D., Skinner, D., Lachicotte, W., Jr., & Cain, C. (1998). *Identity and agency in cultural worlds*. Cambridge, MA: Harvard University Press.

Hong, G., & Hong, Y. (2009). Reading instruction time and homogeneous grouping in kindergarten: An application of marginal mean weighting through stratification. *Educational Evaluation and Policy Analysis*, 31(1), 54–81. <https://doi.org/10.3102/0162373708328259>

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, 85(4), 512–552. <https://doi.org/10.3102/0034654314561338>

Jussim, L., & Harber, K.D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9(2), 131–155. https://doi.org/10.1207/s1527957pspr0902_3

Kamil, M.L., Borman, G.D., Dole, J., Kral, C.C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE 2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Karlson, K.B. (2015). Expectations on track: High school tracking and adolescent educational expectations. *Social Forces*, 94(1), 115–141. <https://doi.org/10.1093/sf/sov006>

Kelly, S. (2010). The prevalence of developmental instruction in public and Catholic schools. *Teachers College Record*, 112(9), 2405–2440.

Kelly, S., & Carbonaro, W. (2012). Curriculum tracking and teacher expectations: Evidence from discrepant course taking models. *Social Psychology of Education*, 15(3), 271–294. <https://doi.org/10.1007/s11218-012-9182-6>

Kelly, S., & Majerus, R. (2011). School-to-school variation in disciplined inquiry. *Urban Education*, 46(6), 1553–1583. <https://doi.org/10.1177/0042085911413151>

Kelly, S., Olney, A.M., Donnelly, P., Nystrand, M., & D'Mello, S.K. (2018). Automatically measuring question authenticity in real-world classrooms. *Educational Researcher*, 47(7), 451–464. <https://doi.org/10.3102/0013189X18785613>

Kelly, S., & Ye, F. (2017). Accounting for the relationship between initial status and growth in regression models. *Journal of*

Experimental Education, 85(3), 353–375. <https://doi.org/10.1080/00220973.2016.1160357>

Kersting, N.B., Sherin, B.L., & Stigler, J.W. (2014). Automated scoring of teachers' open-ended responses to video prompts: Bringing the Classroom-Video-Analysis assessment to scale. *Educational and Psychological Measurement*, 74(6), 950–974. <https://doi.org/10.1177/0013164414521634>

Kieffer, M.J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, 100(4), 851–868. <https://doi.org/10.1037/0022-0663.100.4.851>

Kieffer, M.J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal*, 48(5), 1187–1225. <https://doi.org/10.3102/0002831211419490>

Kim, Y.-S., & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly*, 49(3), 269–281. <https://doi.org/10.1002/rrq.74>

Lauen, D.L., & Gaddis, S.M. (2016). Accountability pressure, academic standards, and educational triage. *Educational Evaluation and Policy Analysis*, 38(1), 127–147. <https://doi.org/10.3102/016237315598577>

Legette, K. (2018). School tracking and youth self-perceptions: Implications for academic and racial identity. *Child Development*, 89(4), 1311–1327. <https://doi.org/10.1111/cdev.12748>

Lewis, A.E., & Diamond, J.B. (2015). *Despite the best intentions: How racial inequality thrives in good schools*. New York, NY: Oxford University Press.

Li-Grining, C.P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, 46(5), 1062–1077. <https://doi.org/10.1037/a0020066>

Luyten, H., & Bruggencate, G.T. (2011). The presence of Matthew effects in Dutch primary education: Development of language skills over a six-year period. *Journal of Learning Disabilities*, 44(5), 444–458. <https://doi.org/10.1177/0022219411410289>

McCoach, D.B., O'Connell, A.A., Reis, S.M., & Levitt, H.A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology*, 98(1), 14–28. <https://doi.org/10.1037/0022-0663.98.1.14>

Mickelson, R.A. (2015). The cumulative disadvantages of first- and second-generation segregation for middle school achievement. *American Educational Research Journal*, 52(4), 657–692. <https://doi.org/10.3102/0002831215587933>

Montt, G. (2011). Cross-national differences in educational achievement inequality. *Sociology of Education*, 84(1), 49–68. <https://doi.org/10.1177/0038040710392717>

Morgan, P.L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, 31(4), 187–198.

Morgan, P.L., Farkas, G., & Maczuga, S. (2015). Which instructional practices most help first-grade students with and without mathematical difficulties? *Educational Evaluation and Policy Analysis*, 37(2), 184–205. <https://doi.org/10.3102/0162373714536608>

Morgan, P.L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities*, 44(5), 472–488. <https://doi.org/10.1177/0022219411414010>

Najarian, M., Pollack, J.M., & Sorongon, A.G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): Psychometric report for the eighth grade (NCES 2009-002)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Nakamoto, J., Lindsey, K.A., & Manis, F.R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing*, 20(7), 691–719. <https://doi.org/10.1007/s11145-006-9045-7>

National Center for Education Statistics. (2017). *The Nation's Report Card: 2017 NAEP mathematics and reading assessments*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Author.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010a). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010b). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects: Appendix A: Research supporting key elements of the standards and glossary of key terms*. Washington, DC: Authors.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>

Northrop, L. (2018). *Reading intervention in middle school: Organization and implementation*. Manuscript submitted for publication.

Northrop, L., Borsheim-Black, C., & Kelly, S. (2018, April). *Matching students to books: The cultural content of eighth-grade literature assignments*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Northrop, L., & Kelly, S. (2018). AYP status, urbanicity, and sector: School-to-school variation in instruction. *Urban Education*, 53(5), 591–620. <https://doi.org/10.1177/0042085915618710>

Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). New Haven, CT: Yale University Press.

Page, R.N. (1991). *Lower-track classrooms: A curricular and cultural perspective*. New York, NY: Teachers College Press.

Pallas, A.M., Entwistle, D.R., Alexander, K.L., & Stluka, M.F. (1994). Ability-group effects: Instructional, social, or institutional? *Sociology of Education*, 67(1), 27–46. <https://doi.org/10.2307/2112748>

Puranik, C.S., & Lonigan, C.J. (2014). Emergent writing in preschoolers: Preliminary evidence for a theoretical framework. *Reading Research Quarterly*, 49(4), 453–467. <https://doi.org/10.1002/rrq.79>

Raghunathan, T.E., Lepkowski, J.M., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology*, 27(1), 85–96.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S.W., Rowan, B., & Cheong, Y.F. (1993). Higher order instructional goals in secondary schools: Class, teacher, and school influences. *American Educational Research Journal*, 30(3), 523–553. <https://doi.org/10.3102/00028312030003523>

Reardon, S.F., & Yun, J.T. (2003). Integrating neighborhoods, segregating schools: The retreat from school desegregation in the South, 1990–2000. *North Carolina Law Review*, 81(4), 1563–1596.

Reed, D.K., Petscher, Y., & Truckenmiller, A.J. (2017). The contribution of general reading ability to science achievement. *Reading Research Quarterly*, 52(2), 253–266. <https://doi.org/10.1002/rrq.158>

Riegle-Crumb, C., & Grodsky, E. (2010). Racial-ethnic differences at the intersection of math course-taking and achievement. *Sociology of Education*, 83(3), 248–270. <https://doi.org/10.1177/0038040710375689>

Risko, V.J., & Walker-Dalhouse, D. (2015). Best practices to change the trajectory of struggling readers. In L.B. Gambrell & L.M.

Morrow (Eds.), *Best practices in literacy instruction* (5th ed., pp. 107–126). New York, NY: Guilford.

Rogers, T., Winters, K.L., Bryan, G., Price, J., McCormick, F., House, L., ... Sinclair, C. (2006). Developing the IRIS: Toward situated and valid assessment measures in collaborative professional development and school reform in literacy. *The Reading Teacher*, 59(6), 544–553. <https://doi.org/10.1598/RT.59.6.4>

Schulte, A.C., Stevens, J.J., Elliott, S.N., Tindal, G., & Nese, J.F.T. (2016). Achievement gaps for students with disabilities: Stable, widening, or narrowing on a state-wide reading comprehension test? *Journal of Educational Psychology*, 108(7), 925–942. <https://doi.org/10.1037/edu0000107>

Shanahan, T., Callison, K., Carriere, C., Duke, N.K., Pearson, P.D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practical guide* (NCEE 2010-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Shavit, Y., & Müller, W. (2000). Vocational secondary education, tracking, and social stratification. In M.T. Hallinan (Ed.), *Handbook of the sociology of education* (pp. 437–452). New York, NY: Springer.

Slavin, R.E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(3), 293–336. <https://doi.org/10.3102/00346543057003293>

Slavin, R.E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, 43(3), 290–322. <https://doi.org/10.1598/RRQ.43.3.4>

Snijders, T.A.B., & Bosker, R.J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Snow, C.E., Burns, M.S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Steenbergen-Hu, S., Makel, M.C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research*, 86(4), 849–899. <https://doi.org/10.3102/0034654316675417>

Sun, J., Anderson, R.C., Lin, T.-J., & Morris, J. (2015). Social and cognitive development during collaborative reasoning. In L.B. Resnick, S.C. Austerhan, & S.N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 63–76). Washington, DC: American Educational Research Association.

Swanson, E., Wanzeck, J., Vaughn, S., Fall, A.-M., Roberts, G., Hall, C., & Miller, V.L. (2017). Middle school reading comprehension and content learning intervention for below-average readers. *Reading & Writing Quarterly*, 33(1), 37–53. <https://doi.org/10.1080/10573569.2015.1072068>

Taylor, B.M., Pearson, P.D., Peterson, D.S., & Rodriguez, M.C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal*, 104(1), 3–28. <https://doi.org/10.1086/499740>

Tourangeau, K., Nord, C., Lê, T., Sorongon, A.G., & Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): Combined user's manual for the ECLS-K eighth-grade and K–8 full sample data files and electronic codebooks* (NCES 2009-004). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Valencia, S.W., & Buly, M.R. (2004). Behind test scores: What struggling readers really need. *The Reading Teacher*, 57(6), 520–531.

Valencia, S.W., Wixson, K.K., & Pearson, P.D. (2014). Putting text complexity in context: Refocusing on comprehension of complex text. *The Elementary School Journal*, 115(2), 270–289. <https://doi.org/10.1086/678296>

Van de Werfhorst, H.G., & Mijs, J.J.B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407–428. <https://doi.org/10.1146/annurev.soc.012809.102538>

Van Houtte, M., Demanet, J., & Stevens, P.A.J. (2013). Curriculum tracking and teacher evaluations of individual students: Selection, adjustment, or labeling? *Social Psychology of Education*, 16(3), 329–352. <https://doi.org/10.1007/s11218-013-9216-8>

Vaughn, S., Swanson, E.A., Roberts, G., Wanzeck, J., Stillman-Spisak, S.J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading Research Quarterly*, 48(1), 77–93. <https://doi.org/10.1002/rrq.039>

Velicer, W.F., & Jackson, D.N. (1990). Component analysis versus common factor analysis: Some issues in selecting appropriate procedure. *Multivariate Behavioral Research*, 25(1), 1–28. https://doi.org/10.1207/s15327906mbr2501_1

Wei, X., Blackorby, J., & Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, 78(1), 89–106. <https://doi.org/10.1177/001440291107800106>

Westat. (2007). *Early Childhood Longitudinal Study, spring 2007 grade 8 English teacher questionnaire*. Rockville, MD: Author. Retrieved from https://nces.ed.gov/ecls/pdf/eighthgrade/teacher_english.pdf

Williamson, G.L., Fitzgerald, J., & Stenner, A.J. (2014). Student reading growth illuminates the Common Core text-complexity standard: Raising both bars. *The Elementary School Journal*, 115(2), 230–254. <https://doi.org/10.1086/678295>

Wilson, J.A., Faggella-Luby, M., & Wei, Y. (2013). Planning for adolescent Tier 3 reading instruction. *Teaching Exceptional Children*, 46(1), 26–34. <https://doi.org/10.1177/004005991304600104>

Wonder-McDowell, C., Reutzel, D.R., & Smith, J.A. (2011). Does instructional alignment matter? Effects on struggling second graders' reading achievement. *The Elementary School Journal*, 112(2), 259–279. <https://doi.org/10.1086/661524>

Worthy, J. (2010). Only the names have been changed: Ability grouping revisited. *The Urban Review*, 42(4), 271–295. <https://doi.org/10.1007/s11256-009-0134-1>

Submitted February 9, 2018

Final revision received October 13, 2018

Accepted October 15, 2018

LAURA NORTHRUP (corresponding author) is an assistant professor in the Department of Teacher Education at Cleveland State University, Ohio, USA; email l.northrop@csuohio.edu. Her work focuses on educational policy, middle school literacy instruction, and struggling readers.

SEAN KELLY is a professor in the Department of Administrative and Policy Studies at the University of Pittsburgh, Pennsylvania, USA; email spkelly@pitt.edu. His research focuses on teacher effectiveness, student engagement, and the social organization of schools.

APPENDIX

Eighth-Grade ELA Instructional Practices in the ECLS-K

Item	M	SD	Regression coefficients	
			Low track	High track
<i>Comprehension instruction items</i>				
Talk with one another about what they read	3.12	0.80	-0.03 (0.07)	0.18*** (0.05)
Write about something that they read	3.10	0.69	-0.06 (0.05)	0.09* (0.04)
Make predictions about what they are reading	3.21	0.76	-0.06 (0.05)	0.08 (0.04)
Make generalizations and draw inferences based on what they read	3.35	0.68	-0.11** (0.04)	0.12** (0.04)
Describe the style or structure of the text that they read	2.72	0.80	-0.17** (0.06)	0.17*** (0.05)
<i>Additional instructional practices</i>				
Work in a reading workbook or on a worksheet	2.43	0.97	0.28*** (0.07)	-0.09 (0.06)
Have assigned homework	3.28	0.82	-0.07 (0.06)	0.11* (0.04)
Take quizzes or tests	2.55	0.58	0.00 (0.04)	0.02 (0.04)
Revise a report or paper that they have wrote	2.28	0.57	0.00 (0.04)	0.07 (0.04)
Do a group activity or project about what they read	2.21	0.69	0.01 (0.05)	0.17*** (0.05)
Watch movies, videos, or television; listen to tapes, compact discs, or records	1.76	0.64	0.12* (0.05)	-0.01 (0.04)
Read aloud	2.98	0.88	0.16* (0.07)	-0.06 (0.06)

Note. Means are for the weighted sample, and standard deviations are for the unweighted sample. Control variables include student, teacher, classroom, and school characteristics (see Table 3 for examples of full regression models). Standard errors are in parentheses.

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



ILA's Literacy Glossary—

Together, we can define
our profession,
one word at a time

Learn more and share your feedback at
literacyworldwide.org/glossary

INTERNATIONAL
LITERACY
ASSOCIATION