INTERNATIONAL LITERACY ASSOCIATION
READING RESEARCH QUARTERLY

The Decline of Comprehension-Based Silent Reading Efficiency in the United States: A Comparison of Current Data With Performance in 1960

Alexandra N. Spichtig
Reading Plus, Winooski, Vermont, USA

Elfrieda H. Hiebert
TextProject, Santa Cruz, California, USA

Christian Vorstius
Bergische Universität Wuppertal, NRW, Germany

Jeffrey P. Pascoe
Reading Plus, Winooski, Vermont, USA

P. David Pearson
University of California, Berkeley, USA

Ralph Radach
Bergische Universität Wuppertal, NRW, Germany

Reading Research Quarterly, 0(0)
pp. 1–21 | doi:10.1002/rrq.137

ABSTRACT

The present study measured the comprehension-based silent reading efficiency of U.S. students in grades 2, 4, 6, 8, 10, and 12. Students read standardized grade-level passages while an eye movement recording system was used to measure reading rate, fixations (eye stops) per word, fixation durations, and regressions (right-to-left eye movements) per word. Eye movement recordings were regarded as valid only if students demonstrated a comprehension level of at least 70% after reading a passage and answering a series of true/false questions. Reading rates increased over grades, with two exceptions: (a) between grades 6 and 8, growth in reading rate appeared to plateau; and (b) between grades 10 and 12, reading rate increases were seen only among students in the upper two quartiles. Changes in the other three efficiency measures reflected similar patterns of reading efficiency development over grades. The reading efficiency of students in this study was also compared with that of a sample of students from 1960, using norms reported by Taylor (1965) and validated by Carver (1989). Comprehension-based silent reading rates in grade 2 were comparable across the 50-year span, but the cross-grade growth trajectory was much shallower in the present study than it was in 1960. These results suggest that present-day students may not achieve the same level of word-reading automaticity as did their 1960 counterparts.

Through the early 19th century, reading pedagogy in the Western world was dominated by the readily observable oral recitation of letter sounds, syllables, words, and passages, largely learned by rote and with religious intent (Smith, 1965). Silent reading of secular material became an increasingly popular pastime during the latter part of the 19th century (Smith, 1965). Yet, despite prominent critics, reading instruction remained firmly rooted in an oral tradition at least until the 1920s (Pearson & Goodin, 2010). Over a century ago, Huey (1908) voiced concerns that reading in school was too often an exercise in speaking at the expense of "thought-getting" (p. 359; i.e., grasping the ideas and meanings in the text), which he argued could only be achieved through silent reading practice in the service of the reader's own purposes (i.e., with an intent to comprehend). Thorndike (1917) reached a similar conclusion after finding that many students he tested were fluent oral readers but did not understand what they read. More recently, Samuels (2007) expressed concerns that reading fluency in schools was too often being operationalized as reading aloud at a good rate ("barking at print"), while disregarding the ability to construct meaning from text. Although there appears to be broad agreement on the importance of proficient silent reading with good comprehension, there is far less agreement on how we can help students attain that goal and, specifically, on the relative value of devoting instructional time to oral and silent reading fluency at different grade and ability levels (for an in-depth overview, see Pearson & Goodin, 2010).

A key development in these deliberations was the report of the National Reading Panel (NRP; National Institute of Child Health & Human Development [NICHD], 2000), which included an evaluation of selected research addressing the efficacy of devoting instructional time to guided, repeated oral reading and/or sustained silent reading. The NRP concluded that there was sufficient evidence to support the efficacy of the former, but not the latter, to the effect that devoting class time to silent reading could not be recommended as an evidence-based approach. This conclusion was followed, during the era of the No Child Left Behind Act (2002), by a shift in emphasis in U.S. education toward oral reading and the use of oral reading fluency assessments. Yet, our best available evidence suggests a declining association between oral reading fluency and reading comprehension as students advance to higher grades (e.g., Denton et al., 2011; O'Brien, Wallot, Haussmann, & Kloos, 2014).

The NRP report (NICHD, 2000) also led to the reexamination of silent reading practices. Concern was expressed that many students were using their silent reading time poorly (Griffith & Rasinski, 2004) and that measures of student accountability and comprehension were often lacking (Reutzel, Fawson, & Smith, 2008). Reutzel et al. also showed that scaffolded silent reading practice with active monitoring was as effective as guided or repeated oral reading in promoting the development of fluency and comprehension in grade 3 students.

The Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSSO], 2010a) mark another major development in reading policy. The Common Core makes explicit the expectation that students develop the capacity to construct meaning in response to increasingly challenging texts, echoing the same goal of reading instruction espoused by Huey (1908) and Thorndike (1917): proficient silent reading with good comprehension. A primary plank of the Common Core is the need for students to have increased exposure to complex texts at every grade level. Hiebert and Mesmer (2013) argued that the new standards are going to be exceedingly challenging for a significant portion of U.S. students who have shown a lack of fluency with texts at current levels of difficulty (Daane, Campbell, Grigg, Goodman, & Oranje, 2005).

The present study sought to evaluate the rates at which 21st-century U.S. students read grade-level text silently with good comprehension—a construct that Hiebert, Wilson, and Trainin (2010) labeled comprehension-based silent reading rates. Previous comprehensive studies on silent reading rates are few, and certainly no contemporary studies exist to document the course of comprehension-based silent reading efficiency development from the early elementary grades through high school.

Previous Studies of Silent Reading Efficiency

Historically, Taylor's (1965; Taylor, Frackenpohl, & Pettee, 1960) work has provided the most robust data on silent reading efficiency. In the 1960 study, approximately 12,000 students from first grade through college were assessed on silent reading rate and comprehension. In addition, eye movement data (fixations per word, fixation durations, and regressions per word) were gathered as students read grade-level passages and responded to true/false comprehension questions. Students had to achieve a standard of 70% correct on the comprehension assessment in order for their efficiency data to be included in the analysis.

In Carver's (1989) review of research on the silent reading efficiency of U.S. students, Taylor et al.'s 1960 results were compared with those of three other projects: (a) a National Assessment of Educational Progress (NAEP) study (Gallo, 1972), (b) rates that Carver interpolated from norms reported for grades 9–16 on the Nelson–Denny Reading Test (Nelson, Brown, & Denny, 1960), and (c) Carver's (1983) own work on the flexibility of reading rate. In the NAEP study, students at ages 9, 13, and 17 and young adults read two passages, each followed by five multiple-choice comprehension questions. Median reading rates at the 25th, 50th, and 75th percentiles were reported in words per minute (wpm) for each age group but without taking comprehension into account. Also reported were the percentages of students who fell into four broad rate bands (<100 wpm, 100–199 wpm, 200–299 wpm, and >300 wpm) and the percentages of students in each rate band who correctly answered at least four of the five comprehension questions.

Carver (1983) studied the reading rates of students in grades 4–12 as they read texts of varying difficulty (i.e., texts meant for grades 1, 4, 7, 10, 13, and 16). Text assignments were based on students' performance on a standardized reading test (National Reading Standards; Carver, 1977). Carver therefore assumed that students read the study passages with good comprehension, but comprehension was not assessed.

Carver's (1989) review concluded that the data in Taylor's (1965) publication provided the most useful standard for national norms because the study involved a national sample with a wide range of grades, used grade-leveled texts, assessed comprehension, and reported comprehension-based silent reading rates (i.e., rates achieved when students demonstrated at least 70% comprehension).

Measuring Visuomotor Behavior During Silent Reading

Although comprehension-based silent reading rate provides a useful measure of reading efficiency, more detailed and valuable insights into the hidden processes of reading can be gained by recording eye movement activity during reading. The observable behavior associated with continuous reading includes a well-coordinated pattern of fast eye movements (referred to as *saccades*) that are interrupted by fixations (periods of relative stability) during which visual information is acquired. This visuomotor behavior reflects the mental workload associated with various aspects of information processing and is therefore closely related to word processing, levels of fluency, comprehension, and individual reading strategies (for overviews, see Radach & Kennedy, 2013; Rayner, 2009).

Fixation durations are typically in the range of 200–250 ms in adult readers and 300–350 ms in children. As text difficulty increases, readers make more fixations, longer fixations, and more regressive eye movements back to regions in the text that had been processed previously. Short-range regressions are frequently the result of either positioning errors in oculomotor control (e.g., the eyes might have overshot the intended saccade target) or problems with lexical processing, typically when the meaning of a word is not accessed immediately. Larger regressions are much less frequent and often indicate processing difficulty on the level of local syntax or semantics or a breakdown in comprehension on a more global level (Inhoff, Weger, & Radach, 2005; Vorstius, Radach, & Lonigan, 2014). The percentage of regressions is often regarded as a key indicator of reading difficulty; in fluent adult readers of English, 10–15% of all eye movements during sentence reading are regressions. In contrast, developing readers' regression rates are more in the order of 30% (e.g., McConkie et al., 1991).

There have been many studies of the eye movements of efficient adult readers but only a few tracking the development of children's eye movements across the course of schooling (for a review, see Blythe & Joseph, 2011). The existing work suggests that younger readers make shorter saccades, more and longer fixations, and more frequent regressions than efficient, adult readers do (e.g., Huestegge, Radach, Corbic, & Huestegge, 2009; McConkie et al., 1991; Taylor, 1965) and that struggling readers of all ages exhibit similar patterns to those of developing readers (Ashby, Rayner, & Clifton, 2005). McConkie et al.'s study is noteworthy in that it was the first extensive developmental study of elementary students (grades 1–5) that utilized eye movement recording technology with letter-level accuracy of measurement. More recently, Vorstius et al. (2014) reported a very detailed analysis of eye movements in a cross-sectional sample of students in grades 1–5 reading identical sentences during oral and silent reading. As of this writing, however, only Taylor has provided comprehension-based silent reading efficiency data that span the elementary through high school grades.

The focus of the present study was to update information on the comprehension-based silent reading efficiency of U.S. students and compare their performance to that of a cohort from a half-century ago (Taylor, 1965). Four efficiency measures were examined: (a) comprehension-based silent reading rate (time required to read a passage with at least 70% comprehension, converted to wpm), (b) fixation count (average number of fixations per word), (c) fixation duration (measured in ms), and (d) regression count (average number of regressions per word).

Methods

The methodology described in this study closely followed the tasks and procedures of Taylor (1965; Taylor, personal communication, July 28, 2009; Taylor et al., 1960). Comparison data on rates, fixations, fixation durations, and regressions were also derived from these sources.

Participants

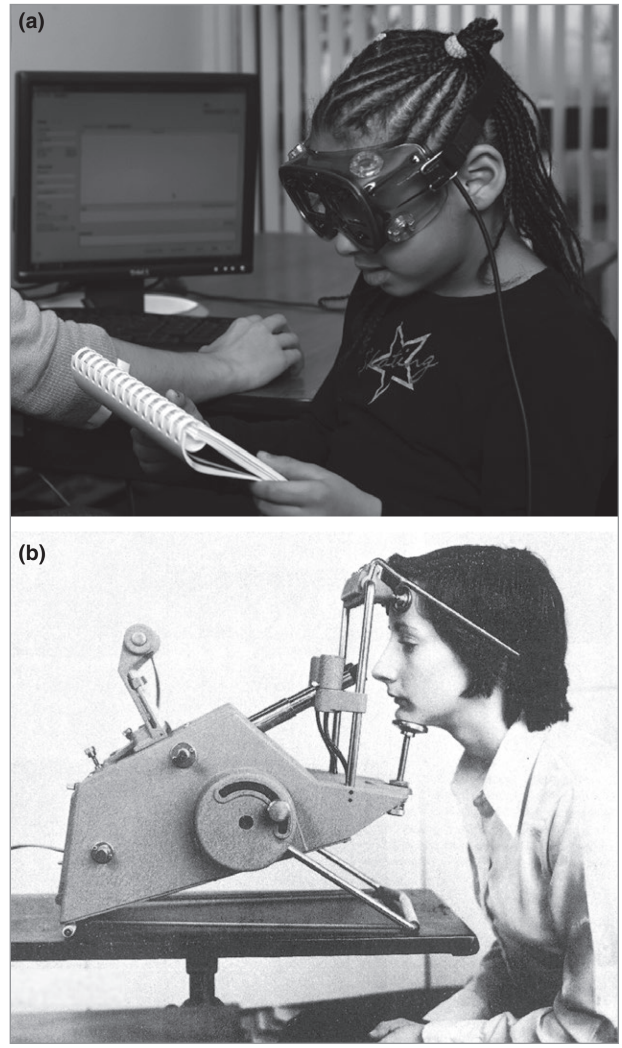
The current study took place during the 2010–2011 school year. It involved 2,203 students in 34 public schools located in 16 states representing all regions of the United States. Students in the sample were distributed across six grades as follows: grade 2: 414, grade 4: 451, grade 6: 316, grade 8: 519, grade10: 265, and grade 12: 238. The racial and ethnic distribution of the sample was 60% white, 20% Hispanic, 16% black, 3% Asian, and 1% other. By comparison, the U.S. Census Bureau (2011) reported the population distribution of young Americans in 2011 as 54% white, 23% Hispanic, 14% black, 4% Asian, and 5% other. Data on free and reduced-price lunch were available for 93% of the schools in the sample. Overall, 49% of the students in the study were eligible for free or reduced-price lunch, as compared with the National Center for Education Statistics (2013a) estimate of 48.1% for the 2010–2011 cohort of U.S.

students. Performance data on state reading/language arts assessments were obtained from 93% of the schools. These data showed that an average of 69.7% of students attained proficiency as defined by their various state standards. Each participating school selected at least 30 students who had performed at below-average, average, and above-average levels on their 2010 state assessment. English learners and special education students were not included in this study. Gender information, available for 94% of the students, showed an approximately equal distribution of males and females in each grade.

Data from Taylor et al.'s (1960) study were originally collected by associate investigators recruited from 39 colleges and universities located in 19 states, also representing all regions of the United States. As in our 2011 study, school personnel selected students for participation in Taylor et al.'s study. The criteria for selecting participating schools were that (a) the students represent the "average socio-economic level for the area" (Taylor et al., 1960, p. 5), and (b) no more than 10% could be private or parochial schools. It was also specified that participating classrooms represented students of mixed-ability levels and a balance of females and males. Beyond this, no demographic or aptitude information were reported in the original study. Reference to the 1960 census data (U.S. Census Bureau, 1960) shows that the states included in the 1960 study provided a fairly representative sample of the U.S. population at the time: 89.7% white (vs. 88.6% nationally), 9.1% African American (vs. 10.5% nationally), 1.2% other (vs. 0.9% nationally), and 23.3% low income (vs. 21.4% nationally).

The data from Taylor et al.'s (1960) study, used for historical comparisons, included grades 2, 4, 6, 8, and 11 (n = 6,429). The 1960 study did not collect data on students in grades 10 and 12; instead, reported values were extrapolated from adjacent grades. Consequently, data for grade 11 students in Taylor et al.'s study were compared with data for grade 12 students in the 2011 sample, and no historical comparisons were made with grade 10 students.

FIGURE 1
Students Reading Sample Texts During Eye Movement Recordings Using (a) the Visagraph and (b) the Reading Eye Camera



Note. Photos courtesy of Reading Plus/Taylor Associates Communications. Reprinted with permission.

Materials

Instrumentation

Data on comprehension-based silent reading rate and visual activity were gathered with the Visagraph Eye Movement Recording System (Taylor, 2009), a portable unit commonly used for this purpose. The Visagraph uses infrared emitters and sensors in a goggle-style apparatus to measure binocular eye movements (corneal reflections) during reading at a sampling rate of 60 Hz. Figure 1a shows a student wearing Visagraph goggles. Spichtig, Vorstius, Greene, and Radach (2009) demonstrated that, with respect to the measures reported within this article, the Visagraph produces data comparable to more sophisticated eye movement recording systems. Eye movement data in Taylor et al.'s (1960) study were collected using a Reading Eye camera[1] (see Figure 1b), a device designed for binocular corneal reflection eye movement photography.

In both studies, measures collected while students read 100-word passages included (a) comprehension-based silent reading rate (measured in wpm), (b) fixation count, (c) fixation duration (measured in ms), and (d) number of regressions. Because of limitations of the recording system, the fixation durations reported in this article include saccade time (i.e., eye movements between fixations). Inclusion of saccade time lengthens fixation duration, but this measure remains a valid indicator of processing workload (Vonk & Cozijn, 2003) in that saccades are generally short (20–40 ms) and processing continues during saccades (Irwin, 1998). Importantly, the fixation durations reported in the 1960 study (Taylor, 1965) also included saccade times, so data from the two studies are comparable. Also in keeping with the 1960 study, the regression data reported in this article reflect only short-range regressions (up to three words in length).

Reading Assessment Content

Participants read five 100-word passages, all of which had been used in the 1960 study (Taylor, 1965). The first passage was two levels below students' designated grade and served as an introduction, and the remaining four passages were at students' grade level. Taylor et al. (1960) used several readability formulas to verify the grade-level status of passages. In the current study, two of the four passages remained unchanged from their original form (Taylor et al., 1960). These original passages were presented first, ensuring comparable procedures to those of the 1960 study. The other two passages were slightly revised if necessary to comply with the accelerated Lexile (L) levels of the Common Core (NGA Center & CCSSO, 2010b).[2] Table 1 provides an overview of the Lexile measures of passages on each grade level and presents the mean word length (in letters) and variability across these passages. Figure 2 shows a passage at readability level 5. Each reading selection was followed by 10 true/false comprehension questions (see Figure 3). These questions were the same as those used in 1960, with the exception of minor wording changes to some questions used with the revised accelerated passages. To maintain a comparable standard with the 1960 study, reading performances were only considered to be successful if a student demonstrated at least 70% comprehension.[3]

FIGURE 2
A Typical 100-Word Visagraph Passage

Whales breathe air just as human beings do. But they do not breathe through their mouths as we can. All of their breathing is through two little nostrils placed on top of their heads. When the whale sinks under the water, special muscles shut his nose to keep out water. This makes it possible for him to swim with his mouth open. He can capture prey and even eat under water without danger of drowning. A whale can hold his breath and remain under water for an hour without any difficulty. As he rises to the surface, he blows the hot air from his lungs with tremendous force. People used to think this fountain of steam was water.

Note. From *Visagraph® Test Booklet* (p. 54), by S.E. Taylor, H. Frackenpohl Morris, and C.E. White, 2004, New York, NY: Taylor Associates Communications. Copyright 2004 by Taylor Associates Communications. Reprinted with permission.

FIGURE 3
An Example of a True/False Comprehension Check

Whales

Answer Yes or No

1. Whales breathe air as human beings do.
2. They breathe through their mouths.
3. Their nostrils are placed on top of their heads.
4. The whale's nostrils close as he sinks under water.
5. A whale can swim under water with his mouth open.
6. He must rise to the surface to eat.
7. A whale holds his breath under water.
8. A whale can remain under water for several hours without difficulty.
9. Whales blow hot water from their lungs when they rise to the surface.
10. When a whale breathes out, it looks like a fountain of steam.

Note. From *Visagraph® Test Booklet* (p. 54), by S.E. Taylor, H. Frackenpohl Morris, and C.E. White, 2004, New York, NY: Taylor Associates Communications. Copyright 2004 by Taylor Associates Communications. Reprinted with permission.

TABLE 1
Lexile Scores and Mean Word Length of Original and Accelerated Passages

	Original passages				Accelerated passages			
	Lexile scores		Word length		Lexile scores		Word length	
Grade	Passage 1	Passage 2	Mean	Standard deviation	Passage 3	Passage 4	Mean	Standard deviation
2	420	390	4.12	1.82	540	540	4.14	1.57
4	750	740	4.43	1.93	790	840	4.31	2.07
6	900	920	4.65	2.05	940	960	4.79	2.06
8	1050	1050	4.81	2.29	1122	1122	4.91	2.32
10	1150	1170	4.99	2.59	1251	1251	4.93	2.66
12	1200	1200	5.38	2.85	1285	1285	5.32	2.96

Note. Two original passages (1 and 2) and two accelerated passages (3 and 4) were used in each grade.

Procedure

Recording Procedure in 2011

Two teachers at each school were trained to conduct recordings using Visagraphs provided for the study. They participated in either a regional on-site training or a two-part webinar in which they learned how to gather recordings with consistency and fidelity. The teachers were provided with a step-by-step recording script to ensure consistency regarding student directions and Visagraph administration. Students first completed a passage as a practice trial. They were then presented two original passages and two accelerated passages, each followed by the associated comprehension questions. To ensure that all recordings represented first exposures, teachers were instructed to never repeat a passage and to move on to the next passage if any problems occurred during the recording process. Recordings were analyzed automatically by the system and reviewed by members of the research team to ensure their adequacy. In addition, research personnel were sometimes present at recording sessions to assess fidelity of implementation. All recordings were collected between January and March of 2011.

Recording Procedure in 1960

In Taylor et al.'s (1960) study, associate researchers were recruited at participating colleges and universities. They received Reading Eye cameras, instructional materials, and training from representatives of Educational Development Laboratories. The associates then collected data at schools between February and March of 1959. As in the present study, students first completed a practice trial. They were then presented with a single passage and comprehension questions. If comprehension was below 70%, a second passage was presented. The associates analyzed the recordings and returned the filmstrips and results to the lead investigators. The research team selected a random sample of 50% of the recordings to check accuracy.

Data Coding

Both studies (1960 and 2011) used identical coding procedures. Each passage for grades 4 and up contained 12 lines and about 120 words. However, the first and last lines were excluded from the analysis to avoid any potential atypical behaviors at the start or completion of a passage. The grade 2 passages only contained seven middle lines (50 words), and the performance scores were doubled to simulate a 100-word reading. Visagraph software automatically calculated the performance measures reported in the present study, whereas in the 1960 study, these calculations were completed manually.

Data Analysis

Focusing first on the 2011 data, analyses of variance (ANOVAs) were used to evaluate increases in comprehension-based silent reading rate and efficiency measures across grades for the original and the accelerated texts. Levene's test was used to test for homogeneity of variance. The Kruskal–Wallis test was used when variances were not homogeneous. Post hoc comparisons with Bonferroni or Games–Howell adjustment were used as appropriate.

Taylor et al. (1960) reported only means and the number of observations for each grade group; no individual data sets or measures of variability were available. To use the 1960 data in the comparison analyses, a data matrix was constructed using the mixedDesign function in the R environment for statistical computing (R Core Team, 2014), following a procedure based on Kliegl (2014). For this purpose, standard deviations for the 1960 data set were estimated on the basis of the 2011 data at each grade level. Due to the much larger sample size of the 1960 data, this estimation can be viewed as conservative (i.e., likely to overestimate the 1960 variability).

To evaluate differences in the development of comprehension-based silent reading rate and efficiency across years (1960 vs. 2011) and grades (2, 4, 6, 8, and 11/12), as well as interactions between these factors, data were analyzed with linear models. The models were fitted using generalized least squares, allowing errors to be correlated and/or have unequal variances. This was achieved using the gls function in combination with the varIdent function from the nlme package (Pinheiro & Bates, 2000) in R. In the models, YEAR (1960 vs. 2011) and GRADE (2, 4, 6, 8, and 11/12) were specified as fixed factors, and successive difference contrasts (Venables & Ripley, 2002) were used to evaluate differences in reading efficiency development from grade to grade and between years. In the model utilized, the varIdent function allows different variances, one for each level of a factor, safeguarding against violations of homogeneity of variance. All of the comparisons were orthogonal, a priori, and within the allowable degrees of freedom offered by the design. The inferential statistics reported are based on the actual results from the analysis. Because differences between 2011 and 1960 at each grade level (2, 4, 6, 8, and 11/12) were not directly tested in the model, a second model was fitted using contrasts specifying these differences. The Benjamini–Hochberg procedure was used to control for the false discovery rate when multiple comparisons were made (Benjamini & Hochberg, 1995): All comparison contrasts were rank ordered by p-values and compared with $(i/m)Q$, where i = rank, m = number of comparisons, and $Q = 0.05$ (false discovery rate).

Results

Passage Comparability

Recall that in the 2011 study, each student read four passages: two from Taylor et al.'s (1960) study and two that had been revised to meet the accelerated Lexile levels of the Common Core. In total, 7,342 valid recordings were obtained (an average of 3.33 recordings per student), of which 52.6% were original passages and 47.4% accelerated passages. The final data set did not include four recordings for every student for one or more of several reasons: (a) Students failed to complete all four recordings, (b) teachers encountered technological problems during a recording, (c) a recording did not save properly, or (d) the system could not process a recording.

Students demonstrated at least 70% comprehension on 5,043 (68.7%) of the 7,342 recordings, of which 49.5% were original passages and 50.5% accelerated passages. The number of students who completed zero to four qualifying recordings at each grade level is shown in Table 2. Of these students, 67.9% had one or more qualifying recordings of each passage type, whereas 16.3% only had qualifying recordings on original passages and 15.8% only on accelerated passages.

Analyses of Data From the 2011 Cohort

An ANOVA examining reading rates across grades for the 2011 cohort demonstrated a significant main effect of grade, $F(5, 2003) = 100.27$, $p < .001$, partial $\eta^2 = .20$. Levene's test indicated, however, that the homogeneity of variance assumption was not met, $F(5, 2003) = 8.38$, $p < .001$. This was because variance in reading rate increased as students advanced toward the upper grades. Nevertheless, a Kruskal–Wallis test confirmed the results of the ANOVA, indicating that reading rates were not the same across grades, $p < .001$. Post hoc grade-to-grade comparisons with Games–Howell adjustment indicated that reading rate significantly increased as students advanced toward the upper grades ($p < .001$) in all but two comparisons: grades 6 to 8 and grades 10 to 12 (see Figure 4).

Differences in reading rate by gender were very small but consistently favored males (males: 166 wpm, females: 159 wpm), $F(1, 1{,}875) = 8.17$, $p < .004$, partial $\eta^2 = .004$. In post hoc comparisons examining gender differences within grades, only the grade 6 comparison reached significance (males: 173 wpm, females: 160 wpm), $F(1, 269) = 5.23$, $p = .023$, and the effect size was quite small (partial $\eta^2 = .019$). This stands in contrast to the gender results reported in the 1960 study (Taylor et al., 1960), where small but consistent differences in reading rate favored females.

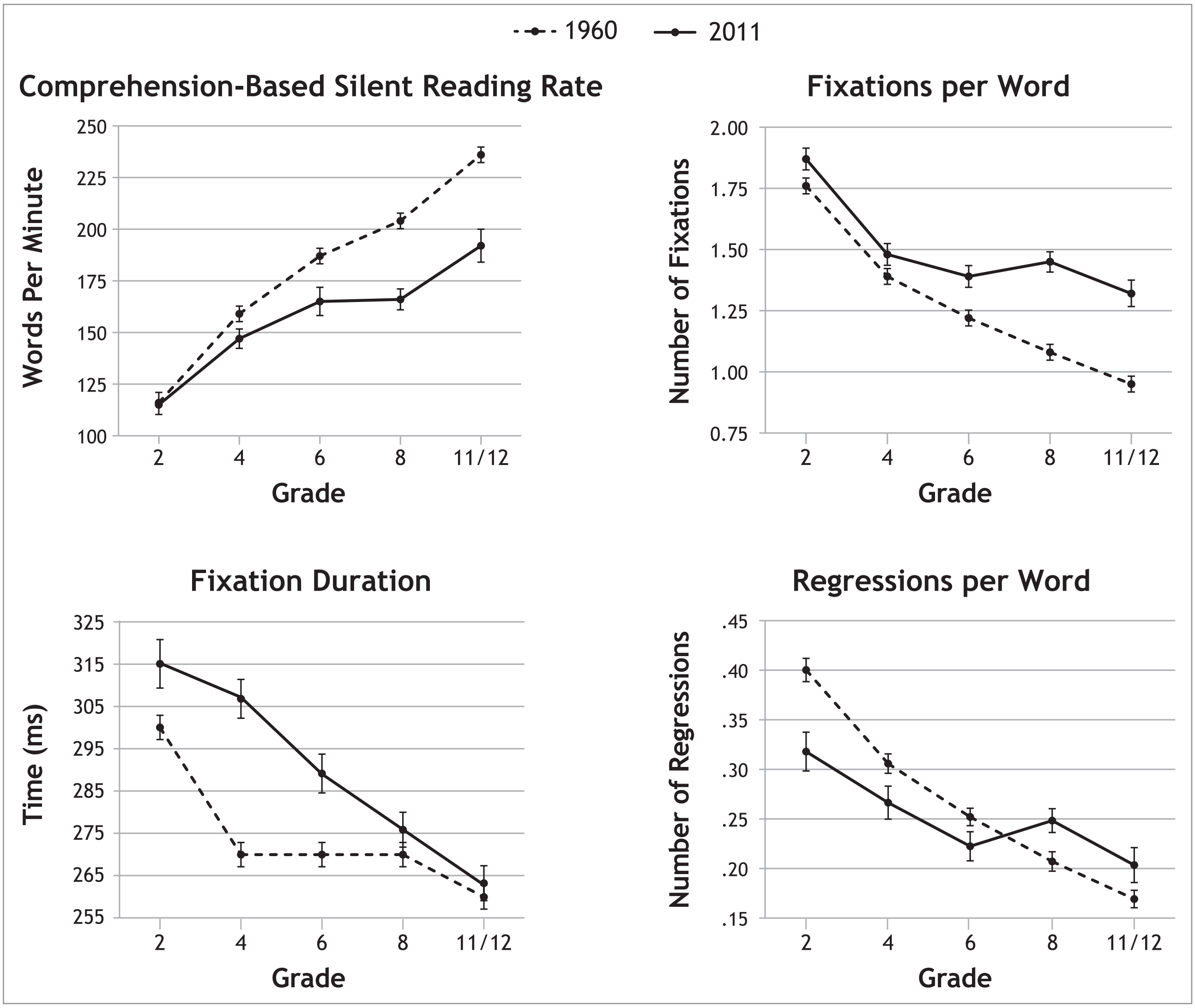
To address a potential source of bias created by students in 2011 having four opportunities to achieve a qualifying recording rather than the two chances afforded to students in 1960 (Taylor et al., 1960), ANOVAs were used to compare the reading performance of students who attained a qualifying passage on an original passage (and thus would have been included in Taylor et al.'s study) with the performance of students who only had qualifying recordings on an accelerated passage (and thus would not have been included in the 1960 study). These analyses showed that there were no significant differences between students in these two groups on any reading performance measure at any grade level. Consequently, all students who achieved

TABLE 2
Percentage of Students Completing Zero to Four Passages With at Least 70% Comprehension, and Final Student Sample

	Initial student sample							Final student sample	
		Percentage of students with qualifying recordings							
Grade	*n*	Zero passages	One passage	Two passages	Three passages	Four passages		*n*	Percentage of initial sample with qualifying performances
2	414	8.5	15.5	21.3	23.2	31.6		379	92
4	451	15.1	28.4	25.3	20.6	10.6		383	85
6	316	7.0	15.8	27.5	27.8	21.8		294	93
8	519	7.7	20.2	28.7	23.9	19.5		479	92
10	265	5.3	18.9	30.2	29.8	15.8		251	95
12	238	6.3	14.3	23.1	31.1	25.2		223	94
All	2,203	8.8	19.6	26.0	25.1	20.5		2,009	91

Note. Regardless of the number of qualifying recordings a student had initially, only a single mean performance measure for each student was used in the efficiency analyses.

FIGURE 4
Means and Standard Errors for Reading Efficiency Measures in 1960[a] and 2011, by Grade



Note. The 1960 standard errors are estimates.
[a]Taylor, S.E., Frackenpohl, H., & Pettee, J.L. (1960). Grade level norms for the components of the fundamental reading skill. *EDL Research and Information Bulletin*, *3*, 22.

qualifying recordings on either original or accelerated passages (or both) were included in the 2011 sample.

Next, a comparison of students' reading performance on the two sets of 2011 passages (i.e., original and Common Core–accelerated texts) was performed. ANOVAs indicated that students read both passage types at essentially identical rates, as shown in Table 3. ANOVAs examining passage effects on fixations per word, fixation duration, and regressions per word at each grade level also showed no significant effects of passage type.

Also shown in Table 3 are the mean comprehension scores of students who completed one or more of the original and/or accelerated passages with at least 70% comprehension and the comprehension score differences between these passage types. These differences and the corresponding effect sizes were small but statistically significant in four grades, where mean comprehension scores were between 2.1% and 4.4% higher on the accelerated passages (partial η^2 ranged from .013 to .056). This higher performance on Common Core–accelerated passages may represent a practice effect because these passages were always presented after the original passages.

Given the nearly identical reading rates and consistent comprehension performance on the original and accelerated passages, a single mean performance score was established for each student based on performances on all texts for which the 70% comprehension standard was met. Numerically, this meant that the performance

TABLE 3
Reading Rates and Comprehension Scores on Original Versus Accelerated Passages

	Original passages			Accelerated passages					
Grade	*n*	Mean	Standard deviation	*n*	Mean	Standard deviation	Difference	*F*	*p*
Reading rate (words per minute)									
2	336	116.1	41.7	317	115.1	40.8	−1.0	0.10	.757
4	309	147.8	45.4	292	148.6	50.3	0.8	0.04	.845
6	242	164.2	54.4	264	164.1	49.8	−0.2	0	.972
8	386	169.9	51.8	413	164.5	51.1	−5.4	2.24	.135
10	227	186.6	53.4	192	177.3	53.9	−9.3	3.10	.079
12	191	187.5	55.5	203	197.3	66.3	9.8	2.50	.114
Comprehension scores (percentage)									
2	336	88.5	9.6	317	87.6	9.2	−0.9	1.49	.222
4	309	79.8	9.1	292	84.3	9.7	4.4	33.53	<.001
6	242	78.9	8.4	264	83.1	8.5	4.1	29.93	<.001
8	386	78.5	8.3	413	80.9	8.3	2.4	17.15	<.001
10	227	81.8	8.5	192	82.1	8.6	0.4	0.17	.676
12	191	81.7	9.1	203	83.8	9.3	2.1	5.08	.025

Note. The upper panel shows the mean reading rates of students who completed either one or both of the passages in one or both passage sets (original and/or accelerated) with at least 70% comprehension. The lower panel shows the mean comprehension scores that students achieved on these qualifying passages. The *n*s represent the number of students who had at least one qualifying passage within the passage group indicated (original or accelerated).

measures of 22% of students were based on four passages, 28% on three passages, 29% on two passages, and 21% on a single passage.

Reading Across the Decades: 2011 Versus 1960

Equivalence of Samples

The sample sizes in the present study and in Taylor et al.'s (1960) study are shown in Table 4. Participants in both studies were intended to represent a cross section of the national population at their respective times. Comparisons between the results of these studies must be made with caution, however, considering the well-documented changes in the demographic character of American society between 1960 and 2011. Further, it cannot be confirmed whether both study samples were equally representative of the national population during the era the data were collected. As noted earlier, the 2011 sample included participants from all regions of the United States but differed somewhat from national norms in having a lower percentage of nonwhite and Hispanic students (40% vs. 46% nationally). The 2011 sample also did not include English learners or students receiving special education services (it is not possible to confirm whether students in these two categories were excluded in the 1960 sample). These differences suggest an overrepresentation in the 2011 sample of students with demographic and socioeconomic characteristics typically associated with higher reading achievement scores. Consistent with this assumption, 67% of the students in the 2011 sample for which performance data were available tested in the top two quartiles on their norm-referenced state reading/language arts assessments, and only 5% tested in the lowest quartile.

Data for Taylor et al.'s (1960) sample were collected in 19 states, representing all regions of the United States and meant to represent the average socioeconomic level for each area. As was often the case in older studies, however, specific demographic and socioeconomic characteristics were not reported. With respect to race and ethnicity, the 1960 sample may have included a proportionally representative sample of the student population, but it cannot be assumed that the educational environment would have been equivalent due to the de facto segregation that was pervasive in the country at that time. However, Taylor's (1965; Taylor et al., 1960) findings were replicated in several later studies and shown to provide a good estimate of comprehension-based silent reading rates across grades (see Carver,

TABLE 4
Distribution of Sample Sizes: 1960[a] and 2011

Grade	2	4	6	8	10	11	12	Total
1960	1,185	1,453	1,636	1,117	—	1,038	—	6,429
2011	379	383	294	479	251	—	223	2,009

Note. Shown here are the numbers of students from which qualifying recordings were obtained. In the 1960 study, no students were tested in grades 10 and 12 (values shown in that study were extrapolated from adjacent grades). In the 2011 study, no students were tested in grade 11.
[a]Taylor, S.E., Frackenpohl, H., & Pettee, J.L. (1960). Grade level norms for the components of the fundamental reading skill. *EDL Research and Information Bulletin*, *3*, 22.

1989). Further, students in both the 1960 and the 2011 samples had nearly identical reading rates in grade 2 and were still fairly close in grade 4. The numbers of fixations per word were fairly similar in grades 2 and 4 as well. The comparable performances of students in the early grades in 1960 and 2011 suggest that the differences in the later grades should be viewed as representing genuine differences in student performance.

Comprehension-Based Silent Reading Rate

Figure 4 (upper left) shows the mean reading rates in the 2011 and 1960 samples for grades in which comparable data were collected. Successive difference contrasts evaluated reading rate development from grade to grade (GRADE) and between studies (YEAR), as well as interactions between these factors. The summary appears in Table 5 (upper left). Estimates of the successive differences between grades (with both years combined) are shown initially. The associated inferential statistics reflect the probability that a given estimate is significantly different from zero. Significant main effects were found for GRADE across all comparisons, indicating reliable reading rate increases from grade to grade (all contrasts, $p < .001$) with effect sizes representing grade to grade increases of between 9 and 38 wpm. A significant main effect was also found for YEAR: Overall, reading rates (with all grades combined) were about 24 wpm lower in 2011 than in 1960 ($p < .001$).

Interactions between GRADE and YEAR were significant for the grade 2 versus grade 4, grade 4 versus grade 6, and grade 6 versus grade 8 comparisons, indicating differences in reading rate growth between successive grades in 2011 versus 1960. These interactions are apparent in Figure 4, which shows that grade 2 students in both samples had a reading rate of 115 wpm, but subsequent grade-to-grade increases in reading rate were nearly always larger in the 1960 sample. This difference in growth was especially striking in middle school, where increases in reading rates between grades 6 and 8 averaged 17 wpm in 1960 but just 1 wpm in 2011.

Differences in reading rate between 1960 and 2011 at each grade level (2, 4, 6, 8, and 11/12) were evaluated using an additional linear model specifying these contrasts. Table 6 presents the inferential statistics, means, standard deviations, and 95% confidence intervals for each grade. The results indicate significant differences between 1960 and 2011 for all grades except grade 2.

To consider the development of reading rates among students with different levels of reading efficiency, participants in the 2011 study were divided into four quartile groups at each grade level based on their silent reading rates. The results are plotted in Figure 5. A linear model was used to compare changes in reading rate across grades within each quartile. When evaluating these interactions, only orthogonal comparisons were made (i.e., comparisons of adjacent quartiles). In addition to the expected main effects of QUARTILE and GRADE, there were four significant QUARTILE-by-GRADE interactions, indicating differences in reading rate growth between adjacent quartiles. These can reasonably be regarded as points at which growth in reading rate in upper quartiles diverges from that occurring in lower quartiles. The first of these interactions indicates that reading rate increases from grade 2 to grade 4 lagged behind in the lowest quartile as compared with the second quartile (by 9.1 wpm, $p < .001$). Two additional interactions indicated that reading rate increases from grade 4 to grade 6 and from grade 10 to grade 12 were more pronounced in the third versus the second quartile (by 3.8 wpm, $p = .039$, and 9.0 wpm, $p < .001$, respectively). The fourth interaction indicated that reading rate increases between grades 8 and 10 were more pronounced in the fourth (the highest) quartile as compared with the third quartile (by 13.4 wpm, $p = .049$).

Fixations per Word

Figure 4 (upper right) shows the mean numbers of fixations per word in the 2011 and 1960 samples for grades with comparable data. Successive difference contrasts were used to evaluate changes from grade to grade (GRADE) and between studies (YEAR), as well as interactions. The output summary is shown in Table 5 (upper right). Displayed first are estimates of the successive differences between grades (with both years combined). The associated inferential statistics reflect the probability that a given estimate is significantly different from zero. As would be expected, there were

TABLE 5
Differences in Comprehension-Based Silent Reading Efficiency Measures: Grades and Years

	Difference in reading rate (wpm)				Difference in fixations per word			
Difference	Estimate	Standard error	*t*	*p*	Estimate	Standard error	*t*	*p*
Grade comparisons								
Intercept	168.8	0.67	252.4	<.001	1.39	0.005	268.5	<.001
4 vs. 2	37.9	1.83	20.7	<.001	−0.38	0.017	−22.7	<.001
6 vs. 4	22.8	2.12	10.8	<.001	−0.13	0.015	−8.5	<.001
8 vs. 6	8.9	2.07	4.3	<.001	−0.04	0.015	−2.5	.011
11/12 vs. 8	29.2	2.51	11.7	<.001	−0.13	0.016	−8.2	<.001
2011 vs. 1960[a]	−23.5	1.39	−16.8	<.001	0.22	0.010	21.3	<.001
Grade × Year interactions: 2011 versus 1960[a]								
4 vs. 2	−10.7	3.66	−2.9	.004	−0.03	0.034	−0.8	n.s.
6 vs. 4	−10.3	4.24	−2.4	.015	0.08	0.030	2.7	.006
8 vs. 6	−16.4	4.13	−4.0	<.001	0.20	0.030	6.6	<.001
11/12 vs. 8	−5.6	5.01	−1.1	n.s.	0.01	0.033	−0.1	n.s.

	Difference in fixation duration (ms)				Difference in regressions per word			
Difference	Estimate	Standard error	*t*	*p*	Estimate	Standard error	*t*	*p*
Grade comparisons								
Intercept	282	0.67	422.6	<.001	0.265	0.0020	130.3	<.001
4 vs. 2	−19	2.36	−8.0	<.001	−0.074	0.0069	−10.7	<.001
6 vs. 4	−9	2.05	−4.5	<.001	−0.049	0.0058	−8.5	<.001
8 vs. 6	−6	1.96	−3.1	.002	−0.008	0.0058	−1.3	n.s.
11/12 vs. 8	−12	1.93	−6.0	<.001	−0.040	0.0065	−6.2	<.001
2011 vs. 1960[a]	16	1.33	11.9	<.001	−0.012	0.0041	−3.0	.003
Grade × Year interactions: 2011 versus 1960[a]								
4 vs. 2	22	4.72	4.7	<.001	0.037	0.0138	2.6	.008
6 vs. 4	−19	4.09	−4.5	<.001	0.015	0.0117	1.3	n.s.
8 vs. 6	−12	3.93	−3.1	.002	0.066	0.0116	5.7	<.001
11/12 vs. 8	−3	3.86	−0.8	n.s.	−0.005	0.0130	−0.4	n.s.

Note. n.s. = not significant; wpm = words per minute. In each quadrant, the upper panel shows linear model estimates of differences in an efficiency measure between grades (1960 and 2011 combined). The associated statistics reflect the probability that a given estimate is significantly different from zero. Using reading rate as an example, the difference in reading rate between grades 2 and 4 was estimated to be 37.9 wpm; that is, reading rate in grade 4 was 37.9 wpm faster than in grade 2, and this increase is significant. Just below this is a comparison of differences in reading rate between 2011 and 1960 with all grades combined. The lower panel shows estimates for the interaction terms that reflect differences in the increase between neighboring grades in 2011 versus 1960. For example, the increase in reading rate from grade 6 to grade 8 in 2011 was smaller than the corresponding increase in 1960 (−16.4 wpm).
[a]Taylor, S.E., Frackenpohl, H., & Pettee, J.L. (1960). Grade level norms for the components of the fundamental reading skill. *EDL Research and Information Bulletin*, *3*, 22.

significant main effects for GRADE across all comparisons, indicating reliable decreases in fixations per word from grade to grade (all contrasts, $p < .02$). There was also a significant main effect for YEAR, with students in 2011 making 0.22 more fixations per word than those studied in 1960 (mean = 1.52 in 2011 vs. 1.29 in 1960).

GRADE-by-YEAR interactions were also significant in the middle grades. Both samples showed a comparable reduction in the number of fixations per word

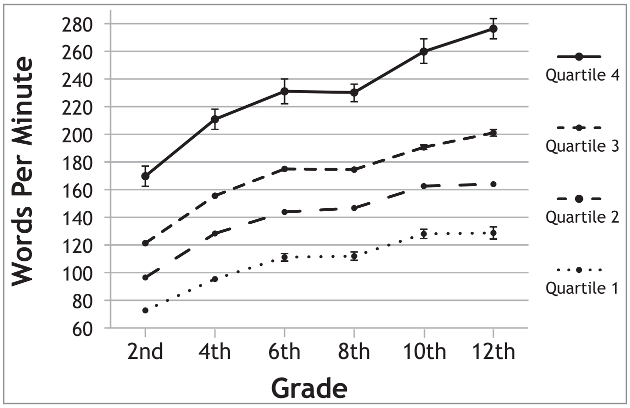
TABLE 6
Comprehension-Based Silent Reading Efficiency Measures Across Grades: 2011 and 1960[a]

Grade	2011 Mean	2011 Standard deviation	2011 95% confidence interval	1960 reported mean	1960 reconstructed Mean	1960 reconstructed Standard deviation	2011 vs. 1960 reconstructed t	2011 vs. 1960 reconstructed p	2011 vs. 1960 reconstructed Hedges' g
Reading rates (words per minute)									
2	115	40	[110, 120]	115	116	39	0.4	n.s.	—
4	147	49	[143, 152]	158	159	49	4.2	<.001	0.24
6	165	50	[159, 171]	185	187	50	7.2	<.001	0.44
8	166	48	[161, 170]	204	204	48	14.6	<.001	0.80
10	185	54	[178, 192]	224[b]	—	—	—	—	—
11/12	192	59	[186, 198]	237	236	53	12.4	<.001	0.81
Fixations per word									
2	1.87	0.46	[1.83, 1.91]	1.74	1.76	0.46	−5.3	<.001	0.26
4	1.48	0.37	[1.44, 1.52]	1.39	1.39	0.37	−4.1	<.001	0.24
6	1.39	0.34	[1.35, 1.43]	1.20	1.22	0.34	−7.2	<.001	0.51
8	1.45	0.38	[1.42, 1.48]	1.09	1.08	0.37	−17.9	<.001	0.99
10	1.32	0.31	[1.28, 1.36]	1.01[b]	—	—	—	—	—
11/12	1.32	0.36	[1.27, 1.37]	0.96	0.95	0.31	−13.4	<.001	1.17
Fixation durations (ms)									
2	315	61	[310, 320]	300	300	60	−5.3	<.001	0.25
4	307	54	[302, 312]	270	270	50	−13.4	<.001	0.73
6	289	44	[283, 294]	270	270	40	−6.1	<.001	0.46
8	276	53	[272, 280]	270	270	50	−2.4	.019	0.12
10	268	33	[264, 272]	260[b]	—	—	—	—	—
11/12	263	37	[257, 269]	260	260	30	0.9	n.s.	—
Regressions per word									
2	0.33	0.18	[0.31, 0.35]	0.40	0.40	0.18	8.7	<.001	0.41
4	0.27	0.15	[0.26, 0.29]	0.31	0.31	0.15	4.6	<.001	0.26
6	0.23	0.12	[0.22, 0.24]	0.25	0.25	0.12	2.6	.009	0.19
8	0.25	0.16	[0.24, 0.27]	0.21	0.21	0.15	−5.2	<.001	0.27
10	0.20	0.11	[0.18, 0.21]	0.19[b]	—	—	—	—	—
11/12	0.21	0.14	[0.19, 0.23]	0.18	0.18	0.11	−3.4	.001	0.33

Note. n.s. = not significant. Data shown for 1960 include the actual means reported in Taylor's 1965 report[c] and the reconstructed statistics for 1960. Also shown are the inferential statistics generated by the linear analysis for comparisons between 2011 and 1960 (reconstructed), and the calculated effect sizes (Hedges' *g*). Grade 12 data in 2011 were compared with grade 11 in 1960.
[a]Taylor, S.E., Frackenpohl, H., & Pettee, J.L. (1960). Grade level norms for the components of the fundamental reading skill. *EDL Research and Information Bulletin*, *3*, 22. [b]Extrapolated value, not reconstructed and compared. [c]Taylor, S.E. (1965). Eye movements in reading: Facts and fallacies. *American Educational Research Journal*, *2*(4), 187–202.

from grade 2 to grade 4 and from grade 8 to grades 11/12. In the 2011 sample, however, the reduction was smaller between grades 4 and 6 (p = .006). Further, the number of fixations per word increased between grades 6 and 8 (p < .001). Finally, because the total number of fixations per word includes fixations following both progressive and regressive eye movements, the number of forward fixations was examined separately. This

FIGURE 5
Means and Standard Errors for Comprehension-Based Silent Reading Rates in 2011, by Grade, for Each Reading Rate Quartile



Note. Standard error bars are too small to represent in quartiles 1–3.

analysis yielded the same pattern of main effects and interactions as found for the total number of fixations.

Differences in number of fixations per word between 1960 and 2011 at each grade level (2, 4, 6, 8, and 11/12) were evaluated using an additional linear model specifying these contrasts. Table 6 presents the inferential statistics, means, standard deviations, and 95% confidence intervals for each grade level. The average number of fixations per word was significantly higher at all grade levels in 2011 in comparison with the 1960 sample ($p < .001$).

Fixation Duration

Figure 4 (lower left) shows mean fixation durations in the 2011 and 1960 samples for grades with comparable data. Successive difference contrasts were used to evaluate changes from grade to grade and between years, as well as interactions between these factors. The output summary is shown in Table 5 (lower left). Displayed first are estimates of the successive differences between grades (with both years combined). The associated inferential statistics reflect the probability that a given estimate is significantly different from zero. Main effects for GRADE were significant across all comparisons ($p < .01$), indicating reliable decreases in fixation duration from grade to grade. A significant main effect was also found for YEAR, indicating that fixation durations were 16 ms longer in 2011 relative to 1960 ($p < .001$).

GRADE-by-YEAR interactions were also significant in the lower grades, reflecting a large decrease in fixation durations between grades 2 and 4 in 1960, followed by little change through grade 8. In the 2011 sample, the decrease occurred more gradually over the elementary and middle school grades. Despite these differences in the time course, fixation durations decreased to a similar extent between grades 8 and 11/12 in both samples.

Differences in fixation duration between 1960 and 2011 at each grade level (2, 4, 6, 8, and 11/12) were evaluated using a linear model specifying these contrasts. Table 6 presents the inferential statistics, means, standard deviations, and 95% confidence intervals for each grade level. Fixation duration means for grades 2, 4, and 6 were significantly longer in 2011 compared with the 1960 sample ($p < .001$), but there were no significant differences between samples in fixation durations in the upper grades (8 and 11/12).

Regressions per Word

Figure 4 (lower right) depicts the mean number of regressions per word across grades in the 2011 and 1960 samples for grades in which comparable data were collected. Successive difference contrasts were used to evaluate changes from grade to grade and between years, as well as interactions. The output summary is shown in Table 5 (lower right). Depicted first are estimates of the successive differences between grades (with both years combined). The associated inferential statistics reflect the probability that a given estimate is significantly different from zero. There was a significant main effect of GRADE in four comparisons ($p < .001$), with the number of regressions per word decreasing between grades in all cases except the comparison between grades 6 and 8 (where the decrease in 1960 was offset by an increase in 2011). There was also a significant main effect for YEAR, with students in 2011 making 3.7% fewer regressions per word than students in the 1960 study ($p = .003$).

GRADE-by-YEAR interactions were significant in two comparisons: (a) between grades 2 and 4, due to a somewhat larger decrease in regressions per word in 1960 ($p = .008$), and (b) between grades 6 and 8, due to a continuing decrease in the number of regressions in 1960 contrasted to an increase in 2011 ($p < .001$).

Differences in the number of regressions per word between 1960 and 2011 at each grade level (2, 4, 6, 8, and 11/12) were evaluated using an additional linear model specifying these contrasts. Table 6 presents the inferential statistics, means, standard deviations, and 95% confidence intervals for each grade. The number of regressions per word in 2011 was smaller in each comparison up to grade 6 but larger in grades 8 and 11/12 ($p < .001$ for all comparisons).

Discussion

The present research constitutes the first comprehensive study of the comprehension-based silent reading efficiency of U.S. students in the 21st century, the last having been conducted in 1960 (Taylor, 1965). Although additional studies of silent reading rates have been published since that time (Carver, 1983; Gallo, 1972; Vorstius et al., 2014), these are less useful for comparative purposes due to their focus on a limited number of grades, not having used grade-leveled texts, and/or not requiring a comprehension standard (see Carver, 1989). The current study was designed to overcome these limitations while also taking advantage of advances in eye-tracking technology.

Trends in Reading Efficiency

The results of the current study suggest that present-day students are less efficient readers than their 1960 counterparts. The data on reading rate and eye movement patterns suggest the strong likelihood of a decline in word recognition automaticity.

Comprehension-Based Silent Reading Rate

Differences in reading rate between the 2011 and 1960 samples were characterized by a gradual divergence in average reading rates over grades. Grade 2 students in the two studies were comparable, but by grade 4, students in the 2011 sample were reading about 12 wpm more slowly than those in the 1960 sample. By grade 6, this difference nearly doubled, and by grade 8, it had more than tripled to a 38 wpm deficit. By grade 12, students in 2011 were reading at an average rate of 192 wpm, 45 wpm more slowly than grade 11 students in 1960. This represents a 19% decline in comprehension-based silent reading rate.

The analysis of reading rate growth for 2011 students with differing levels of reading efficiency (i.e., the lowest to highest quartiles for reading rate) also showed several interesting patterns. First, the reading rate increase between grades 2 and 4 was significantly smaller in the lowest quartile relative to the upper quartiles, suggesting that these students are still developing decoding skills, vocabulary, and reading stamina. Second, growth in reading rate was stalled in all quartiles between grades 6 and 8. Given the uniformity of this pattern across all efficiency groups, further investigation of this phenomenon is warranted. Challenges associated with transitioning from elementary to middle school may be a contributing factor (e.g., Schwerdt & West, 2013). Third, reading rate development in students in the lower quartiles of the 2011 sample progressed slowly (if at all) over the high school years. Between grades 10 and 12, for example, students in the two highest quartiles increased their reading rate by an average of 13 wpm (reaching 201 and 276 wpm, respectively), whereas those in the two lowest quartiles increased their reading rate by an average of only 1 wpm (reaching just 129 and 164 wpm, respectively). These latter students will end their high school careers with reading rates that are well below or at best comparable to typical conversational speaking rates in English (150–170 wpm; Yuan, Liberman, & Cieri, 2006). These low reading rates likely reflect such low levels of automaticity that comprehension and memory are taxed during reading tasks. When the process of decoding words is laborious and time consuming, short-term memory may begin to dispose of information before it can be integrated into the evolving text base and situation model (Hudson, Lane, & Pullen, 2005; Kintsch, 1998; LaBerge & Samuels, 1974; Perfetti, 1985).

Fixations per Word

The analysis of fixations per word also revealed a pattern of declining reading efficiency between 1960 and 2011. Less skilled readers frequently resort to the analysis of sublexical units (e.g., sounding out words), thereby requiring multiple fixations to identify a given word. More capable readers tend to identify common words in a single fixation and even skip shorter words that are predictable from the context, only resorting to sublexical analysis for less familiar and low-frequency words (Ashby et al., 2005; Joseph, Nation, & Liversedge, 2013; Samuels, LaBerge, & Bremer, 1978). Consistent with a pattern of increasing efficiency, students in 1960 averaged 1.76 fixations per word in grade 2, followed by a steady reduction in the number of fixations per word through grade 12, when students averaged 0.95 fixations per word (i.e., less than one fixation per word).

Students in 2011 did not achieve such efficiency. Students in grades 2 and 4 were roughly comparable to their 1960 counterparts, but reductions in fixations per word were significantly smaller from grades 4 to 6 and were followed by an increase in fixations per word between grades 6 and 8, coinciding with the middle school period during which growth in reading rate was stalled. By grade 12, students in 2011 were averaging 1.32 fixations per word, a level of efficiency not substantially different from grade 6 students in the same year (1.39 fixations per word) and nearly 40% less efficient than grade 11 students in the 1960 study. If a large number of fixations per word is regarded as an indication of more serial word processing with a focus on sublexical units (e.g., Ablinger, Huber & Radach, 2014), an interpretation of the 2011 results is that many students in grade 6 and beyond continue to engage in the same sort of sequential decoding strategy that younger readers typically use.

Fixation Duration

Longer fixations are characteristic of less skilled readers and reflect, for example, less familiarity with vocabulary and an inability to use textual context (Joseph et al., 2013; Williams & Morris, 2004). Students in 1960 and 2011 had similar fixation durations by the end of high school but followed different developmental paths. In the 1960 cohort, the length of fixation duration decreased substantially from grade 2 to grade 4. After this point, changes were not substantial until grade 11. Further, although the fixation duration remained more or less constant during this period, the number of fixations per word was steadily decreasing. Such a shift is consistent with holistic lexical processing (i.e., taking in more text per fixation). In contrast, becoming a more efficient reader in 2011 took the form of a steady reduction in fixation duration, whereas the number of fixations per word remained more or less the same (i.e., taking in the same amount of text per fixation but doing so in a shorter period of time). In other words, students in 2011 appeared to sustain a less efficient, predominantly sublexical word-processing strategy for a much longer time than was the case with students in the 1960 sample.

Regressions per Word

With respect to number of regressions (right-to-left saccades), students in 2011 made 3.7% fewer regressions per word than their 1960 counterparts, but this overall effect is complicated by the patterns of regressions across the grades. Students in 2011 made fewer regressions per word through grade 6, but beyond this point, they made more regressions per word than their 1960 counterparts. In other words, students in the upper grades in 2011 not only required more fixations per word but also more frequently went back and refixated on words. Taking regressions as an indicator of processing difficulty (e.g., Reichle, Rayner, & Pollatsek, 2003), these observations support the conclusion that many present-day students maintain a sublexical processing strategy well into the middle school grades and beyond.

Possible Explanations

The present results suggest that typical high school seniors in 2011 read more slowly than their counterparts in 1960 but also read less efficiently, persisting in word identification and systematic decoding of text rather than reading holistically and with automaticity. Additional research will be required to test causal explanations as to why students are less efficient now than in the past. The literature, however, suggests plausible contributing factors, including a less than optimally calibrated approach to text complexity and insufficient silent reading practice.

Text Complexity

Since a decline in verbal SAT scores was reported in the early 1970s, decreases in the complexity of textbooks have been offered as a potential explanation for the poorer reading performances of high school graduates (Chall, Conard, & Harris, 1977; Hayes, Wolfer, & Wolfe, 1996). This same explanation for poor performances by the current generation of high school students (ACT, 2006) was raised by the Common Core writers, who stated, "Despite steady or growing reading demands from various sources, K–12 reading texts have actually trended downward in difficulty in the last half century" (NGA Center & CCSSO, 2010b, p. 3). They cited Chall et al.'s (1977) and Hayes et al.'s (1996) studies as evidence for this downward trend. Williamson (2008) analyzed 75 texts used in grades 11 and 12 and a representative sample of texts that individuals are likely to encounter in college and career settings, including military and government publications. The level of high school texts was significantly less challenging than texts of the workplace (typically 125L higher), community college (169L higher), and universities (259L higher).

Several recent analyses indicated that texts have decreased in difficulty over time in some grade-level bands but not in others. For example, the typical texts used for first-grade instruction have increased in difficulty, as evidenced by an analysis of the complexity of texts from each of the past seven decades from a continually best-selling first-grade core reading program (Fitzgerald, Elmore, Relyea, Hiebert, & Stenner, 2016). In these data, the overall text complexity measure trended toward more complexity over the seven decades, especially in the last three (i.e., 1990s–2010s). Analyses of reading textbooks for grades 3 and 6 used over the past century show that text complexity (lexical sophistication, lexical diversity, and syntactic complexity) generally decreased during the early part of the last century and then stabilized at a lower level in the 1950s through the 1970s (Gamson, Lu, & Eckert, 2013; Stevens et al., 2015). Since that time, however, the lexical sophistication and diversity of third-grade texts has increased markedly, reaching an apex in the 2000s and by some measures reaching higher levels than at any time previously. Changes in the syntactic complexity of third-grade texts followed a different pattern, declining during the early part of the last century but remaining relatively stable since then. The analyses of sixth-grade texts showed that after declines in the early part of the last century, lexical and syntactic complexity have by most measures remained fairly stable. Save for Williamson's (2008) analysis, there is little work at the high school level. If texts have increased in difficulty through grade 3 but remained static in the middle school years, the patterns of declining performances in the middle grades reported in this study would appear to provide corroboration for the need to accelerate text levels for the middle- and high-level grades, as recommended in the Common Core's staircase of text complexity (NGA Center & CCSSO, 2010b, p. 8).

Finally, a serendipitous piece of evidence on the impact of text complexity emerged from the practice of revising (upward) some of Taylor et al.'s (1960) original passages to meet the Common Core expectations for grade-level complexity. Recall that the passages that were revised to meet these standards did not elicit lower comprehension scores among the 2011 cohort of examinees. This provides at least a small piece of evidence that students can handle more complexity than is commonly provided, at least in a testing situation. Of course, our data do not speak to the issue of the impact of exposing students to a steady diet of more challenging texts over an extended period of time, such as an entire school year. Indeed, the Common Core authors caution that students differ in their reading growth trajectories and "need opportunities to stretch their reading abilities but also to experience the satisfaction and pleasure of easy, fluent reading within them" (NGA Center & CCSSO, 2010b, p. 9). Clearly, multiple factors feed into the optimal calibration of text complexity for each student (Williamson, Fitzgerald, & Stenner, 2013). The present results suggest that additional research in this area would be useful.

Reading Practice

Students who are able to develop strong language and reading skills during the primary grades tend to become engaged and proficient readers, whereas students who continually struggle with text tend to become frustrated and avoid reading (Mol & Bus, 2011; Sparks, Patton, & Murdoch, 2014). The data suggest that many U.S. students fall into the latter category (e.g., Klauda, Wigfield, & Cambria, 2012; Melekoğlu & Wilkerson, 2013; Rideout, 2014). In the 2011 PIRLS, 27% of grade 4 students in the United States reported that they liked reading, whereas nearly half indicated that they only read when they had to (Mullis, Martin, Foy, & Drucker, 2013). Among older students, data from the National Center for Education Statistics (2013b) suggest that one third of 13-year-olds and 45% of 17-year-olds rarely or never engage in recreational reading. These infrequent readers also scored lower on the NAEP when compared with students who engaged in recreational reading once or more weekly. These data are consistent with a large body of evidence documenting the Matthew effect: Students who do not engage in sufficient reading practice fall further and further behind (Sparks et al., 2014; Stanovich, 1986). Against this backdrop, it seems clear that an effective program to improve reading development outcomes must nurture both students' reading skills and their motivation to read (Morgan & Fuchs, 2007).

Why a Decrease in Efficiency Matters

A long-standing national priority has been to increase the academic achievement of U.S. students. At the same time, declining, or at best stagnant, scores on national assessments have prompted a host of educational initiatives, such as formation of the NRP in 1997, the Reading First program under the No Child Left Behind Act (2002), and the Common Core State Standards Initiative (NGA Center & CCSSO, 2010a). Each of these undertakings has placed considerable emphasis on improving reading performance, recognizing that reading volume and content play a key role in developing the vocabulary, fluency, and declarative knowledge that students need to comprehend more complex texts such as those they are likely to encounter in college and career settings (e.g., Allington, 2014; Cunningham & Stanovich, 1998; Schmitt, Jiang, & Grabe, 2011; Sparks et al., 2014; Williamson, 2006).

Considered in this light, the significance of a decline in reading efficiency since 1960 becomes apparent: Evidence suggests that time spent reading in the United States is declining—and especially so among teens and young adults (National Center for Education Statistics, 2013b; National Endowment for the Arts, 2007). Yet, even in a situation in which the amount of time spent reading was held constant, the observed 19% (45 wpm) decline in the comprehension-based silent reading rates of U.S. grade 12 students (in comparison with grade 11 students in 1960) represents a substantial decline in reading volume; slower readers inevitably read less in a given amount of time. Lower reading volume in turn is associated with lower levels of word recognition efficiency (as evidenced in the present study by the 17% increase in regressions and 38% increase in fixations), as well as more sluggish development of vocabulary, fluency, and declarative knowledge. A decline in reading comprehension would also be expected, given evidence for a significant correlation between reading rate and reading comprehension scores (Rasinski et al., 2005).

The academic achievement of U.S. students is lagging behind in relation to other nations. If restoring the country's performance (and ranking) in this regard is a national priority, then reversing declines in reading volume and efficiency would seem to be of special importance.

Limitations

Considerable effort was expended to assure that the data for the present 2011 experiment were collected in as rigorous a manner as possible and in a way that rendered them comparable to the data collected in the 1960 study. There are, however, a number of limitations suggesting that the results should be considered with caution.

Comparability of Samples

A major limitation was the inability to ensure that the samples from the two time periods were equivalent. Students in both studies were selected by the participating school personnel who were asked to follow defined criteria. Yet, fidelity cannot be confirmed; the 1960 study did not report demographic data, and demographic data for the 2011 sample were provided at the student level for only about 60% of the students. For the remaining participants, demographic data could not be associated with individual student records. Moreover, the 1960 study did not include an independent measure of reading ability. For the 2011 study, independent measures of reading ability were insufficient. As such, the possibility of sample bias cannot be ruled out in either cohort. A mitigating factor, however, is the good agreement between the results of the 1960 study and the results of replications that followed in the 1970s and 1980s (e.g., Carver, 1983; Gallo, 1972). Confidence in the results of the 2011 study will also benefit from replications with more complete documentation of demographic data and reading ability.

Technology

An additional limitation concerns differences in the eye movement technology used in the two studies. The 1960 data were analyzed by hand from filmstrips, whereas the 2011 data came from automated software routines. Both procedures shared the aim of accurately measuring reading rate, fixations, fixation durations, and regressions. The possibility cannot be excluded, however, that systematic discrepancies between the two methods may have led to a constant difference in one or more eye movement measures. However, all of the parameters included in this study are relatively simple to quantify using either of the methods described and are unlikely to be subject to systematic distortion. Supporting this view is the fact that measures of reading rate and fixations per word in the 1960 and 2011 cohorts were nearly identical in grade 2 and did not diverge substantially until later grades.

Stimulus Materials

Various methods were used to validate the original passages and questions from Taylor's (1965) study, including the use of readability formulas and item analyses of the comprehension assessments. For the passages used in the 2011 study, level-appropriate Lexile scores were confirmed. No additional item analyses were performed because there was no reason to expect the characteristics of the items to change over the 50-year interval. This practice seemed justified in that only minor adjustments were made to some passages, only small comprehension differences were found across passages, and these differences had no significant effect on reading rate or other measures of reading efficiency.

The use of true/false questions paralleled the procedure in the 1960 study, so there were good reasons to retain those items. However, we readily admit that there are well-known limitations on true/false assessments (for summaries of research on true/false items, see Downing, Baranowski, Grosso, & Norcini, 1995; Frisbie & Becker, 1991), namely, that they tend to tap lower level rather than inferential or critical comprehension.

To maximize comparability across the 1960 and 2011 studies, we decided to administer the original 1960 versions of passages before having students respond to the passages revised to meet Common Core difficulty standards. In doing so, we were aware that this might introduce order effects that might account for the finding that the accelerated passages were no more difficult for students than the original passages. It was decided, however, that maintaining comparability with the 1960 administration was the more important consideration.

Grade-Level Designations

Additional complications were introduced by differences in the grades sampled in the 2011 and 1960 studies. In particular, Taylor et al. (1960) extrapolated data for grade 10 from grades 8 and 11 and for grade 12 from grades 11 and college. The 2011 study did not sample grade 11. As such, it was decided to forgo comparisons at grade 10 and to compare grade 12 in 2011 with grade 11 in 1960 in the successive difference contrasts. This provided a one-year experiential advantage for grade 12 students in the 2011 sample in these comparisons. Even so, the grade 11 students in 1960 were generally more efficient than the grade 12 students in 2011.

Data Reporting

Taylor's (1965) study did not report measures of variance, nor could we find them in any reports or fugitive documents compiled in the 1960s when the work was completed. In order to compare performance across these two time periods, we had to find a way to extrapolate variance estimates for the 1960 sample. We ended up using a procedure that employed the 2011 standard deviations for grade-level performances to estimate the variance in 1960 data sets. This approach provided separate estimates of variation for each eye movement measure within and across different grade levels. This is a conservative approach to estimating variance because of the characteristics of the 2011 sample. We begin with the observation that standard deviations decrease with increasing sample size. Then, we note that the 1960 sample was substantially larger than the 2011 sample. Hence, using the 2011 standard deviation values can be regarded as a conservative approach to estimating variance, only likely to overestimate the variance in the 1960 sample and make it more difficult to find true differences between the samples. Adding to confidence in this procedure is the observation that the 1960 means did not fall within the 95% confidence intervals around the 2011 means where there were significant differences.

Constructs Measured

Finally, as important as these findings about efficiency may be, readers are reminded that any conclusions one might draw about declines in performance should be limited to the construct of silent reading efficiency and not extended to global indicators of reading proficiency or other, more specific indicators, such as critical reading or vocabulary knowledge.

Conclusion

The present research adds to evidence suggesting that the silent reading efficiency of U.S. students, especially older students, is declining, stagnant, or at least inadequate to meet the current literacy challenges faced in schools and the workplace (see, e.g., Adams, 2010; Chall, 1996). It is important to evaluate these declines in reading efficiency in the context of shifting demographics in the United States since 1960. There have, for example, been documented increases in racial and ethnic diversity, single-parent families, and the proportion of students from first- or second-generation U.S. families (e.g., Kena et al., 2015). Additionally, there is a distinct possibility that poverty rather than race or ethnicity may be a key factor: First, there is increasing evidence that the proportion of our public school population of children living in poverty has increased since the 1970s (Annie E. Casey Foundation, 2015) and that the sheer effect of poverty on achievement has been on the rise over the last few decades while the effect of race and ethnicity is declining (Reardon, Valentino, & Shores, 2012). Of course, over the same period, there have been fundamental changes in reading pedagogy and philosophy as well. Given the multiplicity of changes in the social, economic, and educational fabric of American society, the present study cannot distinguish the impact of these and many other factors on the results obtained in reading efficiency across the 50 years from 1960 to 2011.

It also appears to be the case, based on the rationale behind the Common Core State Standards (see Williamson, 2006), that most high school graduates lack adequate reading proficiency and have insufficient experience with the sorts of challenging texts they will face in postsecondary educational and workplace settings. These findings are rendered all the more serious when we consider that the United States ranks well behind many other industrialized nations in terms of the reading performance of 15-year-olds and the number of young adults with postsecondary degrees (Organisation for Economic Co-operation and Development, 2011, 2013). In the face of these shortfalls is the sobering realization that with increasing automation and globalization of the economy, nearly two thirds of the jobs available in 2018 will likely require an associate's or bachelor's degree (Carnevale, Smith, & Strohl, 2010) and, more importantly, more efficient and more critical reading skills than high school graduates currently bring to college and work. This shortcoming presents a dilemma both for employers seeking qualified workers and for the majority of job seekers who may lack the necessary qualifications for those jobs. Effectively addressing the root causes of this dilemma should be a national priority.

Notes

This research was supported by Reading Plus, also known as Taylor Associates Communications, Winooski, Vermont, USA. All work was conducted as part of an individual's regular Reading Plus employment, as part of an individual's contract as a Reading Plus advisor, or voluntarily. Replicating Taylor's 1960 study was one of Spichtig's long-standing goals, which she was able to complete as part of her employment at Reading Plus; she also devoted a great deal of her personal time to this project. Hiebert was a paid consultant for Reading Plus at the time of the study, contributing to its design. Vorstius was not paid by Reading Plus; his voluntary participation was motivated solely by academic interest. Pascoe contributed to the literature review and write-up as part of his duties as an employee in the Reading Plus Research Department. Pearson serves as a paid advisor to Reading Plus on matters of curriculum and assessment; his participation in the research project that led to this article was voluntary. Radach has received grant funding from Reading Plus for work not related to the present study; his participation in the research project that led to this article was voluntary.

We would like to express our heartfelt thanks to the many students and teachers who gave of their time to participate in this research and to Stanford E. Taylor for his insight and guidance during this project. We also gratefully acknowledge the significant contributions of Reading Plus staff and others who assisted in this research: John Ferrara, Angela Zinno, Elias Tousley, Michael Lascelles, Jennifer Marsala, Ginger Irish, Alex Rollins, and Ehsan Ansari.

[1] Eye movement data in Taylor et al.'s (1960) study were collected using a Reading Eye camera (Educational Development Laboratories, Huntington, NY). A student positioned his or her head in a brace with a chin rest. Light was then reflected off each cornea and photographed through lenses positioned just below the text selection. The reflected light was captured on filmstrips that moved at a constant rate so the time, direction, and duration of events could be calculated. Each eye stop appeared as a vertical line on the film, each saccade as a horizontal step, and each return sweep as a longer horizontal line in the reverse direction. Analysis of these recordings yielded the same measures of eye movement as the Visagraph.

[2] Some passages were slightly revised to comply with the accelerated Lexile levels of the Common Core (NGA Center & CCSSO, 2010b). Taylor et al.'s (1960) passages for grades 2 and 4 did not require

revisions to comply with Common Core–accelerated Lexile requirements. At each of grades 6, 8, and 10, one passage required minor revisions. Because level 10 passages had been used for all testing in grade 11 and above in the 1960 study (Taylor, 1965), two original level 10 passages were identified that met the level 12 criteria without requiring adjustment, and two others were adjusted to serve as accelerated level 12 passages. On average, the two accelerated passages in a grade set were 82 Lexile levels higher than the two original passages.

[3] Taylor and colleagues (1960) first tested comprehension questions with a sample of 500 subjects and performed an item analysis on each question. After adjustments, the questions were retested, this time with a cohort of 250 students, where it was found that scores obtained by guessing averaged 56% and those obtained after reading averaged 88%. These results were the basis for 70% being selected as the criterion for adequate comprehension.

REFERENCES

Ablinger, I., Huber, W., & Radach, R. (2014). Eye movement analyses indicate the underlying reading strategy in the recovery of lexical readers. *Aphasiology*, *28*(6), 640–657. doi:10.1080/02687038.2014.894960

ACT. (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City, IA: Author. Retrieved from https://www.act.org/research/policymakers/pdf/reading_report.pdf

Adams, M.J. (2010). Advancing our students' language and literacy: The challenge of complex texts. *American Educator*, *34*(4), 3–11, 53.

Allington, R.L. (2014). How reading volume affects both reading fluency and reading achievement. *International Electronic Journal of Elementary Education*, *7*(1), 13–25.

Annie E. Casey Foundation (2015). *Kids count data book: State trends in child well-being*. Baltimore, MD: Author.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065–1086. doi:10.1080/02724980443000476

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, *57*(1), 289–300.

Blythe, H.I., & Joseph, H.S.S.L. (2011). Children's eye movements during reading. In S.P. Liversedge, I.D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643–662). Oxford, UK: Oxford University Press.

Carnevale, A.P., Smith, N., & Strohl, J. (2010). *Help wanted: Projections of jobs and education requirements through 2018*. Washington, DC: Georgetown University Center on Education and the Workforce. Retrieved from ERIC database. (ED524310)

Carver, R.P. (1977). *Technical manual for the National Reading Standards*. Kansas City, MO: Revrac.

Carver, R.P. (1983). Is reading rate constant or flexible? *Reading Research Quarterly*, *18*(2), 190–215. doi:10.2307/747517

Carver, R.P. (1989). Silent reading rates in grade equivalents. *Journal of Reading Behavior*, *21*(2), 155–166. doi:10.1080/10862968909547667

Chall, J.S. (1996). American reading achievement: Should we worry? *Research in the Teaching of English*, *30*(3), 303–310.

Chall, J.S., Conard, S., & Harris, S. (1977). *An analysis of textbooks in relation to declining SAT scores*. Princeton, NJ: Educational Testing Service & College Entrance Examination Board.

Cunningham, A.E., & Stanovich, K.E. (1998). The impact of print exposure on word recognition. In J.L. Metsala & L.C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 235–262). Mahwah, NJ: Erlbaum.

Daane, M.C., Campbell, J.R., Grigg, W.S., Goodman, M.J., & Oranje, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading* (NCES 2006-469). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Denton, C.A., Barth, A.E., Fletcher, J.M., Wexler, J., Vaughn, S., Cirino, P.T., … Francis, D.J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading*, *15*(2), 109–135. doi:10.1080/10888431003623546

Downing, S.M., Baranowski, R.A., Grosso, L.J., & Norcini, J.J. (1995). Item type and cognitive ability measured: The validity evidence for multiple true-false items in medical specialty certification. *Applied Measurement in Education*, *8*(2), 187–207. doi:10.1207/s15324818ame0802_5

Fitzgerald, J., Elmore, J., Relyea, J.E., Hiebert, E.H., & Stenner, A.J. (2016). Has first-grade core reading program text complexity changed across six decades? *Reading Research Quarterly*, *51*(1), 7–28. doi:10.1002/rrq.115

Frisbie, D.A., & Becker, D.F. (1991). An analysis of textbook advice about true-false tests. *Applied Measurement in Education*, *4*(1), 67–83. doi:10.1207/s15324818ame0401_6

Gallo, D.R. (1972). *National Assessment of Educational Progress: Reading rate and comprehension: 1970–71 assessment* (Report No. 02-R-09). Denver CO: Education Commission of the States. Retrieved from ERIC database. (ED076934)

Gamson, D.A., Lu, X., & Eckert, S.A. (2013). Challenging the research base of the Common Core State Standards: A historical reanalysis of text complexity. *Educational Researcher*, *42*(7), 381–391. doi:10.3102/0013189X13505684

Griffith, L.W., & Rasinski, T.V. (2004). A focus on fluency: How one teacher incorporated fluency with her reading curriculum. *The Reading Teacher*, *58*(2), 126–137. doi:10.1598/RT.58.2.1

Hayes, D.P., Wolfer, L.T., & Wolfe, M.F. (1996). Schoolbook simplification and its relation to the decline in SAT-Verbal scores. *American Educational Research Journal*, *33*(2), 489–508. doi:10.3102/00028312033002489

Hiebert, E.H., & Mesmer, H.A.E. (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. doi:10.3102/0013189X12459802

Hiebert, E.H., Wilson, K.M., & Trainin, G. (2010). Are students really reading in independent reading contexts? An examination of comprehension-based silent reading rate. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 151–167). Newark, DE: International Reading Association.

Hudson, R.F., Lane, H.B., & Pullen, P.C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, *58*(8), 702–714. doi:10.1598/RT.58.8.1

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S.M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, *49*(24), 2948–2959. doi:10.1016/j.visres.2009.09.012

Huey, E.B. (1908). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.

Inhoff, A.W., Weger, U.W., & Radach, R. (2005). Sources of information for the programming of short- and long-range regressions during reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance*, (pp. 33–52). New York, NY: Oxford University Press.

Irwin, D.E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, *36*(1), 1–27. doi:10.1006/cogp.1998.0682

Joseph, H.S.S.L., Nation, K., & Liversedge, S.P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review*, *42*(2), 207–222.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., … Dunlop Velez, E. (2015). *The condition of education 2015* (NCES 2015-144). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Klauda, S.L., Wigfield, A., & Cambria, J. (2012). Struggling readers' information text comprehension and motivation in early adolescence. In J.T. Guthrie, A. Wigfield, & S.L. Klauda (Eds.), *Adolescents' engagement in academic literacy* (pp. 295–351). College Park: University of Maryland.

Kliegl, R. (2014). *Linear mixed models workshop: Analysis of experiments with R*. Retrieved from Potsdam Mind Research Repository website: read.psych.uni-potsdam.de/pmr2

LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293–323. doi:10.1016/0010-0285(74)90015-2

McConkie, G.W., Zola, D., Grimes, J., Kerr, P.W., Bryant, N.R., & Wolff, P.M. (1991). Children's eye movements during reading. In J.F. Stein (Ed.), *Vision and visual dysfunction: Vol. 13. Vision and visual dyslexia* (pp. 251–262). London, England: Macmillan.

Melekoğlu, M.A., & Wilkerson, K.L. (2013). Motivation to read: How does it change for struggling readers with and without disabilities? *International Journal of Instruction*, *6*(1), 77–88.

Mol, S.E., & Bus, A.G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267–296. doi:10.1037/a0021890

Morgan, P.L., & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional Children*, *73*(2), 165–183. doi:10.1177/001440290707300203

Mullis, I.V.S., Martin, M.O., Foy, P., & Drucker, K.T. (2013). *PIRLS 2011 assessment results: Student background data almanac by reading achievement (weighted)* [Data file]. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Retrieved from timssandpirls.bc.edu/pirls2011/international-database.html

National Center for Education Statistics. (2013a). Number and percentage of public school students eligible for free or reduced-price lunch, by state: Selected years, 2000–01 through 2011–12 [Table]. In *Common core of data (CCD), Public Elementary/Secondary School Universe Survey, 2000–01, 2005–06, 2010–11, and 2011–12*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://nces.ed.gov/programs/digest/d13/tables/dt13_204.10.asp

National Center for Education Statistics. (2013b). *The Nation's Report Card: Trends in academic progress 2012* (NCES 2013-456). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Endowment for the Arts. (2007). *To read or not to read: A question of national consequence* (Research Report No. 47). Washington, DC: Author. Retrieved from https://www.arts.gov/sites/default/files/ToRead.pdf

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, glossary of key terms*. Washington, DC: Authors.

National Institute of Child Health & Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Nelson, M.J., Brown, J.I., & Denny, M.J. (1960). *The Nelson–Denny Reading Test: Vocabulary, comprehension, rate*. Boston, MA: Houghton Mifflin.

No Child Left Behind Act of 2001, Pub. L. 107-110, 115 Stat. 1425 (2002).

O'Brien, B.A., Wallot, S., Haussmann, A., & Kloos, H. (2014). Using complexity metrics to assess silent reading fluency: A cross-sectional study comparing oral and silent reading. *Scientific Studies of Reading*, *18*(4), 235–254. doi:10.1080/10888438.2013.862248

Organisation for Economic Co-operation and Development. (2011). *Education at a glance 2011: OECD indicators*. Paris, France: OECD. Retrieved from www.oecd.org/education/school/educationataglance2011oecdindicators.htm

Organisation for Economic Co-operation and Development. (2013). *Programme for International Student Assessment (PISA) results from PISA 2012: United States*. Paris, France: OECD. Retrieved from www.oecd.org/unitedstates/PISA-2012-results-US.pdf

Pearson, P.D., & Goodin, S. (2010). Silent reading pedagogy: A historical perspective. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 3–23). Newark, DE: International Reading Association.

Perfetti, C.A. (1985). *Reading ability*. New York, NY: Oxford University Press.

Pinheiro, J.C., & Bates, D.M. (2000). *Mixed-effects models in S and S-PLUS* (pp. 209–210). New York, NY: Springer.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology*, *66*(3), 429–452. doi:10.1080/17470218.2012.750676

Rasinski, T.V., Padak, N.D., McKeon, C.A., Wilfong, L.G., Friedauer, J.A., & Heim, P. (2005). Is reading fluency a key for successful high school reading? *Journal of Adolescent & Adult Literacy*, *49*(1), 22–27. doi:10.1598/JAAL.49.1.3

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. doi:10.1080/17470210902816461

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Reardon, S.F., Valentino, R.A., & Shores, K.A. (2012). Patterns of literacy among U.S. students. *The Future of Children*, *22*(2), 17–37. doi:10.1353/foc.2012.0015

Reichle, E.D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*(4), 445–476. doi:10.1017/S0140525X03000104

Reutzel, D.R., Fawson, P.C., & Smith, J.A. (2008). Reconsidering silent sustained reading: An exploratory study of scaffolded silent reading. *The Journal of Educational Research*, *102*(1), 37–50. doi:10.3200/JOER.102.1.37-50

Rideout, V. (2014). *Children, teens, and reading: A common sense media research brief*. San Francisco, CA: Common Sense Media. Retrieved from https://www.commonsensemedia.org/research/children-teens-and-reading

Samuels, S.J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, *42*(4), 563–566.

Samuels, S.J., LaBerge, D., & Bremer, C.D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 715–720. doi:10.1016/S0022-5371(78)90433-4

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, *95*(1), 26–43. doi:10.1111/j.1540-4781.2011.01146.x

Schwerdt, G., & West, M.R. (2013). The impact of alternative grade configurations on student outcomes through middle and high school. *Journal of Public Economics*, *97*, 308–326. doi:10.1016/j.jpubeco.2012.10.002

Smith, N.B. (1965). *American reading instruction: Its development and its significance in gaining a perspective on current practices in reading.* Newark, DE: International Reading Association.

Sparks, R.L., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of '10 years later'. *Reading and Writing*, *27*(1), 189–211. doi:10.1007/s11145-013-9439-2

Spichtig, A., Vorstius, C., Greene, A., & Radach, R. (2009, August). *Visagraph III and EyeLink 2K: A comparison study.* Poster session presented at the European Conference on Eye Movements, Southampton, UK.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. doi:10.1598/RRQ.21.4.1

Stevens, R.J., Lu, X., Baker, D.P., Ray, M.N., Eckert, S.A., & Gamson, D.A. (2015). Assessing the cognitive demands of a century of reading curricula: An analysis of reading text and comprehension tasks from 1910 to 2000. *American Educational Research Journal*, *52*(3), 582–617. doi:10.3102/0002831215573531

Taylor, S.E. (1965). Eye movements in reading: Facts and fallacies. *American Educational Research Journal*, *2*(4), 187–202. doi:10.3102/00028312002004187

Taylor, S.E. (2009). *Visagraph III: Implementation guide.* Huntington Station, NY: Taylor Associates Communications.

Taylor, S.E., Frackenpohl, H., & Pettee, J.L. (1960). Grade level norms for the components of the fundamental reading skill. *EDL Research and Information Bulletin*, *3*, 22.

Taylor, S.E., Frackenpohl Morris, H., & White, C.E. (2004). *Visagraph® test booklet.* New York, NY: Taylor Associates Communications.

Thorndike, E.L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, *8*(6), 323–332. doi:10.1037/h0075325

U.S. Census Bureau. (1960). *Censuses of population and housing: 1960. Part I: 1960 population and housing censuses.* Retrieved from https://www.census.gov/prod/www/decennial.html#y1960popv1

U.S. Census Bureau. (2011). *Statistical abstract of the United States: 2011* (130th ed.). Washington, DC: U.S. Government Printing Office.

Venables, W.N., & Ripley, B.D. (2002). *Modern applied statistics with S.* New York, NY: Springer-Verlag.

Vonk, W., & Cozijn, R. (2003). On the treatment of saccades and regressions in eye movement measures of reading time. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 291–311). Oxford, UK: Elsevier.

Vorstius, C., Radach, R., & Lonigan, C.J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition*, *22*(3/4), 458–485. doi:10.1080/13506285.2014.881445

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, *16*(1/2), 312–339. doi:10.1080/09541440340000196

Williamson, G.L. (2006). *Aligning the journey with a destination: A model for K–16 reading standards.* Durham, NC: MetaMetrics. Retrieved from https://www.georgiastandards.org/resources/Documents/AligningJourneyWithDestination_MetaMetricsWhitepaper.pdf

Williamson, G.L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, *19*(4), 602–632. doi:10.4219/jaa-2008-832

Williamson, G.L., Fitzgerald, J., & Stenner, A.J. (2013). The Common Core State Standards' quantitative text complexity trajectory: Figuring out how much complexity is enough. *Educational Researcher*, *42*(2), 59–69. doi:10.3102/0013189X12466695

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. In *Proceedings of the ninth International Conference on Spoken Language Processing (INTERSPEECH 2006)* (pp. 541–544). Baixas, France: International Speech Communication Association. Retrieved from http://ldc.upenn.edu/myl/llog/icslp06_final.pdf

Submitted April 25, 2015
Final revision received November 13, 2015
Accepted December 7, 2015

ALEXANDRA N. SPICHTIG (corresponding author) is the chief research officer at Reading Plus, Winooski, Vermont, USA; e-mail alex@readingplus.com.

ELFRIEDA H. HIEBERT is the president and CEO of TextProject and a research associate at the University of California, Santa Cruz, USA; e-mail hiebert@textproject.org.

CHRISTIAN VORSTIUS is a senior lecturer in the Division of General and Biological Psychology of the School of Human and Social Sciences at Bergische Universität Wuppertal, NRW, Germany; e-mail vorstius@uni-wuppertal.de.

JEFFREY P. PASCOE is a research scientist at Reading Plus, Winooski, Vermont, USA; e-mail jeffrey.pascoe@readingplus.com.

P. DAVID PEARSON is a professor in the Language and Literacy and Human Development programs in the Graduate School of Education at the University of California, Berkeley, USA, and a member of the Reading Plus Advisory Team; e-mail ppearson@berkeley.edu.

RALPH RADACH is a professor of general and biological psychology in the School of Human and Social Sciences at Bergische Universität Wuppertal, NRW, Germany, and a member of the Reading Plus Advisory Team; e-mail radach@uni-wuppertal.de.