Making Connections

June 2016

School reading performance and the extended school day policy in Florida

Jessica Sidler Folsom
Yaacov Petscher
La'Tara Osborne-Lampkin
Stephan Cooley
Sarah Herrera
Mark Partridge
Kevin Smith
Florida State University

Key findings

Florida law requires the 100 lowest performing elementary schools in reading to extend the school day by one hour to provide supplemental reading instruction. This study found that those schools were smaller than other elementary schools and served a higher proportion of racial/ethnic minority students and students eligible for the school lunch program. The lowest performing schools reported increasing the number of minutes of reading instruction provided to students, increasing staff, and providing instruction in the extra hour that differed from instruction during the rest of the day. When growth in performance is measured, initially low scores generally rise, even in the absence of an intervention, because of natural year-to-year variations. While average school reading performance improved among the lowest performing schools, the increase did not exceed the small year-to-year variations expected when measuring initially low student performance.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL SOUTHEAST
Regional Educational Laboratory
At Florida State University

U.S. Department of Education
John B. King, Jr., *Secretary*

Institute of Education Sciences
Ruth Neild, *Deputy Director for Policy and Research*
Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance
Joy Lesnick, *Acting Commissioner*
Amy Johnson, *Action Editor*
Sandra Garcia, *Project Officer*

REL 2016–141

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

June 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0011 by Regional Educational Laboratory Southeast administered by Florida State University. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Folsom, J. S., Petscher, Y., Osborne-Lampkin, L., Cooley, S., Herrera, S., Partridge, M., & Smith, K. (2016). *School reading performance and the extended school day policy in Florida* (REL 2016–141). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

Summary

Since the 2012/13 school year Florida law has required the 100 lowest performing elementary schools in reading to extend the school day by one hour to provide supplemental reading instruction. All of Florida's elementary schools are evaluated each year, and the 100 lowest are identified based on the most recent year's school reading performance.

Regional Educational Laboratory Southeast's Improving Literacy Alliance and Improving Low Performing Schools Alliance—both of which have Florida Department of Education administrators as members—requested an analysis of the lowest performing elementary schools. Specifically, the alliances were interested in how the lowest performing schools implemented the extended school day policy (for example, the methods used to add the additional hour, staffing, and delivery of instruction) and in the trends in school reading performance among the lowest performing schools and other elementary schools.

Based on publicly available data on school reading performance between 2011/12 and 2013/14 and survey data collected by the Florida Department of Education, the study found:

- The lowest performing elementary schools were smaller than other elementary schools and enrolled a larger proportion of racial/ethnic minority students and students eligible for the school lunch program.
- The elementary schools that implemented the extended school day policy in 2013/14 reported increasing reading instruction time, increasing staff, providing instruction in the extra hour that differed from instruction during the rest of the day, and complying with the extended school day policy.
- When measuring growth in performance, initially low scores generally increase over time toward the mean, even in the absence of an intervention, because of a statistical phenomenon called regression to the mean. While average school reading performance improved among the lowest performing schools, the increase did not exceed the small year-to-year variations expected because of regression to the mean when measuring initially low student performance.

Contents

Figures

Map

Tables

Why this study?

Since the 2012/13 school year Florida law has required the 100 lowest performing elementary schools in reading to extend the school day by one hour (see box 1 for definitions of key terms used in the report).[1] The additional hour must be spent on intensive, research-based reading instruction provided by effective teachers. However, the law does not provide guidance on how the lowest performing schools are selected, how to incorporate the extra hour, or how to define effective teachers.

The Florida Department of Education—through the Just Read, Florida! office and the Bureau of School Improvement—is responsible for overseeing implementation of the extended school day policy. It developed a system to identify the lowest performing elementary schools and notify them of the requirement to extend the school day by an hour. Policy implementation was determined by the district or school (depending on the district and number of lowest performing schools).

Since the 2012/13 school year Florida law has required the 100 lowest performing elementary schools in reading to extend the school day by one hour

Regional Educational Laboratory Southeast's Improving Literacy Alliance and Improving Low Performing Schools Alliance—both of which have Florida Department of Education administrators as members—requested an analysis of the lowest performing elementary

Box 1. Key terms

Extended school day policy. The implementation of the law requiring the 100 lowest performing elementary schools in reading to extend the school day by an hour. The extra hour must be spent on intensive, research-based reading instruction provided by effective teachers.

Lowest performing schools. The 159 lowest performing elementary schools in reading that were required to implement the extended school day policy. This includes 68 schools that implemented the policy in 2012/13 only, 66 schools that implemented the policy in 2013/14 only, and 25 schools that implemented the policy in both 2012/13 and 2013/14. The Florida Department of Education identified the lowest performing schools based on school reading performance. Schools were identified during the summer (typically in August) of the year that they would be required to extend the school day. While 100 schools were identified each year, some schools restructured or closed before implementing the policy, so the number of schools that implemented the policy each year does not total 100.

Other elementary schools. The 1,578 elementary schools that were not identified as a lowest performing school in reading in either 2012/13 or 2013/14. These schools were not required to implement the extended school day policy between 2011/12 and 2013/14.

School reading performance. The sum of the percentage of students scoring at or above achievement level 3 (satisfactory) on the reading component of the prior year's Florida Comprehensive Assessment Test 2.0 (FCAT 2.0) at each school and the percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0 at each school. It has a theoretical range of 0–200, but the observed range in this study is 54–195. School reading performance is the dependent variable of this study.

Title I. Title I provides additional resources to schools with economically disadvantaged students (Florida Department of Education, n.d.). Schools implementing schoolwide reform models provide all students with access to services. In this report these schools are referred to as Title I schools.

schools. Specifically, the alliances were interested in how the lowest performing schools implemented the extended school day policy (for example, the methods used to add the additional hour, staffing, and delivery of instruction) and in the trends in school reading performance among the lowest performing schools and other elementary schools.

This study addresses three research questions:

- Where were the lowest performing schools located, and how did the demographic composition of the lowest performing schools compare with that of other elementary schools?
- How did districts and schools incorporate the extra hour of instruction in 2013/14? Was there evidence of additional reading instruction? How was the additional time added, and how was instruction provided?
- What was the nature and variability of growth in school reading performance among the lowest performing schools and other elementary schools between 2011/12 and 2013/14? How much growth was expected in the absence of the extra hour of instruction, and did the lowest performing schools exhibit more growth than expected?

The data and methodology behind the study are summarized in box 2.

Box 2. Study data and methodology

Data

The study examined 1,737 elementary schools, 159 of which were identified by the Florida Department of Education as the lowest performing schools and implemented the extended school day policy in 2012/13, 2013/14, or both. The study data cover only open, regular education elementary schools that received a school grade in all years between 2011/12 and 2013/14. Schools that were exempt from grading because they had been open for only one year and alternative schools and special education center schools that receive school improvement ratings rather than school grades were excluded.

The Florida Department of Education provided the study team with a list of the lowest performing elementary schools in reading that implemented the extended school day policy for each year of the study. These schools were identified by the department's Just Read, Florida! office based on school reading performance, which is the sum of the percentage of students scoring at or above achievement level 3 (satisfactory) on the reading component of the prior year's Florida Comprehensive Assessment Test 2.0 (FCAT 2.0) at each school and the percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0 at each school. In some cases schools identified as among the 100 lowest performing schools closed or substantially restructured (for example, became a K–2 school or combined into an elementary/middle school) after the spring administration of the FCAT 2.0. When a school did not implement the extended school day policy because it closed or restructured, the Florida Department of Education did not identify a replacement school. Thus, in any given year fewer than 100 schools may have implemented the extended school day policy.

Of the 100 lowest performing schools identified for 2012/13, 7 schools closed (2 closed prior to the 2012/13 school year and are excluded from the study, and 5 were among the lowest performing schools in 2013/14 as well but closed prior to the 2013/14 school year and are excluded from the study), and 93 schools implemented the extended school day policy in 2012/13. Of those 93 schools, 68 implemented the policy in 2012/13 only (referred to as the lowest performing schools in 2012/13 only), and 25 implemented the policy in both 2012/13 and 2013/14 (referred to as the lowest performing schools in both years).

(continued)

Box 2. Study data and methodology *(continued)*

Of the 101 lowest performing schools identified for 2013/14, 10 closed (5 were among the lowest performing schools in 2012/13 and implemented the policy that year but closed prior to the 2013/14 school year and are excluded from the study, and 5 closed prior to the 2013/14 school year and are excluded from the study), and 91 implemented the extended school day policy in 2013/14. Of those 101 schools, 66 implemented the policy in 2013/14 only (referred to as the lowest performing schools in 2013/14 only), and 25 schools implemented the policy in both 2012/13 and 2013/14 (referred to as the lowest performing schools in both years).

For both the lowest performing schools and other elementary schools, School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a), Master School Identification files (Florida Department of Education, 2012b, 2013b, 2014c), and Membership by School by Grade files (Florida Department of Education, 2012c, 2013c, 2014d) provided data on school type, grade configuration, and status (active or inactive), percentage of students eligible for the school lunch program, percentage of racial/ethnic minority students, region, Title I status, total school enrollment, percentage of students scoring at or above achievement level 3 (satisfactory) on the reading component of the FCAT 2.0 at each school, and percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0. The study team calculated school reading performance using the same method as the Just Read, Florida! office, and the values were used as the dependent variable in all growth analyses. See tables A1 and A2 in appendix A for descriptive statistics.

The Florida Department of Education also provided data from a survey that the Just Read, Florida! office administered to districts to collect information about the lowest performing schools in 2013/14. In some cases the district completed the surveys on behalf of the schools; in other cases the schools themselves completed the surveys. The survey asked about schools' implementation of the extended school day policy. Of the 91 schools that implemented the policy in 2013/14, 86 returned surveys—a 95 percent response rate. However, not every question was answered on each survey. Of the 86 responding schools, 63 were lowest performing schools in 2013/14 only, and 23 were lowest performing schools in both years. See tables A3 and A4 in appendix A for descriptive statistics and appendix B for the survey itself.

Data analysis

Descriptive statistics (means, standard deviations, and cross-tabulations) were calculated to describe the location and demographic composition of the lowest performing schools and other elementary schools. Similar descriptive statistics were calculated to summarize survey findings on how the 2013/14 lowest performing schools implemented the policy.

The study team calculated the expected change in school reading performance among the lowest performing schools through a regression to the mean analysis. Because the expected change was an estimation and not a precise value, it was also important to calculate a plausible range with confidence intervals. A plausible range of expected change was constructed using 95 percent confidence intervals.

School reading performance growth was calculated using multilevel random effects growth curve modeling. Time-varying covariates were used to account for differential growth in school reading performance among the lowest performing schools in the year in which the extended school year was implemented (McCoach & Kaniskan, 2010). A series of models were built and tested, and model 4a was used for interpretation (see table C3 in appendix C). The model intercept provided a baseline (2011/12) average school reading performance for other elementary schools and three mean adjustments—one for each cohort (lowest performing school in 2012/13 only, lowest performing school in 2013/14 only, and lowest performing school in both 2012/13 and 2013/14). The model yielded an overall growth estimate for other elementary schools and a growth difference for the lowest performing schools while the policy was implemented. The overall growth estimate and the growth difference estimate were summed to calculate school reading performance growth for the lowest performing schools the year a school implemented the policy. A plausible range of growth was constructed using 95 percent confidence intervals. Results of the growth modeling were compared with the results of the regression to the mean analysis to determine whether the observed growth was beyond what would be expected if schools had not implemented the policy.

Details on the data and methodology are in appendix C; tests of significance are in appendix D.

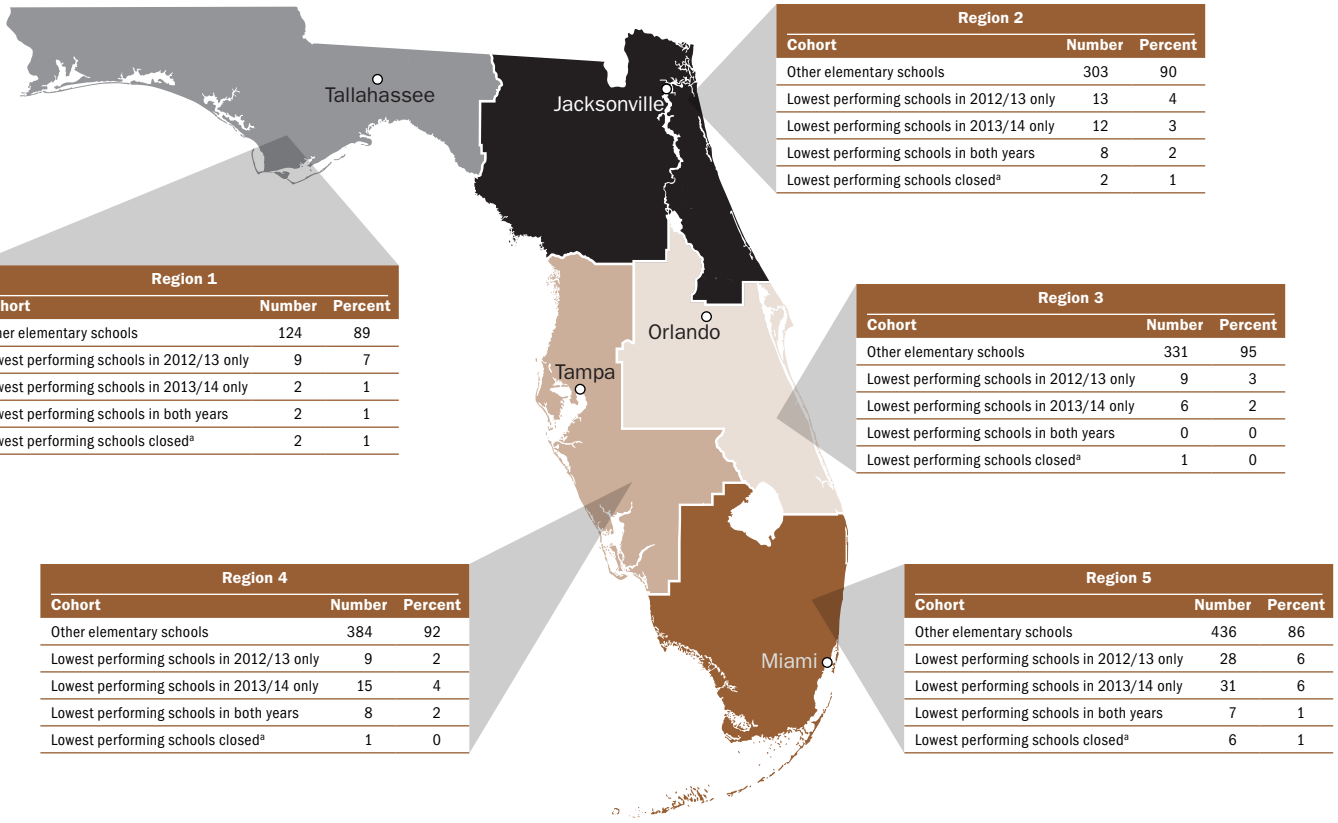
What the study found

This section details the results of the analyses. It describes the characteristics of Florida's elementary schools, implementation of the extended school day policy, and growth in school reading performance from 2011/12 through 2013/14.

The lowest performing elementary schools were smaller than other elementary schools and enrolled a larger proportion of racial/ethnic minority students and students eligible for the school lunch program

Region. The lowest performing schools were located in 28 of Florida's 67 school districts and in all five of Florida's regions (map 1). However, the lowest performing schools were not evenly distributed across regions. In each region approximately 5–14 percent of elementary schools were identified as the lowest performing schools. All regions had lowest performing schools in each year, but only region 3 (the east-central region, which includes Orlando) did not have any schools that were identified as lowest performing schools in both years. The differences between regions were statistically significant (see appendix D).

Map 1. The lowest performing elementary schools in reading were located in all five of Florida's regions



Region 1

Cohort	Number	Percent
Other elementary schools	124	89
Lowest performing schools in 2012/13 only	9	7
Lowest performing schools in 2013/14 only	2	1
Lowest performing schools in both years	2	1
Lowest performing schools closed[a]	2	1

Region 2

Cohort	Number	Percent
Other elementary schools	303	90
Lowest performing schools in 2012/13 only	13	4
Lowest performing schools in 2013/14 only	12	3
Lowest performing schools in both years	8	2
Lowest performing schools closed[a]	2	1

Region 3

Cohort	Number	Percent
Other elementary schools	331	95
Lowest performing schools in 2012/13 only	9	3
Lowest performing schools in 2013/14 only	6	2
Lowest performing schools in both years	0	0
Lowest performing schools closed[a]	1	0

Region 4

Cohort	Number	Percent
Other elementary schools	384	92
Lowest performing schools in 2012/13 only	9	2
Lowest performing schools in 2013/14 only	15	4
Lowest performing schools in both years	8	2
Lowest performing schools closed[a]	1	0

Region 5

Cohort	Number	Percent
Other elementary schools	436	86
Lowest performing schools in 2012/13 only	28	6
Lowest performing schools in 2013/14 only	31	6
Lowest performing schools in both years	7	1
Lowest performing schools closed[a]	6	1

Note: Percentages may not sum to 100 because of rounding.

a. Some schools closed or substantially restructured (for example, became a combined elementary and middle school) after the spring Florida Comprehensive Assessment Test 2.0 administration and prior to identification as a lowest performing school (box 2); these schools were not included in the analyses.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2013a, 2014a).

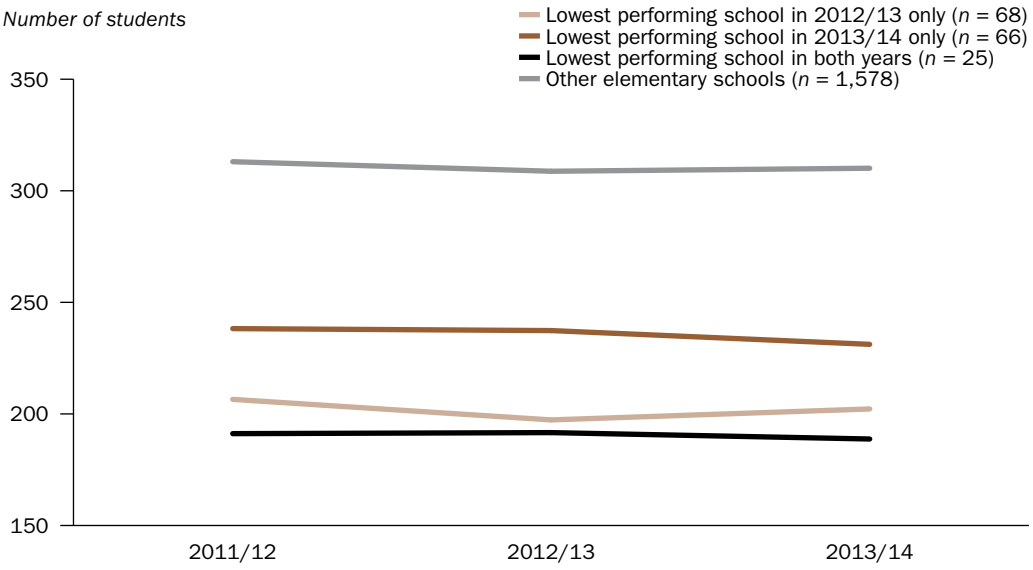
Enrollment was stable across years, and the lowest performing schools enrolled fewer students than other elementary schools. Enrollment averaged 189–238 students in the lowest performing schools and 309–313 students in other elementary schools (figure 1; see also table A1 in appendix A). The difference was statistically significant (see appendix D). Enrollment and school reading performance were moderately correlated (.28–.31; see table A2 in appendix A, and appendix D).

All the lowest performing elementary schools were Title I schools and enrolled a high proportion of racial/ethnic minority students and students eligible for the school lunch program. On average, in the lowest performing schools 88–94 percent of students were a racial/ethnic minority, and 92–95 percent of students were eligible for the school lunch program (figure 2; see also table A1 in appendix A). In contrast, 62 percent of other elementary schools were Title I schools, and on average 58 percent of students were a racial/ethnic minority and 64 percent were eligible for the school lunch program. All differences were statistically significant (see appendix D).

The elementary schools that implemented the extended school day policy in 2013/14 reported increasing reading instruction time, increasing staff, providing instruction in the extra hour that differed from instruction during the rest of the day, and complying with the extended school day policy

In general, the survey responses of the lowest performing schools in 2013/14 only and the lowest performing schools in both years were not significantly different (see appendix D). Descriptive statistics for both the aggregated and disaggregated samples are in tables A2 and A4 in appendix A.

Figure 1. Enrollment averaged 189–238 students in the lowest performing elementary schools and 309–313 students in other elementary schools

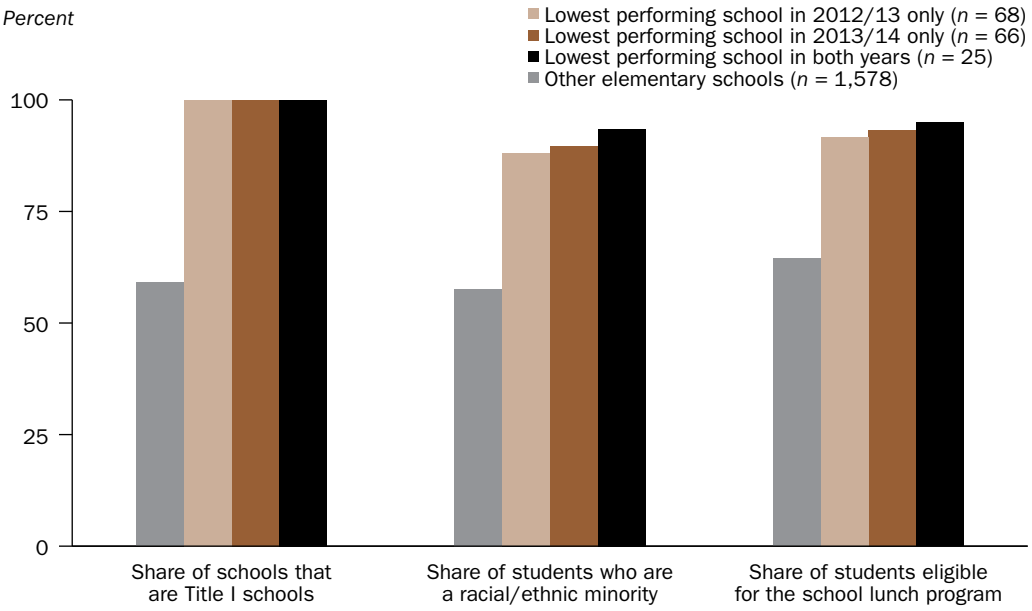


Note: See table A1 in appendix A for exact values.

Source: Authors' analysis of data from School Membership by Grade files (Florida Department of Education, 2012c, 2013c, 2014d).

Enrollment averaged 189–238 students in the lowest performing schools and 309–313 students in other elementary schools

Figure 2. The lowest performing elementary schools enrolled a higher proportion of racial/ethnic minority students and students eligible for the school lunch program than other elementary schools did



On average, in the lowest performing schools 88–94 percent of students were a racial/ethnic minority, and 92–95 percent of students were eligible for the school lunch program

Note: See table A1 in appendix A for exact values.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

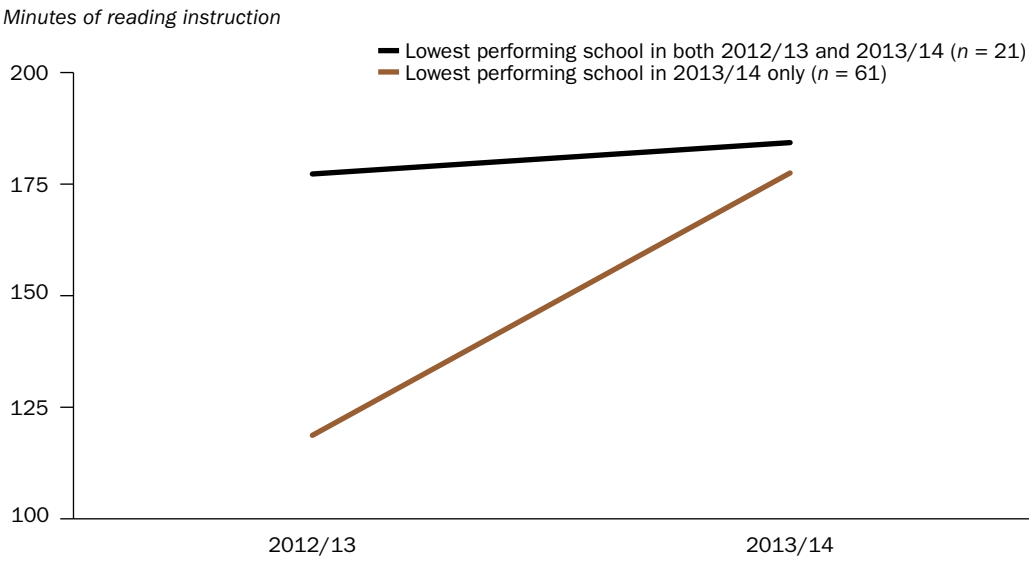
The schools that implemented the extended school day policy in 2013/14 reported increasing the number of minutes of reading instruction and staff over prior years (when the extended school day policy was not implemented). They also reported providing instruction in the extra hour that differed from instruction during the rest of the day and complying with the extended school day policy.

Schools reported increasing the number of minutes of reading instruction provided each day. The schools that implemented the extended school day policy in 2013/14 reported providing an average of 134 minutes of reading instruction a day in 2012/13 and 179 minutes in 2013/14. Among schools with data for both years there was a statistically significant increase of 46 minutes (see table A3 in appendix A, and appendix D).[2]

Of the schools that implemented the extended school day policy in 2013/14, those that implemented it in 2013/14 only (and not in both years) reported providing an average of 119 minutes of reading instruction a day in 2012/13 and 177 minutes in 2013/14. Among schools with data for both years there was a statistically significant increase of 60 minutes of reading instruction (figure 3; see also table A3 in appendix A, and appendix D).

Of the schools that implemented the extended school day policy in 2013/14, those that implemented it in both years (and thus were in their second year of implementation) reported providing an average of 177 minutes of reading instruction a day in 2012/13 and 184 minutes in 2013/14, a statistically significant increase of 7 minutes (see figure 3; see also table A3 in appendix A, and appendix D).

Figure 3. Between 2012/13 and 2013/14 time for reading instruction each day increased 60 minutes among the lowest performing elementary schools in 2013/14 only and 7 minutes among the lowest performing elementary schools in both years



Note: See table A3 in appendix A for exact values.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

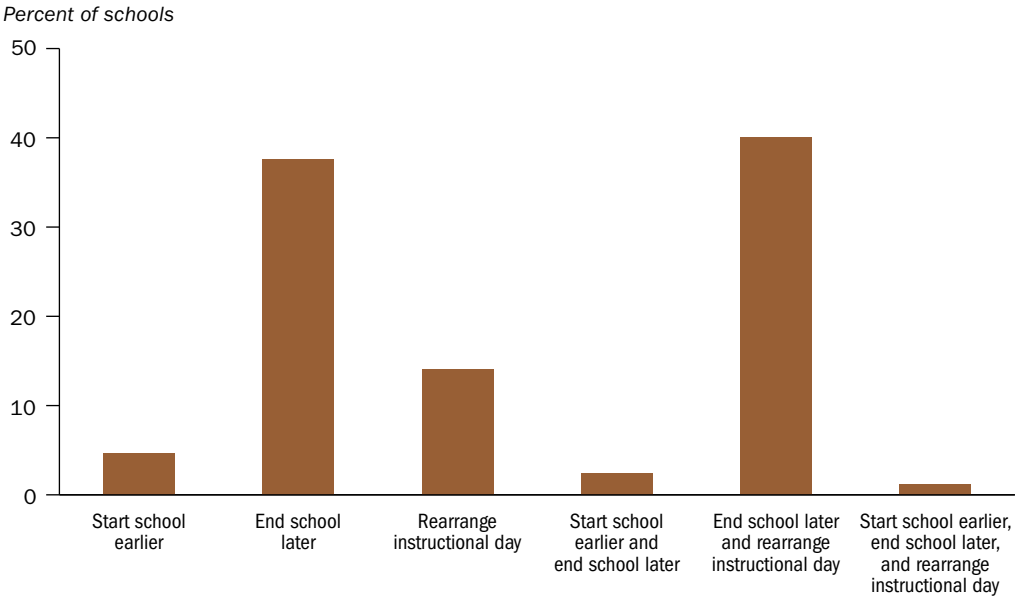
Of the schools that implemented the extended school day policy in 2013/14, those that implemented it in 2013/14 only (and not in both years) reported providing an average of 119 minutes of reading instruction a day in 2012/13 and 177 minutes in 2013/14

Schools reported using a variety of strategies to incorporate the extra hour of instruction. The lowest performing schools in 2013/14 only and the lowest performing schools in both years did not report statistically different approaches to incorporating the extra hour (see appendix D). The most common strategy was to end the school day later and rearrange the instructional day (figure 4; see also table A4 in appendix A).

Schools reported using a combination of students' regular classroom teachers and other staff to provide the extra hour of instruction. There was not a statistically significant difference in who provided the extra instruction between the lowest performing schools in 2013/14 only and the lowest performing schools in both years (see appendix D). Some 29 percent of schools reported using only the students' regular teachers, and 13 percent used only other staff (figure 5; see also table A4 in appendix A).

Schools reported adding staff to meet the increased instructional needs. Among schools that implemented the extended school day policy in 2013/14, added staff included reading coaches, teachers, paraprofessionals, volunteers, and other staff (figure 6; see also table A4 in appendix A). There was not a statistically significant difference in staff changes for reading coaches, paraprofessionals, or volunteers between the lowest performing schools in 2013/14 only and the lowest performing schools in both years. However, there was a statistically significant difference in staff changes for teachers and for other staff (see appendix D). A larger percentage of schools that were lowest performing in both years lost teachers and other staff (see table A4 in appendix A). A larger percentage of schools that were lowest performing schools in 2013/14 only gained teachers and other staff (see table A4 in appendix A).

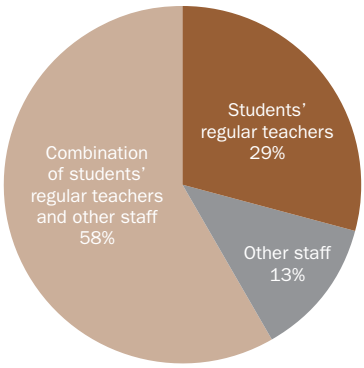
Figure 4. The most common strategy for elementary schools that implemented the extended school day policy in 2013/14 to incorporate the extra hour of instruction was to end the school day later and rearrange the instructional day



Note: Data are for 85 schools. See table A4 in appendix A for exact values.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

Figure 5. Elementary schools that implemented the extended school day policy in 2013/14 reported using a combination of students' regular classroom teachers and other staff to provide the extra hour of reading instruction

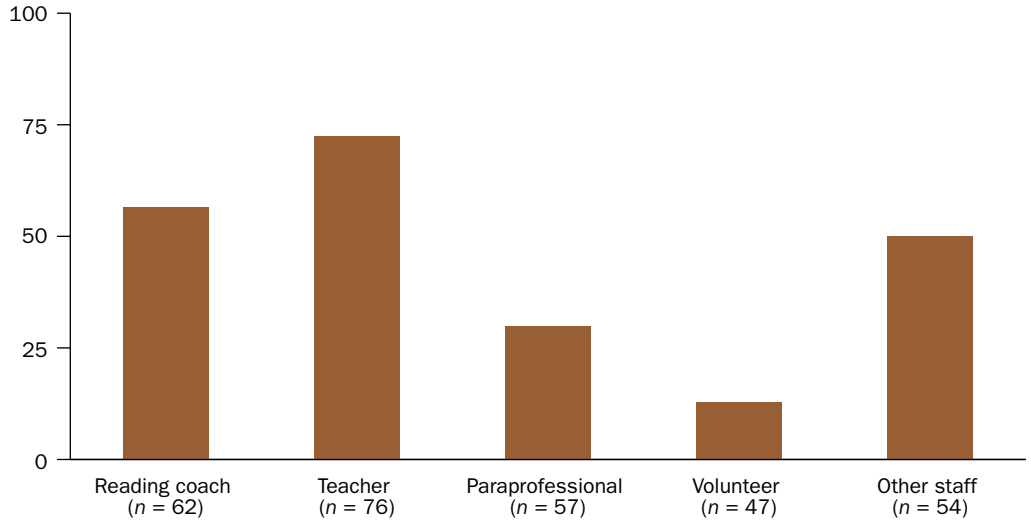


Note: Data are for 48 schools. See table A4 in appendix A for exact values.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

Figure 6. Elementary schools that implemented the extended school day policy in 2013/14 reported adding staff such as reading coaches, teachers, paraprofessionals, and volunteers to meet the increased instructional needs

Percent of schools that reported an increase in the type of staffing



There was not a statistically significant difference in staff changes for reading coaches, paraprofessionals, or volunteers between the lowest performing schools in 2013/14 only and the lowest performing schools in both years

Note: *n* values refer to the number of schools that responded to the survey question on changes in each type of staffing. For example, the first bar indicates that 62 schools identified as the lowest performing schools in 2013/14 provided information on changes in reading coach staffing and that approximately 58 percent of those schools hired additional reading coaches. See table A4 in appendix A for exact values.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

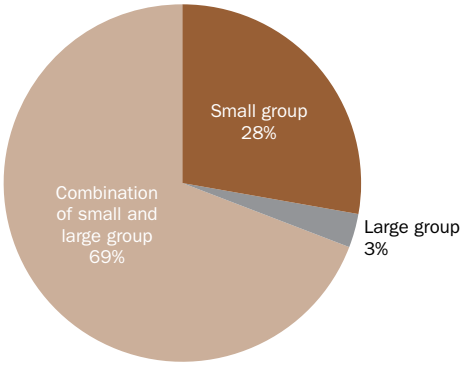
Schools reported using a variety of grouping strategies to provide the extra hour of reading instruction. While it was not possible for the study team to determine whether schools provided differentiated instruction as required by the extended school day policy, the fact that schools reported using small groups of students sorted by ability level for the extra hour of instruction suggests that schools may have attempted to provide differentiated instruction. There were no statistically significant differences in grouping of students for instruction between the lowest performing schools that implemented the extended school day policy in 2013/14 only and the lowest performing schools that implemented the policy in both years (see appendix D).

Nearly all schools reported using small-group instruction either exclusively or in combination with large-group instruction (figure 7; see also table A4 in appendix A). Very few schools reported using large-group instruction exclusively.

Nearly all schools reported grouping students for instruction by ability either exclusively or in combination with grouping students with mixed abilities (figure 8; see also table A4 in appendix A). Very few schools reported only grouping students with mixed abilities.

Schools reported complying with the extended school day policy. Compliance with the policy was based on five broad categories: student enrollment, teachers, curriculum, assessment, and instruction. A majority of schools (77–97 percent across criteria) reported meeting or exceeding each criterion for compliance (figure 9; see also table A4 in appendix A). The weakest reported area of compliance was using teachers who displayed

Figure 7. Nearly all elementary schools that implemented the extended school day policy in 2013/14 reported using small-group reading instruction either exclusively or in combination with large-group instruction

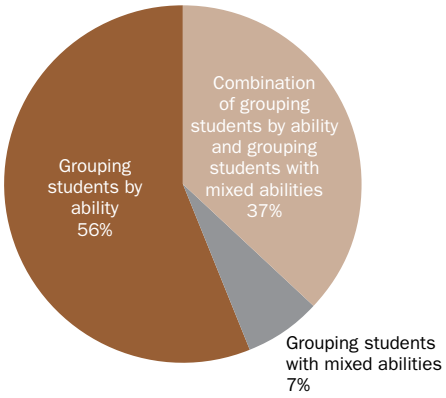


Note: See table A4 in appendix A for exact values.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

Nearly all schools reported using small-group instruction either exclusively or in combination with large-group instruction and grouping students for instruction by ability either exclusively or in combination with grouping students with mixed abilities

Figure 8. Nearly all schools reported grouping students for instruction by ability either exclusively or in combination with grouping students with mixed abilities



Note: See table A4 in appendix A for exact values.

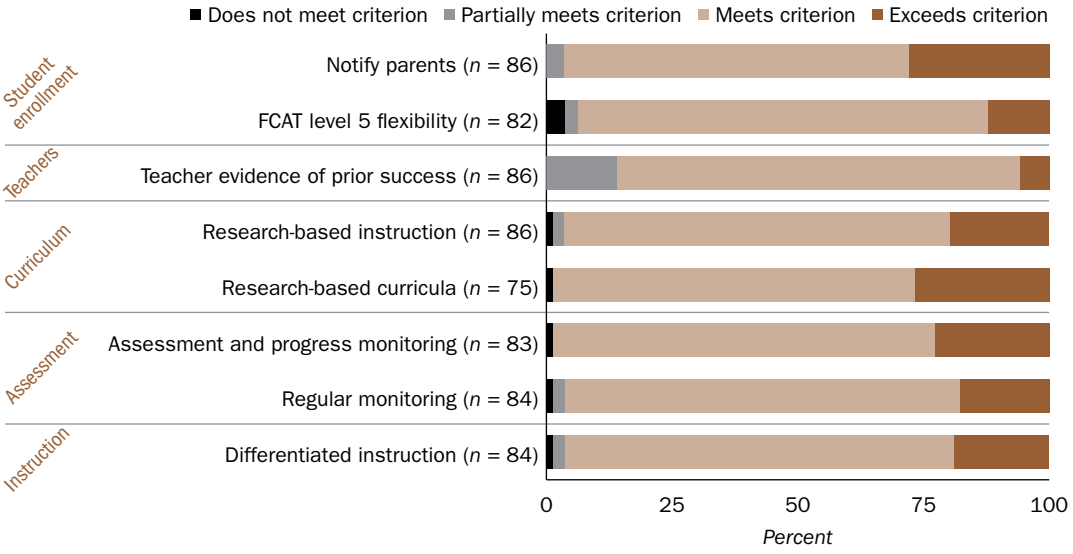
Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

evidence of prior success in teaching reading. The strongest reported areas of compliance were providing research-based instruction and using assessment and progress monitoring to determine instructional priorities. There were no statistically significant differences in compliance between the lowest performing schools that implemented the extended school day policy in 2013/14 only and those that implemented the policy in both years (see appendix D).

Average growth in school reading performance among the lowest performing elementary schools did not exceed what would have been expected in the absence of the extra hour of instruction

Growth in school reading performance was analyzed in two phases (see appendix C). The first phase estimated the expected change for the lowest performing schools assuming that

Figure 9. The majority of schools that implemented the extended school day policy in 2013/14 reported meeting or exceeding each criterion for compliance with the policy



FCAT is Florida Comprehensive Assessment Test 2.0.

Note: The *n* values refer to the number of schools that responded to the survey question on each criterion. For example, the first bar indicates that 86 schools provided information on whether they notified parents that their student had been enrolled in an extra hour of reading instruction each day. See table A4 in appendix A for exact values.

a. Schools were given the flexibility to allow students scoring at Level 5 to opt out of the extended reading instruction. Level 5 is the highest achievement level and indicates that a "student has success with the most challenging content of the Sunshine State Standards. A student scoring in Level 5 answers most of the test questions correctly, including the most challenging questions" (Florida Department of Education, 2008).

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

The weakest reported area of compliance was using teachers who displayed evidence of prior success in teaching reading. The strongest reported areas of compliance were providing research-based instruction and using assessment and progress monitoring to determine instructional priorities

the extra hour of reading instruction had no impact (box 3). The second phase compared observed growth in school reading performance with the expected change.

The expected change in school reading performance among the lowest performing schools was 7.1 points, with a plausible range of 6.4–7.7 points (see appendix C).

Initial school reading performance and observed growth were estimated in a single model for all schools in the study (model 4a; see table C3 in appendix C). As expected, all the lowest performing schools had lower school reading performance than other elementary schools prior to implementing the extended school day policy (figure 10).[3] These differences were statistically significant (see table C3 in appendix C).

The model did not provide a direct estimate of growth in school reading performance, so growth was calculated by adding the overall growth estimate to the growth difference for the lowest performing schools. The overall growth estimate (–1.2 points) was statistically significant, and the growth difference for the lowest performing schools (8.9 points) was statistically significant (see table C3 in appendix C), which yields a growth estimate of 7.7 points for the lowest performing schools that implemented the extended school day policy. Using an appropriate statistical procedure generated a plausible range of 6.2–9.3 points

Box 3. Accounting for the effect of regression to the mean and calculating expected change

When measuring growth in performance, initially low scores generally increase over time toward the mean, even in the absence of an intervention specifically targeting those schools, because of a statistical phenomenon called regression to the mean, the effect of which is referred to as expected change. The farther a score is below the mean, the higher the likelihood it will increase (and the farther a score is above the mean, the higher the likelihood it will decrease). It is important to calculate expected change, particularly when an intervention is provided to a group with low initial scores, so that expected change can be compared with observed change to determine whether the observed change is beyond what would have been expected.

Imagine a school performance measure administered twice to all schools in a state. Without intervention or policy changes, the overall state average score and variation around that score will remain consistent between the two assessment points. Now imagine two schools within this sample: one high performing (that is, with a high score at the first performance measurement) and one low performing (that is, with a low score at the first performance measurement). If between the first and second measurement the high-performing school's score decreased and the low-performing school's score increased, in both cases the second measurement is closer to the overall state average. The changing of scores in the direction of the state average, outside of any intervention or policy change, is the effect of regression to the mean, and the value of the change is the expected change.

In the current study, because the initial school reading performance of the lowest performing schools was in the 5th percentile or below, which was substantially below the mean, it was likely that school reading performance would increase without implementing the extended school day policy. Thus, data for 2011/12–2013/14 were used to estimate the expected annual change in school reading performance.

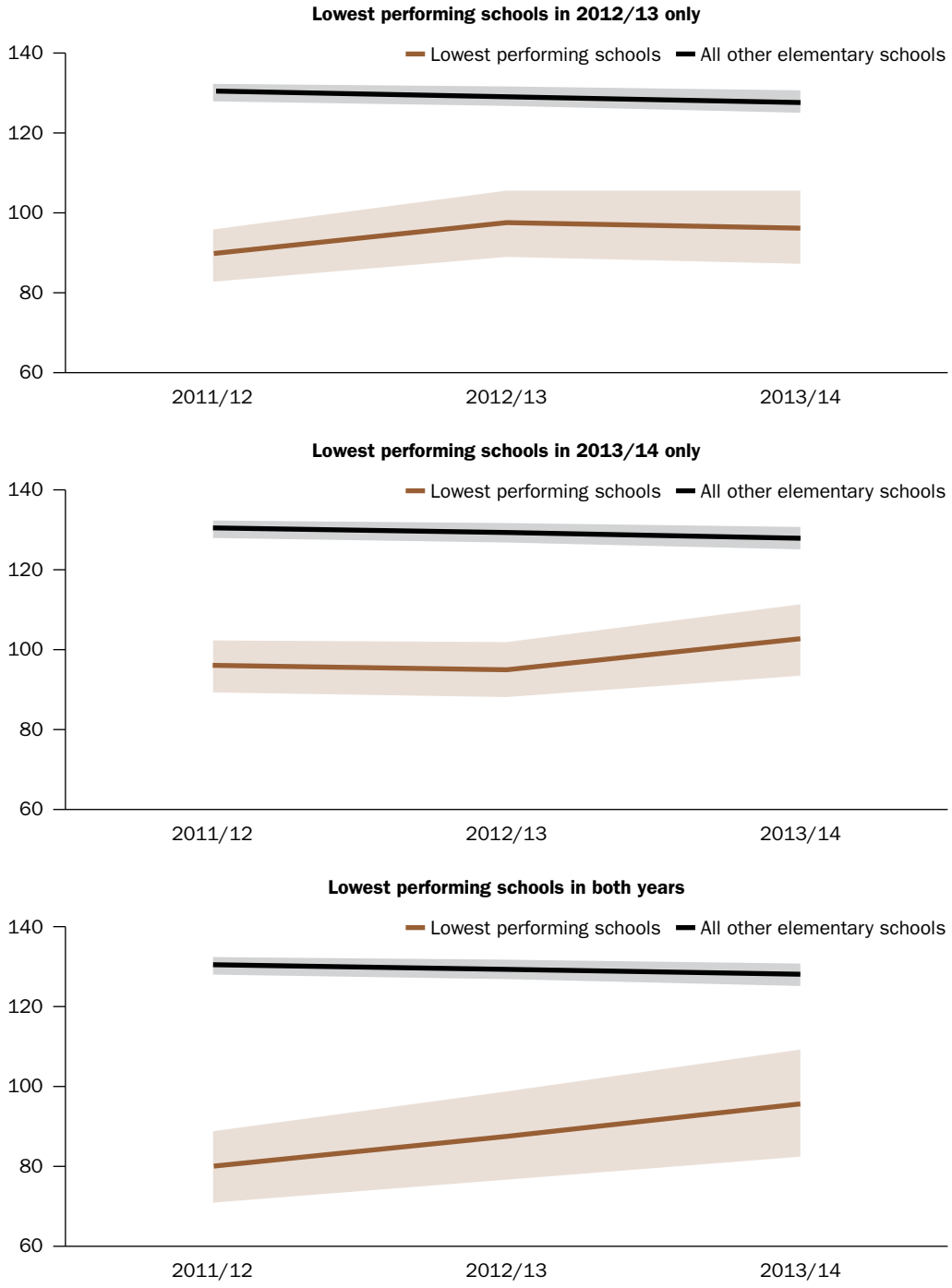
for the true growth in school reading performance among the lowest performing schools the year the school implemented the policy. That is, between the year before the extended school day policy was implemented and the year the policy was implemented, school reading performance increased an average of 6.2–9.3 points (see figure 10).

Growth in school reading performing in each cohort of lowest performing schools was different from growth in school reading performance in each cohort of other elementary schools the year the extended school day policy was implemented. However, given that the expected school reading performance change was 6.4–7.7 points, the overlap between the range of observed growth (6.2–9.3 points) and expected change (6.4–7.7 points) among the lowest performing schools while the extended school day policy was implemented means that school reading performance was not higher than expected in the absence of increased instruction.[4]

The overlap between the range of observed growth and expected change among the lowest performing schools while the extended school day policy was implemented means that school reading performance was not higher than expected in the absence of increased instruction

Figure 10. School reading performance increased the year the lowest performing schools were required to extend the school day

School reading performance (points)



Note: The shaded area is the plausible range (95 percent confidence interval).

Note: Data are results of model 4a (see table C3 in appendix C). All lowest performing schools cohorts were estimated simultaneously but are presented separately to avoid overlap of lines and confusion in interpretation.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

Implications of the study findings

This study highlights several implications for practice and directions for future research.

Accounting for expected change is important when measuring growth and making policy decisions

While school reading performance grew considerably, this change cannot be attributed to the policy because it did not exceed what was expected. If this study had not accounted for expected change, it would have appeared that growth in school reading performance among the lowest performing schools was attributed to the extra hour of instruction provided. Unsupportable conclusions might have been made about the relationship between extending the school day and improved student performance.

If this study had not accounted for expected change, it would have appeared that growth in school reading performance among the lowest performing schools was attributed to the extra hour of instruction provided

The Florida Department of Education may want to consider using a cutpoint on school reading performance rather than rank order when identifying the lowest performing elementary schools

This study was not designed to determine which method—rank order or cutpoint—is better but to bring to light some considerations. The impetus for identifying the lowest performing schools based on rank order of school reading performance rather than a cutpoint (for example, a school reading performance of 100 out of 200) is likely based on funding, as there is a specific allocation that can serve only a specific number of schools.

When using rank order, a school could have the same school reading performance in multiple years and be considered a lowest performing school one year and not the next. For example, if the highest school reading performance among the 100 lowest performing schools was 90 points in one year and 93 points among the next year's 100 lowest performing schools, a school with school reading performance of 92 points in both years would be considered a lowest performing school in one year and not the other, even though school reading performance did not change. Therefore, using exit from designation as a lowest performing school as the only benchmark for success of the extended hour of instruction may be misleading, as it reflects only change in ranking, not change in school reading performance score. If a consistent cutpoint across years were used, designation would be based on the cutpoint, and changes in designation would reflect change in school reading performance.

The Florida Department of Education may want to consider identifying the lowest performing elementary schools using multiple years of data

Using more years of data improves accuracy of identification and reduces measurement error because school reading performance changes over time. Schools that are consistently the lowest performing may benefit more from the extra hour of instruction than schools that are the lowest performing for only a single year.

In the current study the lowest performing schools served fewer students than other elementary schools. Although enrollment is moderately correlated with school reading performance, smaller schools are more easily influenced by variations within a single student or small group of students. For example, a single year of poor performance could be related to internal factors such as substantial staff changes or restructuring. A single year of poor performance could also be related to endogenous factors such as death of a student or

natural disasters, such as a hurricane (Sacerdote, 2008). The year following such an event, scores may return to normal.

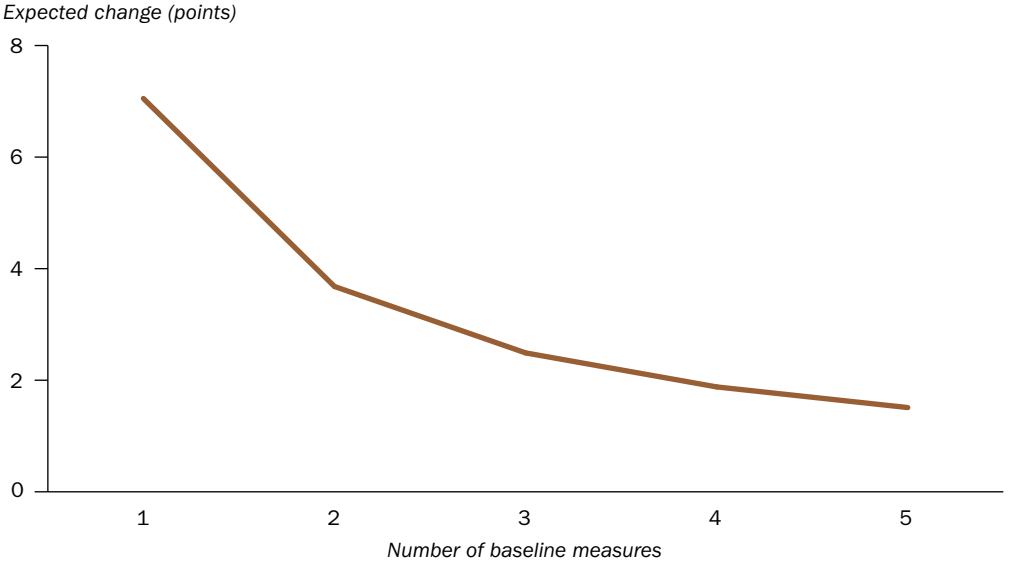
There is precedent for including multiple years of baseline data in identifying and evaluating schools and teachers. For example, the Florida Department of Education's Bureau of School Improvement implements the Differentiated Accountability program, which is a statewide network of strategic support provided to schools and districts, differentiated by need. In the program, schools are classified based on school grades from the previous three years (Florida Department of Education, 2015). Similarly, Florida's teacher value-added scores are also based on three years of student performance (Florida Department of Education, 2011).

Moreover, when multiple baseline measurements are used to select a group for an intervention, the expected change decreases (Barnett, van der Pols, & Dobson, 2005; Davis, 1976; Johnson & George, 1991). When the expected change is smaller than the observed growth, the growth exceeds what is expected and may be attributed to policy changes. For example, in this study, using the two-year average of school reading performance to identify the lowest performing schools and selecting only schools with an average school reading performance of 93 or below decreases the expected change from 7.1 points to 3.7 points.[5] The expected change decreases even more as the number of baseline measurements increases (figure 11).

Using the two-year average of school reading performance to identify the lowest performing schools and selecting only schools with an average school reading performance of 93 or below decreases the expected change from 7.1 points to 3.7 points

Using multiple years of baseline data to identify the lowest performing schools could result in schools that are truly lowest performing being identified rather than schools that are having a poor year. Growth among schools identified that way would likely be beyond what was expected and could indicate true improvement in school reading performance related to implementation of the extended school day policy.

Figure 11. Expected change among the lowest performing schools decreases when multiple baseline measurements are used



Source: Authors' analysis of data from the Florida Department of Education school grades database (2012–14).

Multiple years of baseline data can also be advantageous for the analysis. For the current study, only the 2013/14 cohort has multiple years of pretreatment data using the same test each year. Those data could be used to capture growth among the lowest performing schools prior to implementing the extended school day policy.[6] The current study compares growth among the lowest performing schools with growth among other elementary schools. A stronger analysis could be to compare growth among the lowest performing schools prior to implementing the extended school day policy to growth during and after. There were insufficient data available for this study to conduct such analysis using consistent tests over time.[7] However, future research could standardize test scores from previous years to enable comparisons of data from before and after the tests changed, allowing for more years of baseline pre-treatment data that could be used to compare growth before and after implementation of the extended school day policy.

Limitations of the study

This study has six main limitations.

First, the study cannot separate changes in school reading performance that are due to student mobility from those that are due to changes in performance of the students who remain at the school. This would require student-level data, which were outside the scope of this study.

Second, the study cannot identify what aspects of policy implementation were effective. Because growth in school reading performance did not exceed expected change, further exploratory analysis of implementation and growth in school reading performance correlates in order to identify promising practices was not advisable. The only way to determine whether the extended school day policy caused an increase in school reading performance would be through an appropriate study design that allows for the study of causal relations (for example, a randomized controlled trial or regression discontinuity design). This may be a consideration for future research.

The only way to determine whether the extended school day policy caused an increase in school reading performance would be through an appropriate study design that allows for the study of causal relations (for example, a randomized controlled trial or regression discontinuity design)

Third, the survey data on implementation may suffer from inaccuracies. The Florida Department of Education requested survey data from the districts as a compliance check (see appendix B), but the surveys were not formally validated. While the surveys were intended to collect information using terminology familiar to Florida educators (for example, "comprehensive intervention program" in question 5), some questions may have been confusing. It may have been unclear to the respondents whether in questions 1–8 they were responding to the statement as written by the department (in the first column) or to the alignment of statutory requirements as written in House Bill 5101 (in the second column). For example, question 8 addresses the statutory requirement, "The intensive reading instruction delivered in this additional hour and for other students shall include: differentiated instruction based on student assessment data to meet students' specific reading needs," while the statement from the department reads, "Instruction must be differentiated and include: small group instruction based on student assessment data to meet students' specific reading needs and include formative assessment with writing, and extensive reading from a wide variety of texts, all verified with data logs" (see appendix B). Furthermore, some questions may seem contradictory. For example, question 1 asks about notification to all parents that all students will have an extended school day, but question 2 asks about providing an exemption to students performing above grade level.

In general, districts, not schools, completed the surveys. Thus, the respondents may not have known the nuances of how the extended school day policy was implemented at each school. And because the survey was administered as a compliance check, there is a high social desirability to have the "right" answers and make schools appear compliant with the policy. Without triangulating the survey data with other data such as interviews and observations, the implementation results are limited. Moreover, survey results were available only for schools that implemented the extended school day policy in 2013/14, so the study could not identify differences in implementation in 2012/13. Additionally, information related to 2012/13 was retrospective based on respondent recall.

Incorporating validated items on specific aspects of extended school day policy or context for implementation into the surveys could provide more accurate data. Validation would also enable the Florida Department of Education to examine whether the survey items for each construct validly measure what they are being used to measure and to what extent they are internally consistent.

The study did not measure quality of instruction

Fourth, the study did not measure quality of instruction. Some research suggests that the quality of instruction rather than time allotted is what makes a difference in student achievement (see, for example, Connor et al., 2009; Guthrie et al., 2004; Hamre & Pianta, 2005; Moats, Foorman, & Taylor, 2006). Surveys provide limited information on day-to-day operations at a school and limited insight into actual instructional practices. The surveys captured information about the extended hour only. The lowest performing schools might have substantially changed their entire reading instruction program beyond simply adding an additional hour. Future research could conduct in-depth classroom observations during both the regular school day and the extra hour.

Fifth, it was not possible to determine whether the lowest performing schools were doing anything different from other elementary schools. Implementation data were provided only by schools identified as the lowest performing schools in 2013/14. Data were unavailable for other elementary schools, which, for example, may have implemented similar extended hour policies or found alternative ways to increase reading instruction. This is particularly possible because House Bill 5101 was worded such that reading coaches could be added to the staff only if all students were provided an additional hour of reading instruction.

Sixth, the estimates of expected change may be biased. The distribution of school reading performance is not perfectly normally distributed; while both skew and kurtosis were very small and within acceptable limits, they were statistically significant (see appendix C).[8] And while expected change is supposed to estimate what would happen in the absence of policy change, the study's estimate of expected change was based on data covering the period when the policy was in place, though the study team obtained similar results when calculating expected change using only data not affected by the policy.[9] While the expected change estimates may be biased, the subanalyses provide support for the overall findings. Nonetheless, it may be worthwhile for future research to estimate expected change in ways that do not require an assumption of normality—for example, by looking at multiple years of data prior to policy implementation and standardizing results as needed to allow for comparisons across years when different tests were given.

Appendix A. Supplemental tables of school characteristics, school reading performance, and survey responses

This appendix provides descriptive statistics for all variables included in this report. Table A1 provides descriptive statistics of school characteristics and school reading performance. Table A2 provides correlations and descriptive statistics of the subcomponents of school reading performance and enrollment. Table A3 provides aggregated and disaggregated descriptive statistics of minutes of instruction provided as reported in the implementation survey. Table A4 provides aggregated and disaggregated descriptive statistics from the implementation survey.

Table A1. Descriptive statistics of school characteristics and school reading performance

Variable	Full sample	Other elementary schools	Lowest performing schools in 2012/13 only	Lowest performing schools in 2013/14 only	Lowest performing schools in both years	Closed schools
Number of schools						
Number	1,737	1,578	68	66	25	12
Percent	100	90.2	3.9	3.8	1.4	0.7
Schools in region 1						
Number	137	124	9	2	2	2
Percent	100	89.2	6.5	1.4	1.4	1.4
Schools in region 2						
Number	336	303	13	12	8	2
Percent	100	89.6	3.8	3.6	2.4	0.6
Schools in region 3						
Number	346	331	9	6	0	1
Percent	100	95.4	2.6	1.7	0.0	0.3
Schools in region 4						
Number	416	384	9	15	8	1
Percent	100	92.1	2.2	3.6	1.9	0.2
Schools in region 5						
Number	502	436	28	31	7	6
Percent	100	85.8	5.5	6.1	1.4	1.2
Percentage of schools that are Title I schools						
Number	1,134	975	68	66	25	na
Percent	65.3	61.8	100	100	100	na
Enrollment in 2011/12						
Mean	304.0	313.0	206.5	238.2	191.2	na
Standard deviation	119.5	119.3	80.9	83.2	53.6	na
Range	16–970	16–970	58–432	62–473	98–294	na
Enrollment in 2012/13						
Mean	300.0	308.8	197.3	237.3	191.6	na
Standard deviation	117.1	116.8	76.6	83.2	59.4	na
Range	24–963	24–963	47–387	57–471	104–304	na
Enrollment in 2013/14						
Mean	301.2	310.1	202.2	231.2	188.8	na
Standard deviation	117.3	116.6	83.3	84.0	59.0	na
Range	10–961	10–961	25–416	48–462	88–288	na
Percentage of racial/ethnic minority students in the school						
Mean	60.5	57.6	88.0	90.0	93.6	na

(continued)

Table A1. Descriptive statistics of school characteristics and school reading performance *(continued)*

Variable	Full sample	Other elementary schools	Lowest performing schools in 2012/13 only	Lowest performing schools in 2013/14 only	Lowest performing schools in both years	Closed schools
Standard deviation	27.2	26.4	17.9	12.9	6.0	na
Range	4.3–100	4.3–100	15.3–100	44.3–100	78.3–100	na
Percentage of students eligible for the school lunch program						
Mean	67.0	64.4	91.6	93.6	94.6	na
Standard deviation	5.5	23.2	8.7	6.0	5.5	na
Range	0–100	0–100	48–100	72–100	79–100	na
Percentage of students at the school reading at or above achievement level 3 (satisfactory) on the FCAT 2.0 in 2011/12						
Mean	59.1	61.8	31.0	37.2	27.7	na
Standard deviation	15.7	13.8	4.9	6.2	5.8	na
Range	18–100	29–100	18–42	25–59	19–40	na
Percentage of students at the school reading at or above achievement level 3 (satisfactory) on the FCAT 2.0 in 2012/13						
Mean	58.4	60.9	38.1	31.3	26.5	na
Standard deviation	15.6	13.9	6.6	5.1	5.0	na
Range	17–99	30–99	27–58	19–44	17–36	na
Percentage of students at the school reading at or above achievement level 3 (satisfactory) on the FCAT 2.0 in 2013/14						
Mean	58.8	61.2	37.0	35.1	27.4	na
Standard deviation	6.7	14.5	8.2	7.0	8.7	na
Range	13–48	23–99	21–59	18–50	13–48	na
Percentage of students at the school making learning gains from the prior year on the reading component of the FCAT 2.0 in 2011/12						
Mean	69.3	70.3	54.6	65.0	56.9	na
Standard deviation	8.0	7.3	6.9	5.4	4.8	na
Range	32–100	44–100	32–66	57–80	48–66	na
Percentage of students at the school making learning gains from the prior year on the reading component of the FCAT 2.0 in 2012/13						
Mean	65.4	66.2	66.9	51.0	53.5	na
Standard deviation	7.8	7.2	8.1	5.4	5.9	na
Range	38–91	45–91	58–88	38–61	42–64	na
Percentage of students at the school making learning gains from the prior year on the reading component of the FCAT 2.0 in 2013/14						
Mean	58.8	69.0	58.9	64.5	57.6	na
Standard deviation	16.0	8.1	9.6	8.6	10.4	na
Range	13–99	28–93	38–87	48–87	41–79	na
School reading performance in 2011/12						
Mean	128.4	132.1	85.7	102.2	84.6	na
Standard deviation	21.5	18.7	7.3	8.0	7.0	na
Range	58–195	94–195	58–93	93–128	71–93	na
School reading performance in 2012/13						
Mean	123.8	127.1	105.0	82.3	80.0	na
Standard deviation	21.4	19.1	12.2	6.9	6.9	na
Range	64–182	91–182	91–138	64–90	65–89	na
School reading performance in 2013/14						
Mean	127.0	130.1	96.0	99.6	85.0	na
Standard deviation	22.7	20.9	14.8	13.5	16.9	na
Range	54–185	63–185	65–136	70–133	54–116	na

FCAT 2.0 is Florida Comprehensive Assessment Test 2.0. na is not applicable.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2013a, 2014a) and from School Membership by Grade files (Florida Department of Education, 2012c, 2013c, 2014d).

Table A2. Correlations among school reading performance components and enrollment

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Percentage of students reading at or above achievement level 3 (satisfactory) on the FCAT 2.0 (2011/12)	1											
2. Percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0 (2011/12)	.60**	1										
3. School reading performance (2011/12)	.96**	.81**	1									
4. Enrollment (2011/12)	.31**	.22**	.31**	1								
5. Percentage of students reading at or above achievement level 3 (satisfactory) on the FCAT 2.0 (2012/13)	.95**	.55**	.90**	.30**	1							
6. Percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0 (2012/13)	.50**	.37**	.51**	.15**	.62**	1						
7. School reading performance (2012/13)	.88**	.54**	.84**	.27**	.96**	.82**	1					
8. Enrollment (2012/13)	.33**	.24**	.33**	.98**	.31**	.15**	.28**	1				
9. Percentage of students reading at or above achievement level 3 (satisfactory) on the FCAT 2.0 (2013/14)	.93**	.55**	.89**	.30**	.95**	.56**	.90**	.32**	1			
10. Percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0 (2013/14)	.59**	.47**	.60**	.20**	.55**	.37**	.54**	.21**	.68**	1		
11. School reading performance (2013/14)	.88**	.57**	.85**	.29**	.88**	.54**	.84**	.30**	.96**	.86**	1	
12. Enrollment (2013/14)	.34**	.25**	.34**	.95**	.33**	.16**	.30**	.98**	.32**	.20**	.30**	1

** Significant at $p < .01$.

FCAT 2.0 is the Florida Comprehensive Assessment Test 2.0.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2013a, 2014a) and from School Membership by Grade files (Florida Department of Education, 2012c, 2013c, 2014d).

Table A3. Descriptive statistics of amount of instruction provided

Variable	All lowest performing schools in 2013/14	Lowest performing schools in 2013/14 only	Lowest performing schools in both years
Minutes of reading instruction provided in 2012/13			
Number of schools reporting	85	63	22
Mean	133.9	118.7	177.3
Standard deviation	36.8	27.0	24.3
Range	90–240	90–210	120–240
Minutes of reading instruction provided in 2013/14			
Number of schools reporting	82	61	21
Mean	179.2	177.5	184.3
Standard deviation	21.4	21.8	19.6
Range	150–250	150–250	150–240
Change in minutes of reading instruction provided in 2012/13 to 2013/14[a]			
Number of schools reporting	82	61	21
Mean	46.3	59.8	7.1
Standard deviation	34.0	26.9	18.7
Range	0–160	0–160	0–60

a. Values are for schools with data for both years.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

Table A4. Frequencies of implementation responses from the implementation survey

Question	Response	All schools identified as the lowest performing schools in 2013/14		Lowest performing schools in 2013/14 only		Lowest performing schools in both years	
		Number	Percent	Number	Percent	Number	Percent
I. Student enrollment							
1. Provided parents notification that all students will have an added hour of reading instruction each day (*n* = 86)	Does not meet criterion	0	0.0	0	0.0	0	0.0
	Partially meets criterion	3	3.5	3	4.8	0	0.0
	Meets criterion	59	68.6	42	66.7	17	73.9
	Exceeds criterion	24	27.9	18	28.6	6	26.1
2. Provided flexibility so that eligible students that scored Florida Comprehensive Assessment Test Reading Level 5 may choose to participate in one hour of extended reading instruction each day, but are not required to do so. (*n* = 82)	Does not meet criterion	3	3.7	2	3.3	1	4.5
	Partially meets criterion	2	2.4	2	3.3	0	0.0
	Meets criterion	67	81.7	47	78.3	20	90.9
	Exceeds criterion	10	12.2	9	15.0	1	4.5
II. Teachers							
3. Evidence of prior success teaching reading to struggling readers as indicated by various assessment data and student work samples over time. Teachers should be highly effective K–5 teachers who are Reading Certified or Endorsed. (*n* = 86)	Does not meet criterion	0	0.0	0	0.0	0	0.0
	Partially meets criterion	12	14.0	10	15.9	2	8.7
	Meets criterion	69	80.2	49	77.8	20	87.0
	Exceeds criterion	5	5.8	4	6.3	1	4.3
III. Reading program/materials							
4. A research-based sequence of intensive reading instruction, including instructional routines that have been proven to accelerate progress of students exhibiting a reading deficiency and the integration of social studies, science, and mathematics-text reading, text discussion, and writing in response to reading. (*n* = 86)	Does not meet criterion	1	1.2	1	1.6	0	0.0
	Partially meets criterion	2	2.3	2	3.2	0	0.0
	Meets criterion	66	76.7	47	74.6	19	82.6
	Exceeds criterion	17	19.8	13	20.6	4	17.4

(continued)

Table A4. Frequencies of implementation responses from the implementation survey *(continued)*

Question	Response	All schools identified as the lowest performing schools in 2013/14		Lowest performing schools in 2013/14 only		Lowest performing schools in both years	
		Number	Percent	Number	Percent	Number	Percent
5. One or more of the following types of research-based curricula designed to support and accelerate student development in the components of learning to read: oral language, including vocabulary; phonological awareness and phonics, and text reading that supports accuracy, fluency, and comprehension with more extensive opportunities for guided practice, error correction, and feedback *Comprehensive intervention program; *Targeted intervention program; *Supplemental reading program, including technology. (*n* = 75)	Does not meet criterion	1	1.3	1	1.8	0	0.0
	Partially meets criterion	0	0.0	0	0.0	0	0.0
	Meets criterion	54	72.0	39	69.6	15	78.9
	Exceeds criterion	20	26.7	16	28.6	4	21.1
IV. Assessment							
6. Use of assessment and progress monitoring results that identify the overall learning needs of each student in order to distinguish whether their instructional priority is decoding (phonemic awareness, phonics, fluency) or text meaning (comprehension, vocabulary). This data should be used to initially place the student and plan for instruction. (*n* = 83)	Does not meet criterion	1	1.3	1	1.7	0	0.0
	Partially meets criterion	0	0.0	0	0.0	0	0.0
	Meets criterion	63	75.9	44	73.3	19	82.6
	Exceeds criterion	19	22.9	15	25.0	4	17.4
7. Regular monitoring (may be informal) of students' progress and subsequent adjustment of instruction, as needed. (*n* = 84)	Does not meet criterion	1	1.2	1	1.6	0	0.0
	Partially meets criterion	2	2.4	1	1.6	1	4.3
	Meets criterion	66	78.6	48	78.7	18	78.3
	Exceeds criterion	15	17.9	11	18.0	4	17.4
V. Instruction							
8. Instruction must be differentiated and include: small group instruction based on student assessment data to meet students' specific reading needs and include formative assessment with writing, and extensive reading from a wide variety of texts, all verified with data logs. (*n* = 84)	Does not meet criterion	1	1.2	1	1.6	0	0.0
	Partially meets criterion	2	2.4	2	3.2	0	0.0
	Meets criterion	65	77.4	46	74.2	19	86.4
	Exceeds criterion	16	19.0	13	21.0	3	13.6
VI. Staff details							
9a. How are you staffing the extra hour of reading instruction? (*n* = 48)	Other staff used for the extra hour of reading instruction	6	12.5	6	17.1	0	0.0
	The students' regular classroom teachers are providing the extra hour of reading instruction	14	29.2	10	28.6	4	30.8
	Both	28	58.3	19	54.3	9	69.2
9b. Staffing changes: Reading coaches (*n* = 62)	Less	2	3.2	1	2.1	1	6.7
	No change	25	40.3	16	34.0	9	60.0
	More	35	56.5	30	63.8	5	33.3
9b. Staffing changes: Teachers (*n* = 76)	Less	5	6.6	2	3.7	3	13.6
	No change	16	21.1	8	14.8	8	36.4
	More	55	72.4	44	81.5	11	50.0
9b. Staffing changes: Paraprofessionals (*n* = 57)	Less	5	8.8	4	9.8	1	6.2
	No change	35	61.4	23	56.1	12	75.0
	More	17	29.8	14	34.1	3	18.8

(continued)

Table A4. Frequencies of implementation responses from the implementation survey *(continued)*

Question	Response	All schools identified as the lowest performing schools in 2013/14		Lowest performing schools in 2013/14 only		Lowest performing schools in both years	
		Number	Percent	Number	Percent	Number	Percent
9b. Staffing changes: Volunteers (*n* = 47)	Less	0	0.0	0	0.0	0	0.0
	No change	41	87.2	32	91.4	9	75.0
	More	6	12.8	3	8.6	3	25.0
9b. Staffing changes: Other staff (*n* = 54)	Less	7	13.0	3	7.3	4	30.8
	No change	20	37.0	14	34.1	6	46.2
	More	27	50.0	24	58.5	3	23.1
VII. Instructional time							
11. How did your school add the required extra hour of reading instruction to the school day? (*n* = 85)	Extended the bell schedule at the beginning of the day	4	4.7	3	4.8	1	4.3
	Extended the bell schedule at the end of the day	32	37.6	24	38.7	8	34.8
	Rearranged the instructional day (provided additional instruction at a different time of day)	12	14.1	9	14.5	3	13.0
	Combination: Extended the bell schedule at the beginning and end of the day	2	2.4	1	1.6	1	4.3
	Combination: Extended the bell schedule at the end of the day and rearranged the instructional day	34	40.0	25	40.3	9	39.1
	Combination: Extended the bell schedule at the beginning and end of the day, and rearranged the instructional day	1	1.2	0	0.0	1	4.3
VIII. Student grouping							
13. Are students grouped differently during the extra hour than during the regular instructional block? (*n* = 61)	No	7	11.5	7	14.6	0	0.0
	Yes	33	54.1	25	52.1	8	61.5
	Unclear	21	34.4	16	33.3	5	38.5
14. How are students grouped during the extra hour of reading instruction? (*n* = 86)	Whole class/large group only (*n* = 83)	3	3.6	2	3.2	1	4.3
	Small group/individual only (*n* = 83)	23	27.7	18	28.6	5	21.7
	Both small group/individual and whole class/large group (*n* = 83)	57	68.7	41	65.1	16	69.6
	Homogeneous only (*n* = 82)	46	56.1	32	50.8	14	60.9
	Heterogeneous only (*n* = 82)	6	7.3	6	9.5	0	0.0
	Both homogeneous and heterogeneous (*n* = 82)	30	36.6	22	34.9	8	34.8

Note: Data are for 86 schools. Headers, questions, and question numbers match the survey provided in appendix B. Questions 10 and 12 are omitted because they are not conducive to table format. Percentages may not sum to 100 because of rounding.

Source: Authors' analysis of data from district extended school day reading instruction plans obtained by special request from the Florida Department of Education in 2014.

Appendix B. Florida Department of Education 100 lowest performing elementary schools in reading extended school day reading instruction plans compliance survey

This appendix provides the compliance survey distributed to districts that had one of the 100 lowest performing elementary schools. The Florida Department of Education, through the Just Read, Florida! office, distributed this compliance survey electronically along with directions for completion and return. This survey was distributed, completed, and returned in early 2014.

District's Extended Day Reading Instruction Plans						
SCHOOL INFORMATION						
School Name:						
School Number:						
DISTRICT INFORMATION						
District Name:						
District Number:						
District Contact Name:						
Contact Phone Number:						
Contact Email:						
Please use the checklist to rate the plan's compliance with the numbered criteria according to this scale: 0-does not meet criterion, 1-partially meets criterion, 2-meets criterion, 3-exceeds criterion. Enter comments only if needed to clarify rating.						
I. Student Enrollment						
Extended Day Reading Instruction Plans A plan has been implemented that:	Alignment of Statutory Requirements	0	1	2	3	District Review Comments
1. Provided parents notification that all students will have an added hour of reading instruction each day. (evidence attached)	For the 2012–2013 and 2013–2014 fiscal years, in each school district that has one or more of the 100 lowest performing elementary schools based on the state reading assessment, priority shall be given to providing an additional hour per day of intensive reading instruction beyond the normal school day for each day of the entire school year for the students in each school.	☐	☐	☐	☐	
Extended Day Reading Instruction Plans A plan has been implemented that:	Alignment of Statutory Requirements	0	1	2	3	District Review Comments
2. Provided flexibility so that eligible students that scored Florida Comprehensive Assessment Test Reading Level 5 may choose to participate in one hour of extended reading instruction each day, but are not required to do so. (evidence attached)	Students enrolled in these schools who have level 5 assessment scores may participate in the additional hour of instruction on an optional basis.	☐	☐	☐	☐	

District's Extended Day Reading Instruction Plans						
II. Teachers						
The district plan includes selection of highly qualified teachers who have:		0	1	2	3	District Review Comments
3. Evidence of prior success teaching reading to struggling readers as indicated by various assessment data and student work samples over time. Teachers should be highly effective K-5 teachers who are Reading Certified or Endorsed.	This additional hour of instruction must be provided only by teachers or reading specialists who are effective in teaching reading.	☐	☐	☐	☐	
III. Reading Program/Materials						
The design of the local reading program and the plan for reading instruction/ intervention includes:		0	1	2	3	District Review Comments
4. A research-based sequence of intensive reading instruction, including instructional routines that have been proven to accelerate progress of students exhibiting a reading deficiency and the integration of social studies, science, and mathematics-text reading, text discussion, and writing in response to reading.	The intensive reading instruction delivered in this additional hour and for other students shall include: research-based reading instruction that has been proven to accelerate progress of students exhibiting a reading deficiency; the integration of social studies, science, and mathematics-text reading, text discussion, and writing in response to reading.	☐	☐	☐	☐	
The design of the local reading program and the plan for reading instruction/ intervention includes:		0	1	2	3	District Review Comments
5. One or more of the following types of research-based curricula designed to support and accelerate student development in the components of learning to reading: oral language, including vocabulary; phonological awareness and phonics, and text reading that support accuracy, fluency, and comprehension with more extensive opportunities for guided practice, error correction, and feedback. *Comprehensive intervention program; *Targeted intervention program; *Supplemental reading program, including technology.	The intensive reading instruction delivered in this additional hour and for other students shall include explicit and systematic reading development in phonemic awareness, phonics, fluency, vocabulary, and comprehension, with more extensive opportunities for guided practice, error correction, and feedback.	☐	☐	☐	☐	

District's Extended Day Reading Instruction Plans							
IV. Assessment							
The plan for assessing students includes:			0	1	2	3	District Review Comments
6. Use of assessment and progress monitoring results that identify the overall learning needs of each student in order to distinguish whether their instructional priority is decoding (phonemic awareness, phonics, fluency) or text meaning (comprehension, vocabulary). This data should be used to initially place the student and plan for instruction.	The intensive reading instruction delivered in this additional hour and for other students shall include: differentiated instruction based on student assessment data to meet students' specific reading needs.		☐	☐	☐	☐	
7. Regular monitoring (may be informal) of students' progress and subsequent adjustment of instruction, as needed.	The intensive reading instruction delivered in this additional hour and for other students shall include: differentiated instruction based on student assessment data to meet students' specific reading needs.		☐	☐	☐	☐	
V. Instruction							
The district must create a reading schedule that facilitates interactive, challenging instruction.			0	1	2	3	District Review Comments
8. Instruction must be differentiated and include: small group instruction based on student assessment data to meet students' specific reading needs and include formative assessment with writing, and extensive reading from a wide variety of texts, all verified with data logs.	The intensive reading instruction delivered in this additional hour and for other students shall include: differentiated instruction based on student assessment data to meet students' specific reading needs.		☐	☐	☐	☐	
VI. Staff Details							
9. How are you staffing the extra hour of reading instruction?							District Review Comments
A. Select all that apply	The student's regular classroom teachers are providing the instruction			☐			
	Other staff used for the extra hour of reading instruction			☐			
B. Please check who the other staff are for the extra hour of reading instruction and indicate how many serve your school. Indicate below the number of staff for each category, year and total	Staff Type	Staff used to provide extra hour instruction	Last year	Additional this year	Total		District Review Comments
	Reading Coaches	☐					
	Teachers	☐					
	Paraprofessionals	☐					
	Volunteers	☐					
	Other	☐					
10. If students' regular teachers do not provide the extra hour of instruction, what are the teachers doing during the extra hour?							

District's Extended Day Reading Instruction Plans					
VII. Instructional Time					
11. How did your school add the required extra hour of reading instruction to the school day? *(Select all that apply)*	Extended the bell schedule at the beginning of the day	Extended the bell schedule at the end of day	Rearranged the instructional day (provided additional instruction at a different time of day)	Other (please explain)	District Review Comments
Please select all that apply by placing an X or comment in the place below the strategies applied at your school	☐	☐	☐	☐	
12. How many minutes of reading instruction are provided during the following blocks per day for an average student in your school?	Standard reading block—this does not include the extra hour (e.g., Tier 1 instruction) Intervention block (e.g., Tier 2 instruction)	Other reading related block including additional hour	Total Reading instruction per day for 2013–14	Total Reading instruction per day for 2012–13	District Review Comments
Please report for all categories that apply					
VIII. Student Grouping					
13. Are students grouped differently during the extra hour than during the regular instructional block?					
14. How are students grouped during the extra hour of reading instruction?	Grouping	Check all that apply	Approximate group size	District Review Comments	
	whole class/large group	☐			
	small group/individual	☐			
	students are grouped homogeneously according to ability/needs	☐			
	students are grouped heterogeneously	☐			

Appendix C. Data and methodology

This appendix describes the data and methodology used in the study.

Data

Study sample. The study sample includes only schools eligible to be selected as a lowest performing school (that is, they were open, general education elementary schools receiving school grades[10]) in all years between 2011/12 and 2013/14. Schools with a special school function or setting (such as virtual, hospital/homebound, and Department of Juvenile Justice schools) were excluded.

Lowest performing schools. The Florida Department of Education, through the Improving Literacy Alliance and Improving Low Performing Schools Alliance, provided the study team with a list of the lowest performing elementary schools in reading for each year of the study.[11] Alliance members also explained how the Florida Department of Education identified the lowest performing schools based on school reading performance, which is the sum of the percentage of students scoring at or above achievement level 3 (satisfactory) on the reading component of the prior year's Florida Comprehensive Assessment Test 2.0 (FCAT 2.0) at each school and the percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0 at each school.

Publicly available data. The list of 100 lowest performing schools each year was matched to publicly available School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a), Master School Identification files (Florida Department of Education, 2012b, 2013b, 2014c), and School Membership by Grade files (Florida Department of Education, 2012c, 2013c, 2014d). All data, including student achievement, demographic, and contextual information, are publicly available either on the Florida Department of Education website or through public records request.

School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a) are maintained by the Florida Department of Education to provide detailed information about student achievement and other components used to calculate school grades. They provided the following information pertinent to the study:

- District and school identifiers (used to link files).
- Percentage of students reading at the satisfactory level or higher (used to compute school reading performance).
- Percentage of students making learning gains in reading (used to compute school reading performance).
- School type (used to select schools for analysis).
- School grades (used to select schools for analysis).
- Percentage of racial/ethnic minority students (used to describe school composition).
- Percentage of students eligible for the school lunch program (used to describe school composition).
- Title I status (used to describe the schools).
- Region (used to describe the distribution of the lowest performing schools).

The Master School Identification files (Florida Department of Education, 2012b, 2013b, 2014c) are maintained by the Florida Department of Education to ensure accurate identification and directory information on all Florida public schools. These files were matched to the

School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a) and School Membership by Grade files (Florida Department of Education, 2012c, 2013c, 2014d) to verify the school type, region, and Title I status from the School Accountability Reports. The Master School Identification files provided the following information pertinent to the study:

- District and school identifiers (used to link files).
- School type, accountability type, and grade configuration variables (used to verify that only elementary schools were included in the analysis).
- Primary service type and school function/setting variables (used to verify that only K–12 general education schools were included in the analysis).
- Activity code (used to determine whether a school was active or closed).

To determine a school's status as closed or restructured, the study team consulted the Florida Master School Identification files (Florida Department of Education, 2012b, 2013b, 2014c). In the files, school number did not change across years, and all schools were flagged as either active or closed. Each year, the grade configuration (for example, K–2) and school type (for example, elementary) were updated in the file as well. Schools whose activity codes or grade configurations changed between years were identified as closed or restructured and excluded from the study.

The School Membership by Grade files (Florida Department of Education, 2012c, 2013c, 2014d) are maintained and updated yearly by the Florida Department of Education and provide total student enrolment by grade and across grades for all Florida schools. These files were matched to the School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a). They provided the following information pertinent to the study:

- District and school identifiers (used to link files).
- Number of students in the school.

School reading performance. The list of the 100 lowest performing schools provided by the Florida Department of Education included the school reading performance value used to identify the 100 lowest performing schools. The study team computed school reading performance each year for other schools included in the study based on the same methodology using SPSS (IBM Corp., 2012). School reading performance has a theoretical range of 0–200, but the observed range in this study was 54–195.

Extended school day survey. The Just Read, Florida! office of the Florida Department of Education, which oversees implementation of the extended school day policy, administered a compliance survey to district personnel in 2013/14 to gather information on how the lowest performing schools were implementing the extended hour. In some cases, the district personnel completed the surveys on behalf of the schools; in other cases the schools themselves completed the surveys. It is unknown which surveys were completed by district versus school personnel.

Neither the study team nor the state piloted or validated the survey—that is, neither group verified or otherwise triangulated the accuracy of the information provided in the surveys. It is unknown whether school personnel had input or otherwise verified the responses when district personnel completed the surveys or whether district personnel had input or otherwise verified the responses when school personnel completed the surveys.

The Florida Department of Education provided the original surveys as electronic spreadsheets for each school. The study team double-blind entered the information from the surveys into a database for analysis.

Initial descriptive analysis. Preliminary analyses suggested that changes in school reading performance might not be associated with implementation of the extended school day policy. These analyses include study team visual inspections of the distribution of school reading performance each year. The distributions appeared similar across all years (figure C1). If there were a systematic influence, the distribution of school reading performance would be normal only in 2011/12, when the extended school day policy was not implemented, and school reading performance in subsequent years would not be normally distributed because the lower tail would shrink.

Analysis of the skewness and kurtosis values also suggested that the distributions were similar across years. Skewness ranged from –.12 to –.02, and kurtosis ranged from –.47 to –.40. These estimates were within the standard error (.06 for skewness and .12 for kurtosis) of the median skewness (–.10) and kurtosis (–.42).

Regression to the mean (or expected change). To calculate the effect of regression to the mean, referred to as expected change in the main text, the study team followed the procedures set forth by Barnett et al. (2005). Regression to the mean can be evaluated when data are normally distributed. The formula to calculate regression to the mean is defined by Davis (1976) and Gardner and Heady (1973) as:

$$\begin{aligned} &\textit{Regression to the mean} \\ &= \frac{\sigma_w^2}{\sqrt{\sigma_w^2 + \sigma_b^2}} C(z) \\ &= \sigma_t(1-\rho)C(z), -1 \le \rho \le 1 \end{aligned} \tag{C1}$$

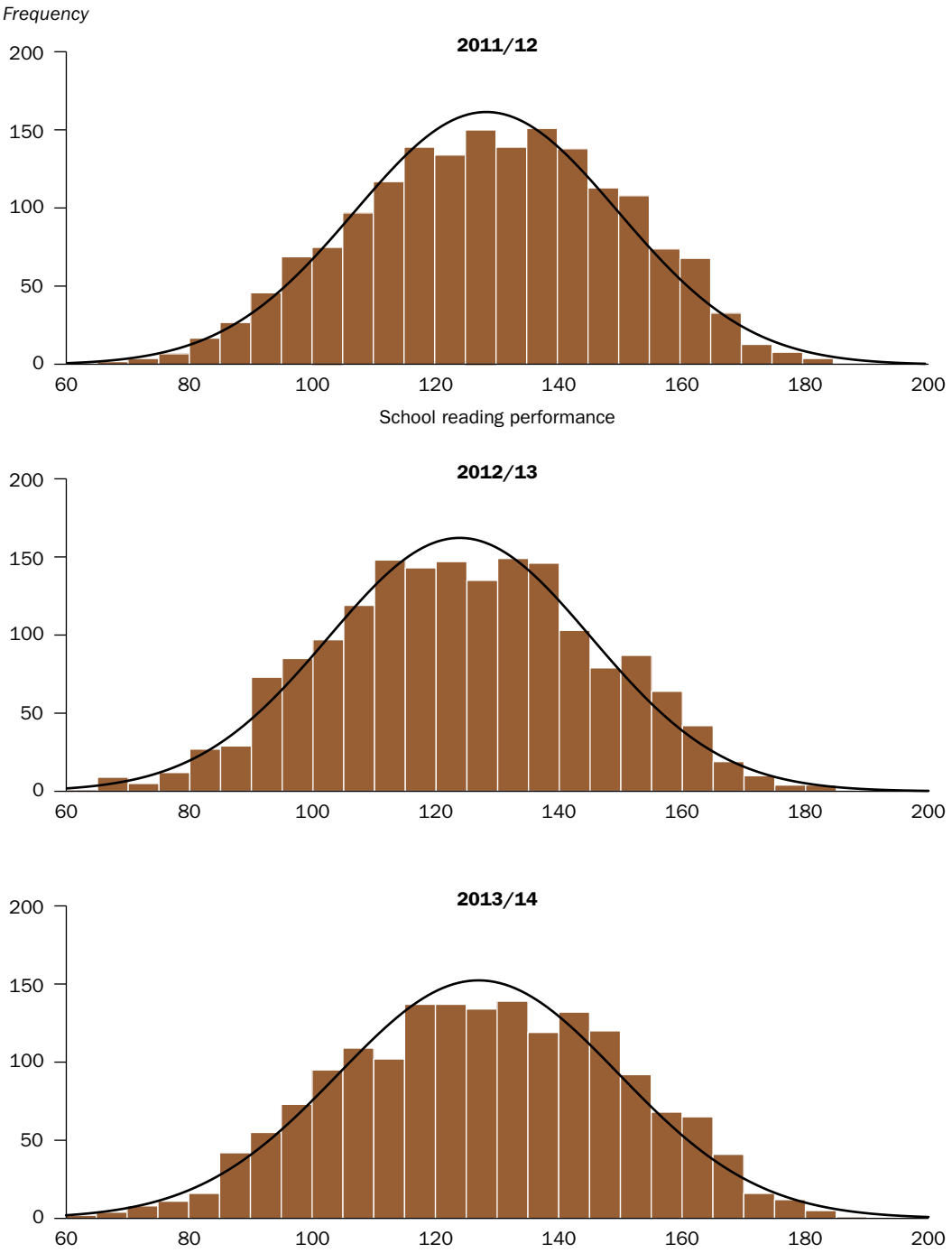
where $\sigma_t^2 = \sigma_w^2 + \sigma_b^2$ is the total variance, $\sigma_w^2 = (1-\rho)\sigma_t^2$ is the within-subject (school) variance, $\sigma_b^2 = \rho\sigma_t^2$ is the between-subject (school) variance, ρ is the correlation between measurement time one and two, and $C(z) = \frac{\phi(z)}{\Phi(z)}$, where $z = \frac{(\mu - c)}{\sigma_t}$ if the subjects (schools) are selected using a baseline measurement less than c and μ is the population mean.[12] $\phi(z)$ is the probability density of the standard normal distribution, and $\Phi(z)$ is the cumulative distribution function of the standard normal distribution.

As variability in a measure increases, so does the estimate of the effect of regression to the mean (Barnett et al., 2005; Davis, 1976; Johnson & George, 1991). When multiple baseline measures are used to determine group membership, the estimated effect of regression to the mean is:

$$\begin{aligned} &\textit{Regression to the mean} \\ &= \frac{\frac{\sigma_w^2}{m}}{\sqrt{\frac{\sigma_w^2}{m} + \sigma_b^2}} C(z) \end{aligned} \tag{C2}$$

where m is the number of baseline measurements and σ_w^2, σ_b^2, and $C(z)$ are as in equation C1.

Figure C1. Distribution of school reading performance within years



Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

Extending the regression to the mean equations. The regression to the mean equations can be extended in a multilevel framework, where $\sigma_b^2 = \sigma_{r0}^2 + \sigma_{u00}^2$, or the between variance is the sum of the between-school variance and the between-district variance; $\sigma_t^2 = \sigma_e^2 + \sigma_{r0}^2 + \sigma_{u00}^2$, or the total variance is the sum of the within-school, the between-school, and the between-district variances; and $p = \frac{\sigma_e^2}{\sigma_t^2}$, or the interclass correlation from a null multilevel model. When using a multilevel framework to obtain the values to calculate

the effect of regression to the mean, the intercept from the null model is considered the mean for use in calculating z.

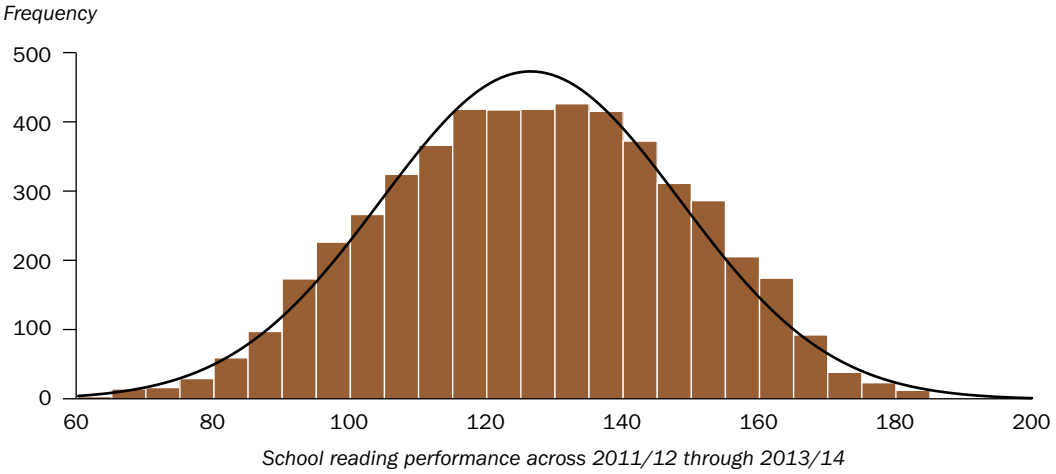
To increase the certainty around the estimates of the effect of regression to the mean, confidence intervals can be calculated by bootstrapping (Linden, 2013). For each bootstrapped estimation of the null model, a new estimate of the effect of regression to the mean is produced. The bootstrapped results can then be used to calculate confidence intervals around the estimate.

The estimate of the effect of regression to the mean and its confidence intervals can be used to compare estimated growth in school reading performance from multilevel growth models to determine whether average growth is within the confidence interval. If the confidence interval of estimated growth in school reading performance does not overlap with the confidence interval of the effect of regression to the mean, the average growth in school reading performance is beyond what would have been expected based on the effect of regression to the mean. However, if the confidence interval of the effect of regression to the mean and the confidence interval of estimated growth in school reading performance overlap, it can be concluded that, on average, growth was roughly what would have been expected because of the effect of regression to the mean.

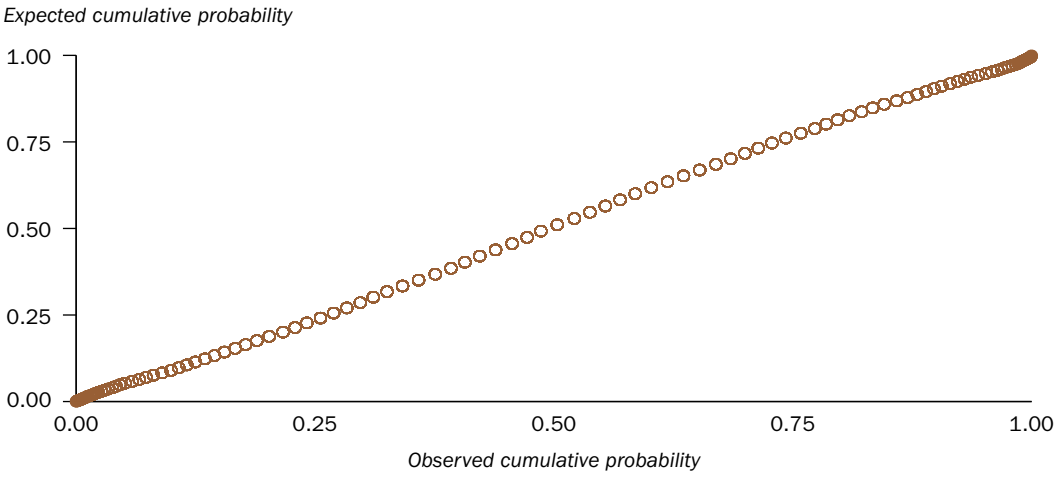
Applying the regression to the mean formula. To determine whether the data were normally distributed—a necessary condition to use the formula for the effect of regression to the mean—the study team inspected the distribution of the school reading performance across all years. That was the primary variable of interest because that variable is entered in the analysis of regression to the mean. Skewness was –0.08 (standard error = .03), and kurtosis was –0.44 (standard error = .07)—only slight deviations from normality but within generally accepted limits, suggesting that many statistics based on these data will suffer from relatively low bias (George & Mallery 2010; Kim, 2013; Tabachnick & Fidell, 2013). With very large sample sizes, as in this report, normality tests can detect very small but statistically different from zero differences from normality. Thus, absolute values of skewness and kurtosis were inspected rather than critical values obtained by dividing the skewness and kurtosis estimates by the standard error. Additionally, visual inspection of the normal distribution, quantile–quantile plot, and probability–probability plots are considered helpful with large sample sizes.[13] Visual inspection also supported a sufficiently normal distribution of school reading performance (figure C2).

After ensuring that the deviations from normality were below commonly accepted standards, the study team estimated the effect of regression to the mean to be 7.05 points using the procedures described above with the intercept and variance components from the null mixed model of school reading performance. Because there was not a set school reading performance cutpoint to identify the lowest performing schools, the study team inferred a cutpoint as the highest school reading performance among the lowest performing schools at the year of identification. The highest school reading performance among the lowest performing schools was 90 points for schools identified as the lowest performing schools in 2012/13 and 93 points for schools identified as the lowest performing schools in 2013/14 (see table A1 in appendix A). The study team used the higher of the cutpoints, 93, because

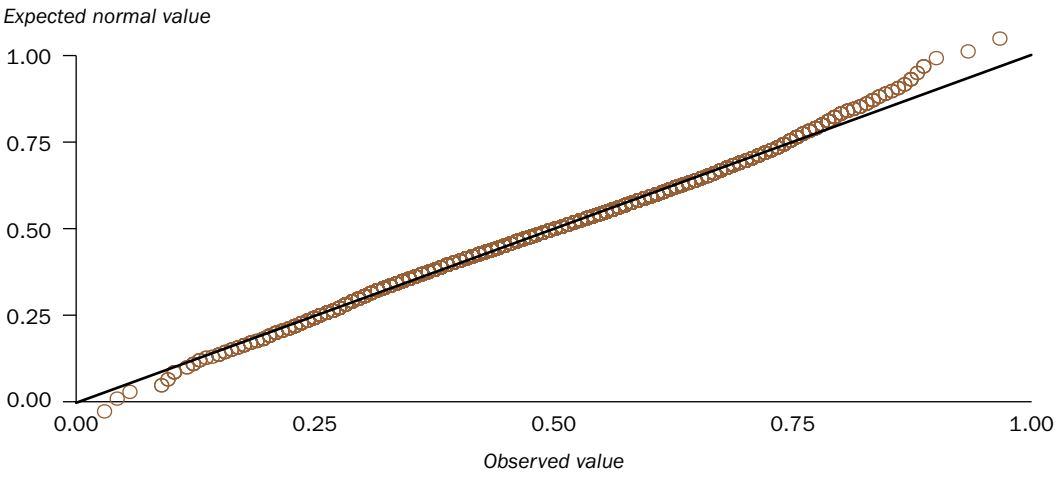
Figure C2. Distribution of school reading performance across years



Normal probability–probability plot of school reading performance across 2011/12 through 2013/14



Normal quantile–quantile plot of school reading performance across 2011/12 through 2013/14



Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

it produced a smaller estimate of the effect of regression to the mean.[14] The study team applied the following values to calculate the effect of regression to the mean:

- $\sigma_b^2 = 416.79$.
- $\sigma_w^2 = 80.87$.
- $\sigma_t = 22.31$.
- $\rho = .84$.
- $\mu = 126.63$.
- $c = 93$.
- $z = 1.51$.
- $\phi(z) = .13$.
- $\Phi(z) = .07$.
- $C(z) = 1.94$.

The study team calculated the effect of regression to the mean in the same way described above on 1,000 bootstrapped samples for the null model. The bootstrapping results revealed a standard error of .33 around the estimate of the effect of regression to the mean, which translates to a confidence interval of 6.44–7.71.

Regression to the mean with multiple years of baseline data. To illustrate an advantage to having multiple years of baseline data to identify the lowest performing schools, the study team calculated the effect of regression to the mean with multiple years of baseline data. This illustrates that an increase in the number of years of baseline data decreases the estimate of the effect of regression to the mean. Assuming that two years of school reading performance were used to identify the lowest performing schools in 2013/14, applying the same means, variance, and cutpoint from equation C1 to equation C2, the estimate of the effect of regression to the mean decreases from 7.05 to 3.68. The estimate decreases even more as the number of baseline measures increases (see figure 11 in the main report).

Multilevel random effects growth curve models with time-varying covariates. To describe growth in school reading performance, the study team built and tested a series of three-level (time, school, and district) random effects models (Raudenbush & Bryk, 2002). District was included in the model to control for the nested nature of schools in districts. Additionally, while schools implemented the extended school day policy, districts had considerable input into how schools implemented the extended hour. Therefore, it was necessary to control for the district level in the analyses. Time-varying covariates were used to account for differential growth in school reading performance among the lowest performing schools in the year in which the extended school year was implemented (McCoach & Kaniskan, 2010).

In the model-building process, the first focus was on establishing the level 1 model that accurately reflected the shape of the growth trajectory. If the level 1 model is mis-specified, incorrect parameter estimates and errors of inferences can occur in the upper levels of the model (McCoach & Kaniskan, 2010). Time was treated as a time-variant predictor and was centered at 2011/12. Growth represented yearly school reading performance change. Because only three years of data were included in the analysis, only linear growth could be tested. An additional time-variant predictor was added to indicate when a school implemented the extended school day. In this step, both persistent and fleeting coding schemes were tested to determine which coding scheme best fit the data (McCoach & Kaniskan, 2010; table C1). Model results suggested that a persistent effect best fit the data (table C2).

Table C1. Level 1 time-varying coding schemes

Year	Linear growth	Persistent effect	Fleeting effect
Other elementary schools			
2011/12	0	0	0
2012/13	1	0	0
2013/14	2	0	0
Lowest performing schools in 2012/13 only			
2011/12	0	0	0
2012/13	1	1	1
2013/14	2	1	0
Lowest performing schools in 2013/14 only			
2011/12	0	0	0
2012/13	1	0	0
2013/14	2	1	1
Lowest performing schools in both years			
2011/12	0	0	0
2012/13	1	1	1
2013/14	2	2	1

Source: Adapted from McCoach and Kaniskan (2010).

Table C2. Model fit indices

Model	Degrees of freedom	Akaike information criteria	Bayesian information criteria	Log likelihood ratio test	Deviance	Chi square test[a]	Degrees of freedom
Null model	4	42,397	42,423	–21,194	42,389	na	na
Model 1	5	42,377	42,410	–21,184	42,367	21.87***	1
Model 2	9	42,324	42,383	–21,153	42,306	61.29***	4
Model 3a	10	42,270	42,335	–21,125	42,250	55.79***	1
Model 3b	10	42,278	42,344	–21,129	42,258	47.75***[b]	1
Model 4a	13	41,632	41,718	–20,803	41,606	651.48***[c]	3
Model 4b	13	41,653	41,738	–20,814	41,627	630.69***[d]	3

*** Significant at $p < .001$.

na is not applicable.

a. Evaluates the difference between model fit of the previous model to the current model except for models 3b, 4a, and 4b. For example, the chi-square test for model 2 evaluates the difference from model 1.

b. Evaluates the difference between model 2 and model 3b, not model 3a, because they are not nested models.

c. Evaluates the difference between model 3a and model 4a.

d. Evaluates the difference between model 3b and model 4b.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

Once the level 1 model (time) was decided, level 2 (school) predictors were added. The only level 2 predictors that could be tested were the intercept adjustments for the lowest performing schools in 2012/13 only, the lowest performing schools in 2013/14 only, and the lowest performing schools in both years. The coefficients for these variables are the adjustment to the mean school reading performance in 2011/12 for each of the lowest performing schools cohorts prior to implementing the extended school day. Because the level 1 model included the time-variant predictor of growth for the lowest performing schools, it did not make sense to include level 2 predictors of growth (McCoach & Kaniskan, 2010). In summary, seven models were built and tested.

- Null model—random intercept; this model was used to identify the variance components.
- Model 1—null model plus fixed growth.
- Model 2—model 1 with random growth.
- Model 3a—model 2 plus persistent effect of the time-varying covariate for the lowest performing schools.
- Model 3b—model 2 plus fleeting effect of the time-varying covariate for the lowest performing schools.
- Model 4a—model 3a plus level 2 intercept adjustments for the lowest performing schools.
- Model 4b—model 3b plus level 2 intercept adjustments for the lowest performing schools.

All models were analyzed using R's Package lme4 (Bates, Maechler, Bolker, & Walker, 2014). In the model-building process, model fit was examined according to Akaike information criteria, Bayesian information criteria, log-likelihood, deviance, and chi-square tests. Both the growth and intercept estimates were tested as fixed and random effects. Because there are only three years of data available, only two parameters could be random effects. The final model used for interpretation (table C3), based on model fit indices (see table C2), was model 4a,[15] represented by:

Level 1 (time)

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}(year_{tij}) + \pi_{2ij}(LPS_{tij}) + e_{tij}$$

Level 2 (school)

$$\pi_{01j} = \beta_{00j} = \beta_{01j}(LPS_{2012/13_{ij}}) + \beta_{02j}(LPS_{2013/14_{ij}}) + \beta_{03j}(LPS_{both_{ij}}) + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + r_{1ij} \qquad \text{(C3)}$$

Level 3 (district)

$$\beta_{00j} = \gamma_{000} + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + u_{10j}$$

where Y_{tij} is the actual school reading performance at time t for school i in district j; $year_{ij}$ is the measurement occasion at time t for school i in district j; π_{2ij} is the difference when the school extends the day; and β_{01j}, β_{02j}, and β_{03j} are the time-invariant dummy code adjustments for schools identified as a lowest performing school in 2012/13, 2013/14, or both 2012/13 and 2013/14.

Table C3. Model results

Fixed effects					Random effects			
Coefficient	Estimate	Standard error	Degrees of freedom	*t*-value	Groups	Coefficient	Variance	Standard deviation
Null model								
(Intercept)	126.62	1.24	49.06	101.90***	School	(Intercept)	362.43	19.04
					District	(Intercept)	55.97	7.48
					Residual		80.87	8.99
Model 1								
(Intercept)	127.33	1.25	51	101.69***	School	(Intercept)	362.59	19.04
Year	–0.71	0.15	3,473	–4.68***	District	(Intercept)	55.97	7.48
					Residual		80.38	8.97
Model 2								
(Intercept)	127.45	1.21	44.51	105.11***	School	(Intercept)	343.92	18.55
Year	–0.77	0.27	39.28	–2.82***		Year	0.27	0.52
					District	(Intercept)	51.49	7.18
						Year	1.77	1.33
					Residual		78.14	8.84
Model 3a								
(Intercept)	127.44	1.22	44.2	104.81***	School	(Intercept)	346.95	18.63
Year	–1.07	0.27	41	–3.93***		Year	1.22	1.11
Persistent effect	6.73	0.77	3,156.5	8.74***	District	(Intercept)	51.78	7.20
						Year	1.75	1.32
					Residual		74.57	8.64
Model 3b								
(Intercept)	127.41	1.22	44	104.62***	School	(Intercept)	347.68	18.65
Year	–0.91	0.27	39	–3.31***		Year	0.71	0.85
Fleeting effect	5.56	0.75	3,500	7.42***	District	(Intercept)	51.93	7.21
						Year	1.79	1.34
					Residual		75.58	8.69
Model 4a								
(Intercept)	130.48	1.00	44	131.07***	School	(Intercept)	228.53	15.12
Year	–1.20	0.27	42	–4.42***		Year	1.19	1.09
Persistent effect	8.92	0.78	3,343	11.49***	District	(Intercept)	33.49	5.79
Lowest performing schools in 2012/13 only	–40.91	2.15	1,798	–19.05***		Year	1.74	1.32
Lowest performing schools in 2013/14 only	–34.45	2.13	1,708	–16.15***	Residual		74.46	8.63
Lowest performing schools in both years	–50.20	3.43	1,730	–14.64***				
Model 4b								
(Intercept)	130.35	1.0	44	130.4***	School	(Intercept)	227.58	15.09
Year	–1.00	0.27	40	–3.68***		Year	1.27	1.13
Fleeting effect	7.95	0.76	3,446	10.51***	District	(Intercept)	33.88	5.82
Lowest performing schools in 2012/13 only	–39.80	2.14	1,778	–18.64***		Year	1.75	1.32
Lowest performing schools in 2013/14 only	–34.29	2.13	1,707	–16.09***	Residual		74.85	8.65
Lowest performing schools in both years	–49.61	3.43	1,726	–14.48***				

*** Significant at $p < .001$.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

The primary parameter of interest in the final model was the time-varying covariate for implementation. In table C3 this parameter is the "persistent effect" in model 4a and is parameter π_{2ij} in equation C3. This parameter identified how growth in school reading performance changed when a school was required to implement the extended school day policy. Importantly, this parameter was compared with the estimate of the effect of regression to the mean. If this parameter were higher than that estimate, further exploratory analyses would be tenable.[16]

The null model suggested that 73 percent of total variance was between schools, 11 percent was between districts, and 16 percent was within schools over time. The final model had a pseudo R^2 of .32. Of the remaining variance, approximately 67 percent was at the school level, 11 percent was at the district level, and 22 percent was within schools.

The final model (model 4a; see table C3) suggested that in 2011/12, other elementary schools had an average school reading performance (β_{00j}) of 130.48 points. The school reading performance of the lowest performing schools in 2012/13 only (β_{01j}) was approximately 40.91 points less than that of other elementary schools, or 89.57 points (130.48–40.91). The school reading performance of the lowest performing schools in 2013/14 only (β_{02j}) was approximately 34.45 points less than that of other elementary schools, or 96.03 points (130.48–34.45). The school reading performance of the lowest performing schools in both years (β_{03j}) was approximately 50.20 points less than that of other elementary schools, or 80.28 points (130.48–50.20).

The final model (model 4a; see table C3) suggested that the school reading performance for other elementary schools decreased about 1.2 points a year (π_{1ij}). The persistent time-varying covariate (π_{2ij}), used to estimate the deflection, or time-specific change in the trajectory, in growth of school reading performance for the lowest performing schools the year the extra hour was implemented, was 8.92. Taken together, the school reading performance of the lowest performing schools increased approximately 7.72 points (–1.2 + 8.92) the year the school was required to extend the day.

Because growth in school reading performance among the lowest performing schools is calculated by adding two coefficients from the final model, confidence intervals cannot be computed directly from the final model. Rather, the model must be bootstrapped, and the growth in school reading performance among the lowest performing schools must be calculated at each bootstrap. The study team used 1,000 bootstrapped sampled results of the final model and computed the school reading performance of the lowest performing schools for each bootstrap. The bootstrapping results revealed a standard error of 0.79 for growth in school reading performance among the lowest performing schools, which translates to an average increase in school reading performance of 6.17–9.31 points for the lowest performing schools the year the policy was implemented.

The 95 percent confidence interval of growth in school reading performance among the lowest performing schools (6.17–9.31) overlapped with the 95 percent confidence interval for the estimate of the effect of regression to the mean (6.44–7.71). Therefore, on average, change within the lowest performing schools was not beyond what would have been expected and could not be statistically quantified in association with implementation of the extended hour.

Replication for each year and cohort of lowest performing schools. This section provides the results of the analysis replicated for each lowest performing schools cohort (schools identified as the lowest performing schools in 2012/13 and schools identified as the lowest performing schools in 2013/14).

To confirm that the findings were consistent across all the lowest performing schools cohorts, analyses were replicated for the schools identified as the lowest performing schools in 2012/13 compared with other elementary schools and for schools identified as the lowest performing schools in 2013/14 compared with other elementary schools. These analyses provide a clean disaggregation of the growth deviation during the year of implementation for only the cohort of interest. The same procedures for estimating the effect of regression to the mean and the multilevel model as described above were used with two subsets of data. The first included schools identified as the lowest performing schools in 2012/13 and other elementary schools. The second included schools identified as the lowest performing schools in 2013/14 and other elementary schools. Schools identified as the lowest performing schools in both years were included in both subsets because the extended school day policy was implemented in both years.

- The estimate of the effect of regression to the mean for schools identified as the lowest performing schools in 2012/13 (which excludes the lowest performing schools in 2013/14 only) was 6.49–7.90.
- The estimate of the effect of regression to the mean for schools identified as the lowest performing schools in 2013/14 (which excludes the lowest performing schools in 2012/13 only) was 6.48–7.80.

Subset analyses (table C4) results were similar to the full data results (see table C3). The intercepts of the lowest performing schools were significantly lower than those of other elementary schools. School reading performance declined each year for other elementary schools and for the lowest performing schools while the extended school day policy was not implemented. The year that the extended school day policy was implemented, school reading performance increased. However, the range of the increase in school reading performance for the lowest performing schools overlapped with the range of the effect of regression to the mean:

- The average increase in school reading performance among schools identified as the lowest performing schools in 2012/13 was 5.67–9.27, which overlapped with the range of the estimate of the effect of regression to the mean of 6.49–7.90.
- The average increase in school reading performance among schools identified as the lowest performing schools in 2013/14 was 1.90–5.48, which overlapped with the range of the estimate of the effect of regression to the mean of 6.48–7.80.

Because findings from the subset analyses do not substantially differ from the overall results from the full model (see table C3) results, the overall results are used for interpretation.

Table C4. Subset analyses final model results

Fixed effects					Random effects			
Coefficient	Estimate	Standard error	Degrees of freedom	*t*-value	Groups	Coefficient	Variance	Standard deviation
Null model								
School identified as the lowest performing schools in 2012/13								
(Intercept)	130.30	1.03	44	126.61***	School	(Intercept)	238.87	15.46
Year	–1.06	0.27	41	–3.87***		Year	1.07	1.03
Persistent effect	8.52	0.91	3,201	9.40***	District	(Intercept)	36.13	6.01
Lowest performing schools in 2012/13 only	–40.80	2.19	1,739	–18.64***		Year	1.81	1.34
Lowest performing schools in both years	–50.00	3.49	1,663	–14.33***	Residual	71.34	8.45	
Schools identified as the lowest performing schools in 2013/14								
(Intercept)	130.35	1.03	44	127.00***	School	(Intercept)	240.72	15.52
Year	–1.23	0.27	46	–4.50***		Year	0.92	0.96
Persistent effect	4.91	0.91	3,218	5.40***	District	(Intercept)	35.64	5.97
Lowest performing schools in 2013/14 only	–34.12	2.17	1,643	–15.73***		Year	1.75	1.32
Lowest performing schools in both years	–48.13	3.50	1,667	–13.76***	Residual		71.18	8.44

*** Significant at $p < .001$.

Source: Authors' analysis of data from School Accountability Reports (Florida Department of Education, 2012a, 2013a, 2014a).

Appendix D. Supplemental statistical tests of significance for school characteristics and implementation of extended school day policy

This appendix provides results of supplemental statistical tests of significance. Tests are listed in the order in which they appear in the main text of the report.

Region

A chi-square test revealed that the lowest performing elementary schools were not evenly distributed across regions ($X^2 = 33.82$, $df = 12$, $p = .001$). Some regions had significantly more lowest performing schools than other regions.

Enrollment

A repeated measures analysis of variance of enrollment revealed a nonsignificant test for time ($F = 2.59$, $df = 2$, $p = .075$), a significant test for cohort ($F = 36.01$, $df = 3$, $p < .001$), and a nonsignificant time by cohort interaction ($F = 1.19$, $df = 6$, $p = .307$). This suggests that enrollment was stable over time, that the lowest performing schools were smaller than other elementary schools, and that enrollment did not change differently for the lowest performing schools compared with other elementary schools. However, the within-year correlation between enrollment and school reading performance ranged from .28 to .31 (see table A2 in appendix A), which suggests a moderate correlation.

Title I status, racial/ethnic minority, and eligibility for the school lunch program

A chi-square test revealed that there were statistically significantly more lowest performing schools than other elementary schools that were Title 1 schools ($X^2 = 92.83$, $df = 3$, $p < .001$). Additionally, analysis of variance revealed that there were statistically significantly more racial/ethnic minority students ($F = 225.29$, $df = 3$, $p < .001$) and students eligible for the school lunch program ($F = 225.31$, $df = 3$, $p < .001$) in the lowest performing schools than other elementary schools.

Minutes of reading instruction

A repeated measures analysis of variance of minutes of reading instruction revealed a significant test for time ($F = 110.64$, $df = 1$, $p < .001$), a significant test for cohort ($F = 42.29$, $df = 1$, $p < .001$), and a significant time by cohort interaction ($F = 68.43$, $df = 1$, $p < .001$). This suggests that the amount of reading instruction in 2012/13 was significantly different from that in 2013/14, that the lowest performing schools in 2013/14 only and the lowest performing schools in both years provided significantly different amounts of instruction, and that the change in the amount of instruction provided was different for the lowest performing schools in 2013/14 only and the lowest performing schools in both years.

Chi-square tests revealed no significant differences in how the lowest performing schools in 2013/14 only and the lowest performing schools in both years added the additional hour (question 11 on the survey; $X^2 = 3.34$, $df = 5$, $p < .648$).

Staffing

Chi-square tests revealed no significant differences in how the lowest performing schools in 2013/14 only and the lowest performing schools in both years were staffing the extra hour (question 9a on the survey; $X^2 = 2.61$, $df = 2$, $p < .272$).

Chi-square tests revealed no significant differences in how the lowest performing schools in 2013/14 only and the lowest performing schools in both years made staffing changes (question 9b on the survey) from 2012/13 to 2013/14 related to reading coaches ($X^2 = 4.5$, $df = 2$, $p < .105$), paraprofessionals ($X^2 = 1.75$, $df = 2$, $p < .418$), or volunteers ($X^2 = 2.17$, $df = 1$, $p < .141$). However, there were significant differences related to teachers ($X^2 = 7.93$, $df = 2$, $p < .019$) and other staff ($X^2 = 7.05$, $df = 2$, $p < .029$). The lowest performing schools in both years lost proportionally more teachers and gained proportionally fewer teachers than the lowest performing schools in 2013/14 only. The lowest performing schools in both years lost proportionally more other staff than the lowest performing schools in 2013/14 only, which gained more staff.

Grouping

Chi-square tests revealed no significant differences if students were grouped differently during the extra hour for the lowest performing schools in 2013/14 only and the lowest performing schools in both years (question 13 on the survey; $X^2 = 2.14$, $df = 2$, $p < .342$). Chi-square tests also revealed no significant differences in grouping strategies for the instruction provided in the extra hour (question 14 on the survey). Chi-square tests ranged from .00 to 2.36 with 1 degree of freedom; *p*-values were all greater than .125.

Compliance

Chi-square tests revealed no significant differences between the lowest performing schools in 2013/14 only and the lowest performing schools in both years regarding compliance with the extended school day policy (questions 1–8 on the survey). Chi-square tests ranged from .81 to 2.55 with 3 degrees of freedom; *p*-values were all greater than .532.

Notes

1. The requirement was introduced when House Bill 5101 amended sections 1011.62(4)(f)(2) and 1011.62(9)(a) of the Florida Statutes (http://laws.flrules.org/files/Ch_2012–133.pdf).
2. Information for 2012/13 was a retrospective report based on respondent recall and may not be exact.
3. Based on the fitted means from model 4a (see table C3 in appendix C). Observed means are provided in table A1 in appendix A.
4. Results were the same when each implementation cohort (schools that were identified as the lowest performing schools in 2012/13 and schools that were identified as the lowest performing schools in 2013/14) were analyzed separately. Disaggregated cohort results can be found in table C4 in appendix C.
5. The school reading performance value of 93 was selected for this example because it was used in the study to calculate regression to the mean. The value was selected for the study because it was the highest value across the years examined of the lowest performing schools. For more details, see appendix C.
6. See table C4 in appendix C for additional subanalyses for each cohort, the results of which are not substantively different across cohorts.
7. In 2010/11 Florida switched from the FCAT to the FCAT 2.0. The percentage of students making learning gains from the prior year on the reading component of the FCAT 2.0—one of the two measures used to calculate school reading performance—was calculated differently in 2010/11 because whether students made gains was based on their performance on different tests, in contrast to the calculations in 2011/12–2013/14, for which whether students made gains was based only on their performance on the FCAT 2.0. The percentage of students scoring at or above achievement level 3—the other measure used to calculate school reading performance—was measured consistently between 2010/11 and 2013/14. While school reading performance could be mathematically computed for 2010/11, the values would not be comparable to the values for 2011/12–2013/14.
8. For a technical discussion of the influence of skewness and kurtosis, see Yuan, Bentler, and Zhang (2014).
9. The expected change was nearly identical when calculated for the schools identified as the lowest performing schools in 2013/14 only, providing the most accurate estimation of expected variation.
10. School grades are assigned by the Florida Department of Education and communicate to the public how well a school is performing relative to state standards based on student achievement on the Florida Comprehensive Assessment Test 2.0 (and are thus assigned only to schools that include grades 3 or higher; Florida Department of Education, 2014b). Schools are exempt from grading the first year they are open, and such schools are also excluded from this report. See http://schoolgrades.fldoe.org for more on Florida's school grades.
11. The list of the 100 lowest performing schools was not published on the Florida Department of Education website, though the list was available to the public on request. Since the 2014/15 school year, the list of the 300 lowest performing schools has been available on the Florida Department of Education website.
12. Ideally, the population mean, variances, and correlations would be obtained prior to policy implementation. Because data prior to implementation were not available, the

correlation across years may be biased downward by impacts of the policy. This could result in an overestimation of the magnitude of the effect of regression to the mean.

13. A quantile–quantile plot compares the quantiles of a data distribution with the quantiles of a standardized normal distribution. A probability-probability plot compares the empirical cumulative distribution function of a data set with the cumulative distribution of a normal distribution.
14. A higher cutpoint produces a lower estimate of the effect of regression to the mean, and a lower cutpoint produces a higher estimate. A lower estimate of the effect of regression to the mean allows for less of an adjustment to the final growth estimate. Using a cutpoint of 90 (the lowest cutpoint ever used between 2011/12 and 2013/14) resulted in an estimate of the effect of regression to the mean of 7.47. The study team was concerned with overestimating the effect of regression to the mean, which could lead to incorrectly attributing growth to regression to the mean instead of meaningful growth among the lowest performing schools.
15. Though models 4a and 4b are not directly comparable by chi-square tests, according to Akaike information criteria, Bayesian information criteria, and log-likelihood, model 4a fit the data better. The results of model 4b are presented and are not substantively different from the results of model 4a.
16. Because of concerns about regression to the mean, the study team first analyzed growth on the school reading performance to determine whether there was any growth outside the effect of regression to the mean. If evidence of growth beyond the effect of regression to the mean had been found, it would have been of interest to examine growth on portions of the school reading performance—the percentage of students reading at level 3 (satisfactory) or above or the percentage of students making learning gains.

References

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34*(1), 215–220.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1–7. Retrieved February 1, 2015, from http://CRAN.R-project.org/package=lme4.

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., et al. (2009). Individualizing student instruction precisely: Effects of child × instruction interactions on first graders' literacy development. *Child Development, 80:* 77–100.

Davis, C. E. (1976). The effect of regression to the mean in epidemiologic and clinical studies. *American Journal of Epidemiology, 104*(5), 493–498.

Florida Department of Education (2008). *FCAT achievement levels*. Tallahassee, FL: Author. Retrieved August 1, 2014, from http://fcat.fldoe.org/pdf/fcAchievementLevels.pdf.

Florida Department of Education. (2011). *Value added model recommendation to the commissioner of education from the student growth implementation committee*. Tallahassee, FL: Author. Retrieved January 29, 2016, from http://www.fldoe.org/core/fileparse.php/7503/urlt/0072159-summaryfinalrecommendation.pdf.

Florida Department of Education. (2012a). *2011–2012 school accountability reports: Detailed information on non-high schools*. Tallahassee, FL: Author. Retrieved August 1, 2014, from http://schoolgrades.fldoe.org/xls/1112/All-districts-11–12.xls.

Florida Department of Education. (2012b). *Master school ID 2011–2012*. Tallahassee, FL: Author. Retrieved August 1, 2012, from http://doeweb-prd.doe.state.fl.us/EDS/MasterSchoolID/.

Florida Department of Education. (2012c). *School membership by grade, preliminary survey 2, 2011–12*. Tallahassee, FL: Author. Retrieved August 1, 2014, from http://www.fldoe.org/core/fileparse.php/7584/urlt/0086502-mem_schl_grd1112.xls.

Florida Department of Education. (2013a). *2012–2013 school accountability reports: Detailed information on non-high schools*. Tallahassee, FL: Author. Retrieved August 1, 2014, from http://schoolgrades.fldoe.org/xls/1213/All-districts-12–13.xls.

Florida Department of Education. (2013b). *Master school ID 2012–2013*. Tallahassee, FL: Author. Retrieved August 1, 2013, from http://doeweb-prd.doe.state.fl.us/EDS/MasterSchoolID/.

Florida Department of Education. (2013c). *School membership by grade, preliminary survey 2, 2012–13*. Tallahassee, FL: Author. Retrieved August 1, 2014, from http://www.fldoe.org/core/fileparse.php/7584/urlt/0086501-mem_schl_grd1213.xls.

Florida Department of Education. (2014a). *2013–2014 school accountability reports: Detailed information on non-high schools.* Tallahassee, FL: Author. Retrieved August 1, 2014, from http://schoolgrades.fldoe.org/xls/1314/All-Districts-13–14.xls.

Florida Department of Education. (2014b). *2014 guide to calculating school grades.* Tallahassee, FL: Author. Retrieved August 1, 2014, from http://schoolgrades.fldoe.org/pdf/1314/SchoolGradesCalcGuide2014.pdf.

Florida Department of Education. (2014c). *Master school ID 2013–2014.* Tallahassee, FL: Author. Retrieved August 1, 2014, from http://doeweb-prd.doe.state.fl.us/EDS/MasterSchoolID/.

Florida Department of Education. (2014d). *School membership by grade, preliminary survey 2, 2013–14.* Tallahassee, FL: Author. Retrieved August 1, 2014, from http://www.fldoe.org/core/fileparse.php/7584/urlt/0069369-mem_schl_grd1314.xls.

Florida Department of Education. (2015). *Office of inspector general differentiated accountability report #A-1314–016.* Tallahassee, FL: Author. Retrieved January 29, 2016, from http://www.fldoe.org/core/fileparse.php/7514/urlt/Final-Report-A-1314–016-Differentiated-Accountability.pdf.

Florida Department of Education. (n.d.). *Improving the academic achievement of the disadvantaged.* Retrieved January 29, 2016, from http://www.fldoe.org/policy/federal-edu-programs/title-i-part-a-improving-the-academic-/improving-the-academic-achievement-of-.stml.

Gardner, M. J., & Heady, J. A. (1973). Some effects of within-person variability in epidemiological studies. *Journal of Chronic Diseases, 26*(12), 781–795.

George, D., & Mallery, P. (2010). *SPSS for Windows: Step by Step.* Boston, MA: Pearson Education, Inc.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology, 96*(3), 403–423.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 75*(5), 949–967.

IBM Corp. (2012). *IBM SPSS Statistics for Windows, Version 21.0.* Armonk, NY: Author.

Johnson, W. D., & George, V. T. (1991). Effects of regression to the mean in the presence of within-subject variability. *Statistics in Medicine, 10*(1), 1295–1302.

Kim, H.-Y. (2013). Statistical notes for clinical researchers assessing normal distribution using skewness and kurtosis. *Restorative Dentistry and Endodontics, 38*(1), 52–54.

Linden, A. (2013). Assessing regression to the mean effects in health care initiatives. *BMC Medical Research Methodology, 13*(1), 1–7.

McCoach, D. B., & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology, 1*(17), 1–12.

Moats, L., Foorman, B., & Taylor, P. (2006). How quality of writing instruction impacts high-risk fourth graders' writing. *Reading and Writing, 19*(4), 363–391.

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

Sacerdote, B. (2008). *When the saints come marching in: Effects of hurricanes Katrina and Rita on student evacuees* (NBER No. w14385). Washington, DC: National Bureau of Economic Research.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (6th ed)*. Boston, MA: Allyn and Bacon.

Yuan, K-H., Bentler, P. M., Zhang, W. (2014). The effect of skewness and kurtosis on mean and covariance structure analysis: The univariate case and its multivariate implications. *Sociological Methods and Research, 34*(2), 240–258.

The Regional Educational Laboratory Program produces 7 types of reports

Making Connections
Studies of correlational relationships

Making an Impact
Studies of cause and effect

What's Happening
Descriptions of policies, programs, implementation status, or data trends

What's Known
Summaries of previous research

Stated Briefly
Summaries of research findings for specific audiences

Applied Research Methods
Research methods for educational settings

Tools
Help for planning, gathering, analyzing, or reporting data or research