Molly Schlesinger,
Senior Research Associate

Giselle Saleet,
Data Analyst

Kristyn Stewart,
Director of Strategic Projects

Katherine Mosher,
Senior Project Manager

Office of Research and Evaluation

October 2021

THE SCHOOL DISTRICT OF PHILADELPHIA

Four-year analysis of 2015-16 Kindergarteners' aimswebPlus reading and PSSA performance from 2015-16 to 2018-19

Key Findings

- Of the over 12,000 kindergarten students enrolled in SDP schools in 2015-16, only a quarter (4,194) took the aimswebPlus core reading assessment in fall and spring each year between 2015-16 and 2018-19 and also took the PSSA English Language Arts (ELA) exam in Spring 2019.
- On average, students ended third grade by performing in the same aimswebPlus performance group as they did in kindergarten.
- Although English Learners (ELs) had lower average aimswebPlus performance than their non-EL counterparts across the four years, ELs had the same average performance starting kindergarten as at the end of third grade, while non-ELs declined in their average aimswebPlus performance across the four years.
- Overall, students in the analytic sample who scored in the Advanced category on the third grade ELA PSSA had higher aimswebPlus performance in third grade than in kindergarten, while students who scored in the Below Basic and Basic categories on the ELA PSSA had lower aimswebPlus performance in third grade than in kindergarten.

Contents

Why and how we track literacy performance of Kindergarten through third-grade students

Reading proficiently by the end of third grade is critical to a student's future academic success.

Research has found that students who are not proficient readers by the end of third grade (when instruction transitions from "learning to read" to "reading to learn") are four times more likely to drop out of high school. The lowest performers in third grade are at the greatest risk of dropping out: 23% of below-basic readers drop out or fail to finish high school on time, compared to 9% of third graders with basic reading skills and 4% of third graders who are proficient readers.[1]

One of the Board of Education's Goals for the District is that 62% of third grade students will score proficient on the State ELA assessment (PSSA) by August 2026 (Goal 2).

As of spring 2019, 32.5% of third-grade students earned a score of proficient or advanced on the State administered English Language Arts (ELA) assessment (ELA PSSA).[2] To track student progress more frequently than once per year, the School District has identified within-year "leading indicators." In order to reach the 62% goal, there must be an increase in the percentage of students in grades K-3 – both overall (Leading Indicator 2.1) and by student group (Leading Indicator 2.2E) – who score at or above grade-level on aimswebPlus each year.[3]

Research Questions

In this report, we explore three primary research questions to better understand the performance of the 2015-16 kindergarten student cohort on aimswebPlus between 2015-16 and 2018-19 and the third grade ELA PSSA in 2018-19:

1. How did student performance on the aimswebPlus assessment change between kindergarten (2015-16) and third grade (2018-19) for the 2015-16 kindergarten cohort?
2. How did student performance on the aimswebPlus assessment change between kindergarten and third grade by demographic characteristics for the kindergarten cohort?
3. How did student performance on the aimswebPlus assessment change between kindergarten and third grade grouped by third grade ELA PSSA performance for the 2015-16 kindergarten cohort?

---

[1] Annie E. Casey Foundation (2013). *Early Warning Confirmed*. http://www.aecf.org/m/resourcedoc/AECF-EarlyWarningConfirmed2013.pdf; Annie E. Casey Foundation (2012). *Double Jeopardy*. https://www.aecf.org/resources/double-jeopardy/
[2] Read more about the PSSA-ELA here: https://www.education.pa.gov/K12/Assessment%20and%20Accountability/PSSA/Pages/default.aspx
[3] For more about SDP's Goals and Guardrails see: https://www.philasd.org/schoolboard/goals-and-guardrails/ and https://www.philasd.org/era/goals-and-guardrails/

Assessment Data Used for this Analysis

AimswebPlus

SDP used aimswebPlus, a universal early literacy screening, benchmarking, and progress-monitoring tool from Pearson, to assess literacy proficiency in SDP for all K-3 students from the 2014-15 to the 2020-21 school year.

The District administered multiple aimswebPlus subtests to students in each grade level three times a year. Students received a score based on the number of cues they correctly identify in a 60-second period. Subtest scores are combined into a "composite score" or general outcome measures (GOM) for assessments prior to 2018-19 which provided teachers with a holistic measure of student literacy performance.

At each grade level, one subtest was designated as the "core" subtest – that is, one required subtest that was taken each fall, winter, and spring. "Core" subtests allow teachers, school leaders, and District staff to track student performance and growth over time. However, because each grade level has its own "core" subtest that measures a discrete literacy skill, comparisons across grade levels should be interpreted with caution (i.e., longitudinal comparisons are not a representation of students' growth on the same set of skills).[4]

The "core" assessments for each grade level are as follows:

- The kindergarten **Letter Naming Fluency (LNF) assessment**
- The first-grade **Nonsense Word Fluency (NWF) assessment**
- The second- and third-grade **Oral Reading Fluency (ORF) assessment**

Key Data Points

For each of these three assessments, we looked closely at the following key data points related to performance on the assessments.

1. **National Percentile Rank (NPR)**: A student's national percentile rank is a norm-referenced performance measure that compares a student's score with a national reference of scores of students in the same grade who were tested on the same content during the same timeframe. NPR is useful for understanding a student's reading skills compared to other students in the same grade nationally. Based on the number of correct responses, each student is assigned a percentile rank in the fall and in the spring. Percentiles range from 0-99. For example, a fall NPR of the 23rd percentile indicates that the student is performing better than 23% of the nationally-normed sample based on their number of correct responses that fall.
2. **AimswebPlus Performance Groups**: Based on their NPR, students fall into one of six performance groupings (Table 1). These performance groups give us a more nuanced

---

[4] In the 2018-19 school year SDP switched from aimsweb to aimswebPlus. aimswebPlus introduced slightly updated assessments, renormed national percentiles, and composite scores. This report uses data from aimsweb and aimswebPlus. For consistency, aimsweb will be referred to as aimswebPlus in this report.

understanding of student performance, while still allowing us to categorize students based on their performance. While there are other ways of grouping students (by Tier level, for example), aimswebPlus performance groups are useful for analyzing how students at various levels of proficiency are performing and improving over time.[5]

Table 1. aimswebPlus core assessment NPR group names and NPR range

Grouping Name	NPR Range
Well Above Average	90-99
Above Average	75-89
High Average	51-75
Low Average	26-50
Below Average	11-25
Well Below Average	1-10

PSSA

The Pennsylvania System of School Assessment (PSSA) is a standards-based, criterion-referenced test administered to all Pennsylvania students in grades 3-8 during spring of each year. The intended purpose of the PSSA is to measure how well students acquire the knowledge and skills described in the Pennsylvania Assessment Anchor Content Standards as defined by the Eligible Content for Mathematics, English and Language Arts (ELA), and Science.[6]

Students receive a scale score for each assessment based on the number and difficulty of the questions they answer correctly. The PSSA is based on an item response theory model, which means that its scale scores adjust for the item difficulties. Using cut-points on the score range, which can vary across grade and academic year, scale scores corresponds to one of four performance tiers: Below Basic, Basic, Proficient, and Advanced.[7]

In this report, student performance on the third grade ELA PSSA exam is analyzed based on the following performance levels: Below Basic, Basic, Proficient, and Advanced.

---

[5] aimswebPlus provides five groups: Well Above Average, Above Average, Average, Below Average, and Well Below Average. We split the Average (25-75%) group into High Average (51-75%) and Low Average (25-50%) because that encapsulates 50% of the national sample, in previous research we found that the bottom and top halves of the Average group perform differently from each other.

[6] For more information see https://www.education.pa.gov/K-12/Assessment%20and%20Accountability/PSSA/Pages/PSSA-Technical-Reports.aspx

[7] For links to more information about cut scores and a description of performance levels at each grade level see https://www.education.pa.gov/K-12/Assessment%20and%20Accountability/PSSA/Pages/DescriptorsCutScores.aspx

Identifying the Student Sample

About one quarter of kindergarten students took all of the assessments during spring and fall of each year between 2015-16 and 2018-19.

Student Enrollment and Participation

In 2015-16, there were 12,399 kindergarten students enrolled in SDP schools (Figure 1).[8] The cohort was smaller in 2016-17 and even smaller in 2017-18. By the 2018-19 school year, there were 8,147 third-grade students enrolled in SDP schools who had been enrolled since 2015-16 as kindergarteners.

Figure 1. The number of students who were kindergarteners in 2015-16 and were enrolled through third grade in 2018-19

Kindergarten (2015-16) 12,399	First Grade (2016-17) 10,621	Second Grade (2017-18) 9,095	Third Grade (2018-19) 8,147

Source: Data from RL-Total Student Enrollment Yearly [v1.0.0], data pulled on January 15, 2021

Of the 8,147 kindergarten students enrolled in 2015-16 who remained enrolled in SDP through 2018-19 as third-grade students, 6,257 students took both the spring and fall core aimswebPlus assessment (Letter Naming Fluency) in 2015-16 (Figure 2). Of those 6,257 students, 4,194 were still enrolled as of 2018-19 AND took both the fall and spring ORF core aimswebPlus assessment (Oral Reading Fluency) in 2018-19 as well as the ELA PSSA in spring 2018-19.

Figure 2. The number of students who were continuously enrolled and took the aimswebPlus core assessments and ELA PSSA

Kindergarten (2015-16) 6,257	First Grade (2016-17) 4,576	Second Grade (2017-18) 4,343	Third Grade (2018-19) 4,194

Source: Data from RL-Total Student Enrollment Yearly [v1.0.0], data pulled on January 15, 2021
Note: The number of third grade students is the analytic sample size.

Student Characteristics

The characteristics of students who were enrolled in SDP schools as kindergarten students in 2015-16 and of those students who were continuously enrolled in SDP schools between 2015-16 and 2018-19 do not differ widely from the characteristics of the analytic sample of students who were enrolled continuously with assessment data each year (Table 2). For example, the percentage of

---

[8] Enrollment is defined as 10 days or more in a District school.

Black/African American students in the analytic sample (43%) is only 6 percentage points lower when compared to overall kindergarten enrollment in 2015-16 (49%).

Table 2. The demographic characteristics of kindergarten students included in the sample compared to the students enrolled in SDP schools as kindergarten students in 2015-16

Student Characteristic	Kindergarten Enrollment 2015-16 (n = 12,399)	Third Grade Enrollment 2018-19 (n = 8,147)	Analytic Sample of students with aimswebPlus and PSSA data (n = 4,194)
	Percent of students	Percent of students	Percent of students
Gender			
Female	46.7%	46.4%	48.8%
Male	53.3%	53.6%	51.2%
Race/Ethnicity			
Asian/Pacific Islander	6.9%	7.6%	8.5%
Black/African American	49.6%	47.0%	43.4%
Hispanic/Latinx	19.4%	22.5%	23.9%
Multi-racial/Other	10.1%	9.2%	9.8%
White	13.9%	13.8%	14.4%
Socio-Economic Status			
Economically Disadv.	69.1%	76.4%	75.9%
Not Economically Disadv.	30.9%	23.6%	24.1%
English Learner Status			
English Learner	10.6%	11.8%	12.5%
Not an English Learner	89.4%	88.2%	87.5%
Special Education Status			
Students with IEPs	8.4%	16.7%	7.8%
Students without IEPs	91.6%	83.3%	92.2%

Source: Data from Qlik RL-Total Student Enrollment Yearly [v1.0.0], data pulled on January 15, 2021

AimswebPlus Performance Groups

The aimswebPlus performance groups were calculated using students' performance on the fall kindergarten aimswebPlus assessment. Of the 4,194 students in the analytic sample, nearly 60% fell into the Low Average, Below Average, or Well Below Average groups, which corresponds to an NPR of 50 percentiles or lower, or that 60% of kindergarten students in the District scored lower on their first assessment than half of kindergarten students in the national sample (Table 3). Less than ten percent of performed in the Well Above Average group, or an NPR of 91 percentiles or higher.

Table 3. aimswebPlus performance group data in fall of kindergarten (2015-16)

aimswebPlus performance group	Well Below Average	Below Average	Low Average	High Average	Above Average	Well Above Average
NPR Percentile Range	0-10	11-25	26-50	51-75	76-90	91-99
Number of students	557	936	1026	815	514	346
Percentage of students	13%	22%	24%	19%	12%	8%

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents the numbers and percentage of students in the analytic sample in each aimswebPlus performance group when they were in kindergarten in fall 2015-16. The National Percentile Rank (NPR) ranges represent the range of NPR within each performance group. For example, a student with an NPR of the 78th percentile falls into Above Average group because their NPR is between the 76th and 90th percentiles.

ELA PSSA Performance Levels

Of the 4,194 students in the analytic sample, 26% scored in the Below Basic category on the 2018-19 ELA PSSA in third grade (Table 4). Thirty-eight percent (38%) of students scored in the Basic category, 30% scored in the Proficient category, and 6% scored in the Advanced category.

Table 4. ELA PSSA performance level data in third grade (2018-19)

ELA PSSA Performance Levels	Below Basic	Basic	Proficient	Advanced
Number of students who scored at each level	1075	1580	1275	264
Percent of students who scored at each level	26%	38%	30%	6%

Source: Data from PSSA State Accountability File, data accessed on January 15, 2021

Findings

How did student performance on the aimswebPlus assessment change between kindergarten (2015-16) and third grade (2018-19) for the 2015-16 kindergarten cohort?

Students in the lowest performing aimswebPlus performance groups experienced an increase in their average NPR over time.

The average NPR trajectories of the six aimswebPlus performance groups between fall of kindergarten (2015-16) and spring of third grade (2018-19) are presented in Figure 3 (see Table 1 for more information about the six performance groups). Student groups were formed based on aimswebPlus performance in fall of kindergarten. The average NPR for students in the six groups is presented for eight timepoints between fall of kindergarten and spring of third grade.

Students who performed in the Well Above Average and Above Average groups in kindergarten had higher average NPR in kindergarten than in third grade. The students in High Average group in kindergarten had the same average NPR in kindergarten as third grade. Students in the Low Average, Below Average, and Well Below Average groups in kindergarten had lower average NPRs in kindergarten than in third grade, which indicates that these students improved their performance relative to the national sample over time.

Figure 3. Average NPR on the core aimswebplus assessment over 4 years disaggregated by fall kindergarten (2015-16) aimswebPlus performance group
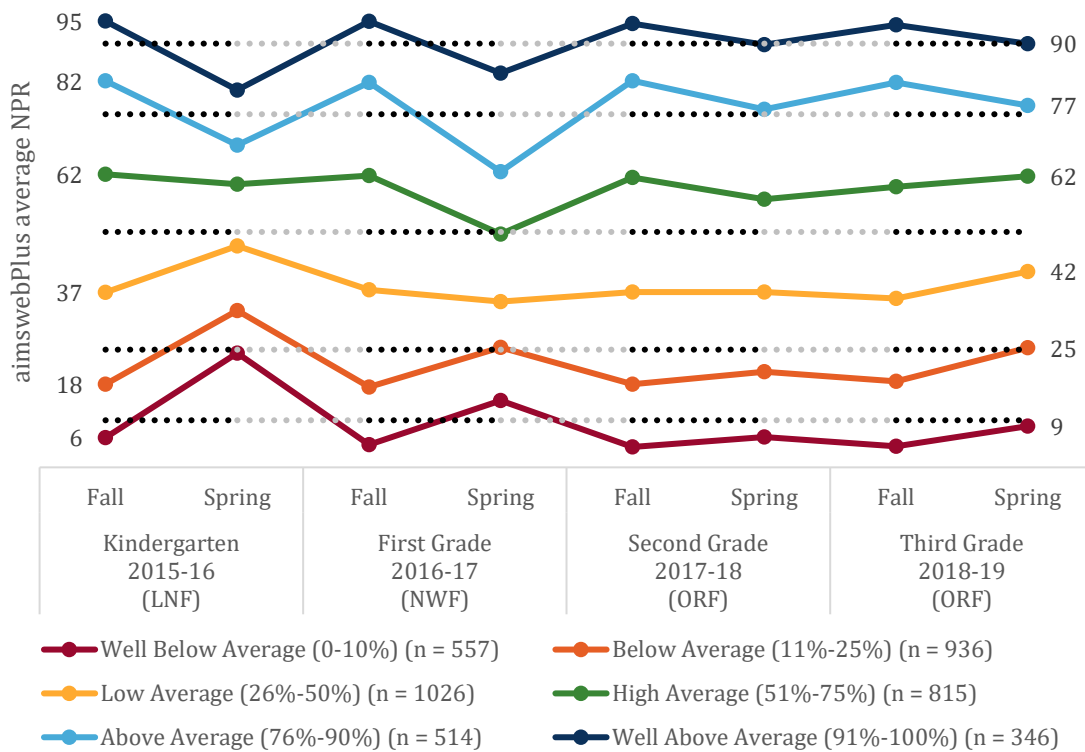
Source: data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR for six aimswebPlus performance groups that were based on performance on the fall of kindergarten year. Average NPRs for these groups are presented for fall and spring windows of the kindergarten through third grade years on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points for the aimswebPlus performance group ranges; for example, any marker above the 25th percentile cut point performed at 26th percentile or higher.[9]

[9] The black dotted lines represent fall to spring timepoints, when students stay in the same grade and typically the same classroom throughout the year, and the grey dotted lines represent spring to fall timepoints when students are promoted from one grade to the next.

Comparing average NPR between groups over time

The figures in this report feature aimswebPlus average NPR results from three different core assessments, or a different test in kindergarten, first grade, and second grade/third grade (see page 5 for information on the different assessments). Each assessment measures slightly different literacy skills, and therefore we would anticipate differences in average NPR from testing window to testing window within groups. It is important to focus on the comparison of different groups on the overall trends across the four years to better understand the role of being in different performance or demographic groups has on average NPR over time.

Students generally stayed in the same aimswebPlus performance group over time.

Average NPR scores of the six kindergarten aimswebPlus performance groups at eight timepoints between fall of kindergarten (2015-16) and spring of third grade (2018-19) are shown in Table 5. The average NPR scores correspond to the markers in Figure 3. For example, students who scored in the Low Average aimswebPlus performance group in fall of kindergarten (the yellow line that starts with the 37th percentile in fall 2015 on the graph) had an average 37th percentile NPR in the fall of their kindergarten year and an average 42nd percentile NPR in the spring of third grade.

Table 5. Average NPR on the core aimswebPlus assessment over 4 years, disaggregated by fall kindergarten (2015-16) aimswebPlus performance group

Kindergarten aimswebPlus performance group	Number of Students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Well Above Average	557	95	80	95	84	94	90	94	90
Above Average	936	82	68	82	63	82	76	82	77
High Average	1026	62	60	62	50	62	57	60	62
Low Average	815	37	47	38	35	37	37	36	42
Below Average	514	18	33	17	25	18	20	18	25
Well Below Average	346	6	24	5	14	4	6	4	9

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores of the six kindergarten aimswebPlus performance groups between fall of kindergarten (2015-16) and spring of third grade (2018-19). The average NPR scores correspond to the markers in Figure 3.

Although average NPR varied across seasons and years, the average NPR of students in the same kindergarten aimswebPlus performance group did not vary much from the fall of kindergarten to the spring of third grade (Table 5). Students in lower aimswebPlus performance groups (Well Below Average, Below Average, and Low Average) in fall of kindergarten experienced increases in their average NPR, and students in higher aimswebPlus performance groups (Above Average, Well

Above Average) in fall of kindergarten experienced decreases in their average NPR by spring of third grade (for the NPR ranges of each performance group, see Table 3).

How did student performance on the aimswebPlus assessment change between kindergarten and third grade by demographic characteristics for the kindergarten cohort?

We also examined student performance over time by demographic characteristics, such as gender, race/ethnicity, English Learner (EL) status, Special Education (IEP) status, and Economically Disadvantaged status.

Both male and female students decreased in their average NPR over time.

Both male and female students had lower NPR in third grade than in kindergarten (Figure 4 and Appendix A Table A1). Both groups declined in their performance relative to the national sample between the four years. Female students (the blue line that starts with the 45th percentile in fall 2015 at the top of the graph) dipped from an average NPR of 45th percentile in the fall of kindergarten to an average NPR of 39th percentile in spring of third grade. Male students followed a similar pattern, decreasing nine percentiles in their NPR from 43rd percentile in the fall of kindergarten to 34th percentile in the spring of third grade.

Figure 4. Average NPR on core assessments over 4 years, disaggregated by gender
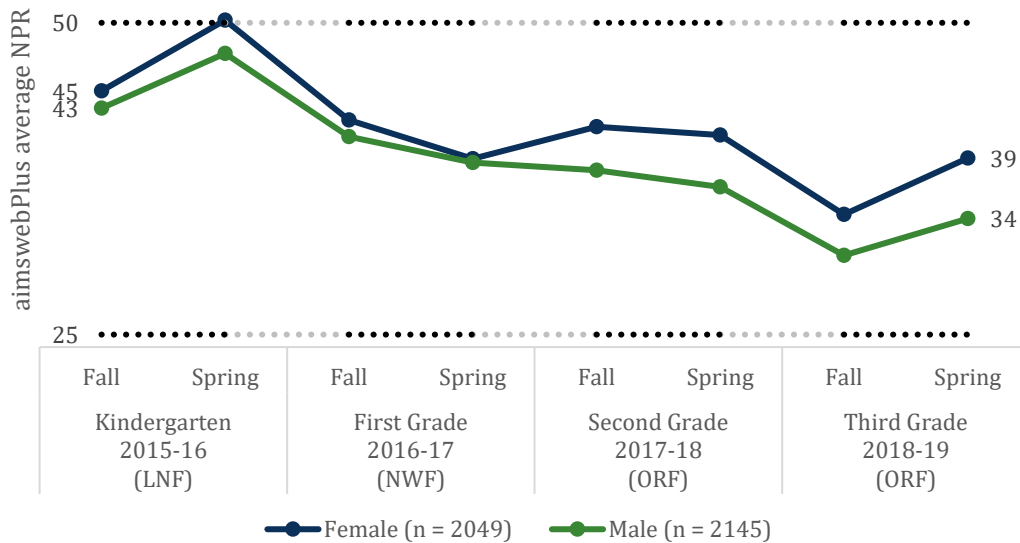
Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR for female and male students in fall and spring of the kindergarten through third grade years in the markers on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points for the aimswebPlus performance group ranges (Table 2). For example, female students (the dark blue line at the top) had an average NPR of 45 in the fall of kindergarten, and an average NPR of 39 by the spring of third grade.

Black/African American students' average NPR decreased 14 percentile pointss from kindergarten to third grade, while their peers of other racial groups did not decrease as much.

Black/African American and Multi-Racial/Other students had lower average NPR in spring of third grade (2018-19) than in fall of kindergarten (2015-16). Asian/Pacific Islander, White, and Hispanic/Latino students' average NPR were largely unchanged, wavering an average of only one percentile from kindergarten to third grade (Figure 5 and Appendix A Table A2). Black/African American and Multi-Racial/Other students experienced the largest NPR changes over time (a decrease of 14 percentile points and 6 percentile points, respectively). Black/African American students, White students, and Asian students' average NPR were within two percentiles of each other in fall of kindergarten, while Black/African American students experience a sharp decrease in average NPR, White students' average NPR declined one percentile, and Asian students' average NPR increased one percentile from fall of kindergarten to spring of third grade.

Figure 5. Average NPR on core assessment over 4 years, disaggregated by race/ethnicity student group
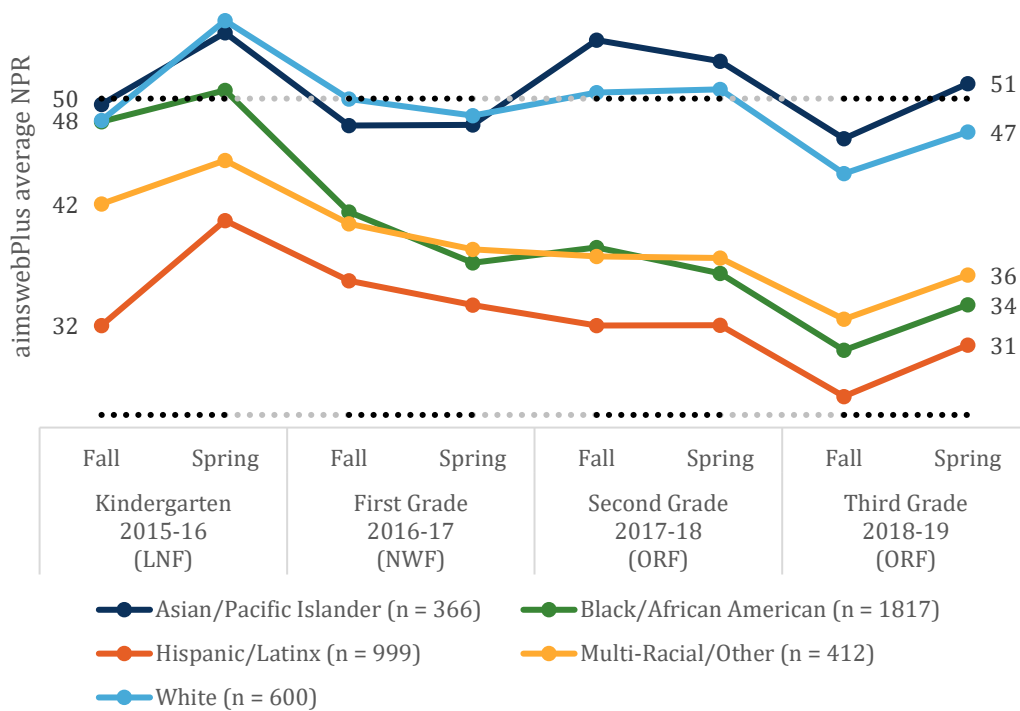

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR of the race/ethnicity student group in fall and spring of the kindergarten through third grade years in the markers on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points for the aimswebPlus performance group ranges (Table 4). For example, Asian/Pacific Islander students (the dark blue line at the top) had an average NPR of 50 in the fall of kindergarten, and an average NPR of 51 by the spring of third grade.

Although English Learners (ELs) had lower average NPR than non-ELs, ELs had the same average NPR in fall of kindergarten and spring of third grade, while non-ELs experienced a decline.

We also examined NPR changes over four years by English Learner (EL) status (see Appendix C for analyses by home language).[10]

Non-ELs had higher average NPR in kindergarten (45th percentile) than in third grade (37th percentile) indicating that their performance declined compared to the national sample, and ELs had the same average NPR (33rd percentile) in fall kindergarten and spring of third grade (Figure 6 and Appendix A Table A3).

Figure 6. Average NPR on core assessment over 4 years, disaggregated by English Learner (EL) status

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

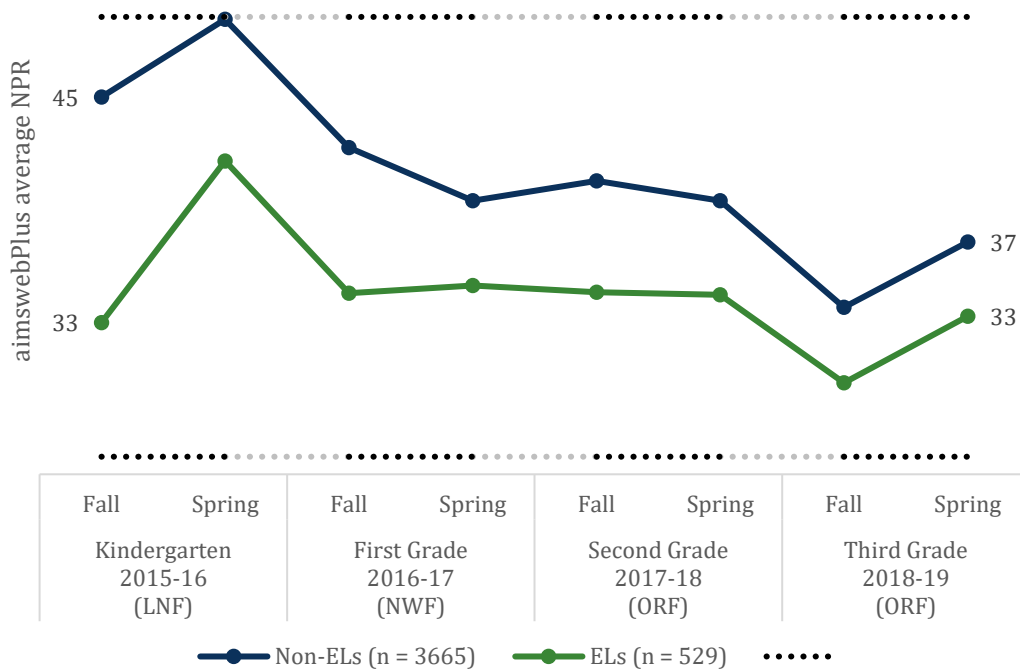**How to read this figure:** This figure presents the average NPR for ELs and non-ELs in fall and spring of the kindergarten through third grade years in the markers on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points for the aimswebPlus performance group ranges (Table 4). For example, ELs (the green line) had an average NPR of 33 in the fall of kindergarten and in spring of third grade.

[10] English Learner (EL) status was determined by whether a student had an LEP indicator in kindergarten (2015-16). Most often, students who may qualify as ELs are tested for EL status when they first enter the school district, therefore, we selected EL status in kindergarten. EL students can test out and may no longer be ELs by third grade – this only impacted 9 students out of the sample.

14

Students with IEPs experienced a larger decrease in average NPR over time than students without IEPs from kindergarten to third grade.

In addition to EL status, we examined NPR performance across the four years by Special Education status, or whether or not students had an IEP.[11] Student with an IEP at any point between kindergarten and third grade are included in the students with IEP group. We also examined average NPR trajectories by primary disability (see Appendix C for primary disability analyses).

Both students with and without IEPs had lower average NPR in third grade than in kindergarten, demonstrating lower performance relative to the national sample over time (Figure 7 and Appendix A Figure A4). For example, students with IEPs (the green line that starts with the 30th percentile in the fall of 2015) declined from an average NPR of 30th percentile in the fall of kindergarten to an average NPR of 18th percentile in spring of third grade. Students without IEPs, on average, experienced a decrease in their NPR as well, from 46th percentile in the fall of kindergarten to 39th percentile in the spring of third grade – a decrease of 7 percentiles.

Figure 7. Average NPR on core assessment over 4 years, disaggregated by Special Education status
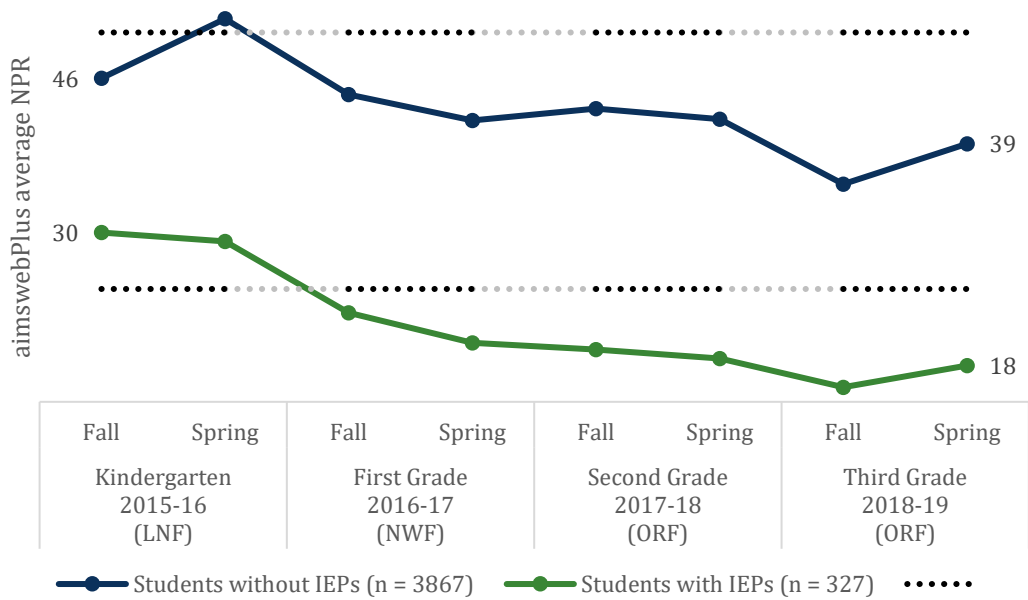
Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents data by IEP status for the average NPR in fall and spring of the kindergarten through third grade years in the markers on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points for the aimswebPlus performance group ranges (Table 2). For example, students with IEPs (the green line) had an average NPR of 30 percentiles in the fall of kindergarten, and an average NPR of 18 percentiles by the spring of third grade.

---

[11] Students can be recommended for evaluation to receive an IEP at any point; receiving an IEP later can retroactively explain low performance prior to having the IEP. Not all students with an IEP are eligible to take aimswebPlus and PSSA, which may explain a smaller sample of students with IEPs who took the aimswebPlus assessment in the four years of interest. Students who are gifted, but do not have a disability are not considered "Students with IEPs." Students who are gifted and have a disability are included in the "Students with IEPs" group.

15

Economically disadvantaged students and non-economically disadvantaged students experienced similar declines in average NPR from kindergarten to third grade.

We also examined student average NPR trajectory by Economical Disadvantage status.[12]

Both economically disadvantaged and non-economically disadvantaged students had lower average NPR in third grade than in kindergarten (Figure 8 and Appendix Table A5). For example, economically disadvantaged students (the green line that starts with the 43rd percentile in the fall of 2015) dipped from an average NPR of 43rd percentile in the fall of kindergarten to an average NPR of the 35th percentile in spring of third grade. Non-economically disadvantaged students experienced a 5-percentile decrease in their average NPR from the 45th percentile in the fall of their kindergarten year to the 40th percentile in the spring of third grade.

Figure 8. Average NPR on core assessment over 4 years, disaggregated by Economically Disadvantaged status

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR for economically disadvantaged students and non-economically disadvantaged students in kindergarten through third grade in the markers on each of the solid lines. For example, economically disadvantaged students (the green line) had an average NPR of 43 in the fall of kindergarten, and an average NPR of 35 by the spring of third grade.

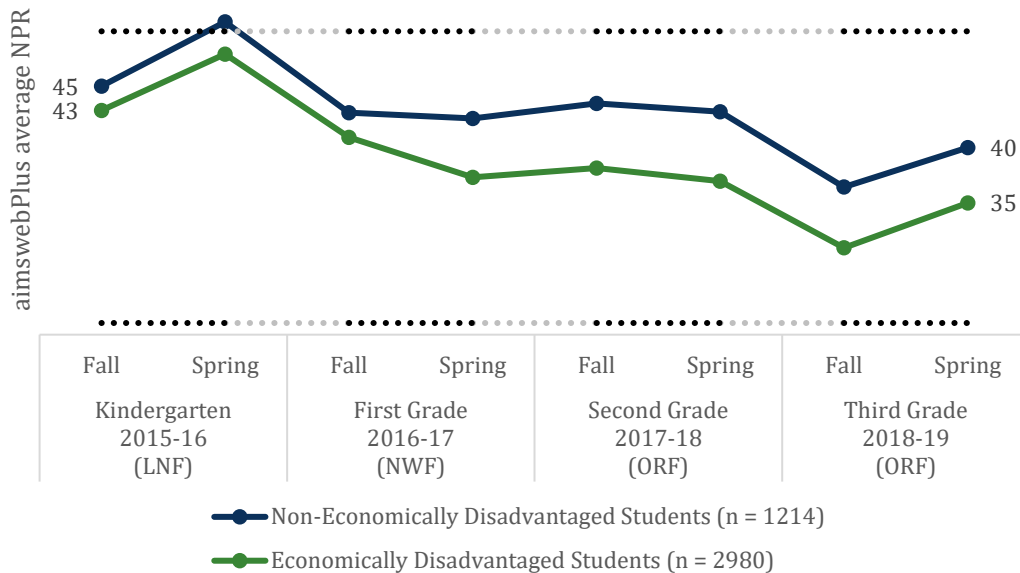[12] Economic Disadvantage status was determined by whether students' families were determined to be eligible for income tested federal assistance programs by the state during their kindergarten year.

How did student performance on the aimswebPlus assessment change between kindergarten and third grade grouped by third grade ELA PSSA performance for the kindergarten cohort?

In addition to examining students' NPR over time by aimswebPlus performance groups and demographic characteristics, we also examined student aimswebPlus performance from kindergarten to third grade grouped by students' third grade ELA PSSA performance. That is, we grouped students by their third-grade performance, and looked backwards at their performance in second grade, first grade, and kindergarten to observe patterns over time. This type of analyses allows us to observe if there were patterns in aimswebPlus performance during the four years that could have potentially predicted third grade ELA PSSA performance.

Students who performed in the Basic or Below Basic categories on the ELA PSSA had lower average NPR in third grade than in kindergarten.

The ELA PSSA performance level groups were formed based on students' third grade ELA PSSA performance, and allows us to look backwards at their average NPR from spring of third grade back to fall of kindergarten.

Students who performed in the Basic or Below Basic categories on the ELA PSSA had lower average NPR on the aimswebPlus core assessment in spring third grade – the same season that students take the ELA PSSA – than they had in fall of kindergarten. There was a general negative trend for students who performed in the Basic or Below Basic categories on the ELA PSSA, with their average NPR decreasing across the four year to lower average NPR percentiles in spring of third grade (Figure 9 and Appendix Table A6). The opposite was true of students who performed in the Advanced or Proficient categories on the third grade ELA PSSA: they had higher average NPR in spring of third grade than in fall of kindergarten. Overall, students who performed in the Advanced or Proficient categories on the third grade ELA PSSA experienced positive trajectories, with lower percentiles in fall of kindergarten slightly increasing to higher average NPR percentiles in spring of third grade.

These changes indicate that students who performed in the Advanced or Proficient categories on the third grade ELA PSSA improved their aimswebPlus performance relative to the national sample over time, whereas students who performed in the Basic or Below Basic categories on the third grade ELA PSSA declined in their aimswebPlus performance relative to the national sample.

Figure 9. Average NPR on core assessments over 4 years, grouped by spring third grade (2018-19) ELA PSSA performance
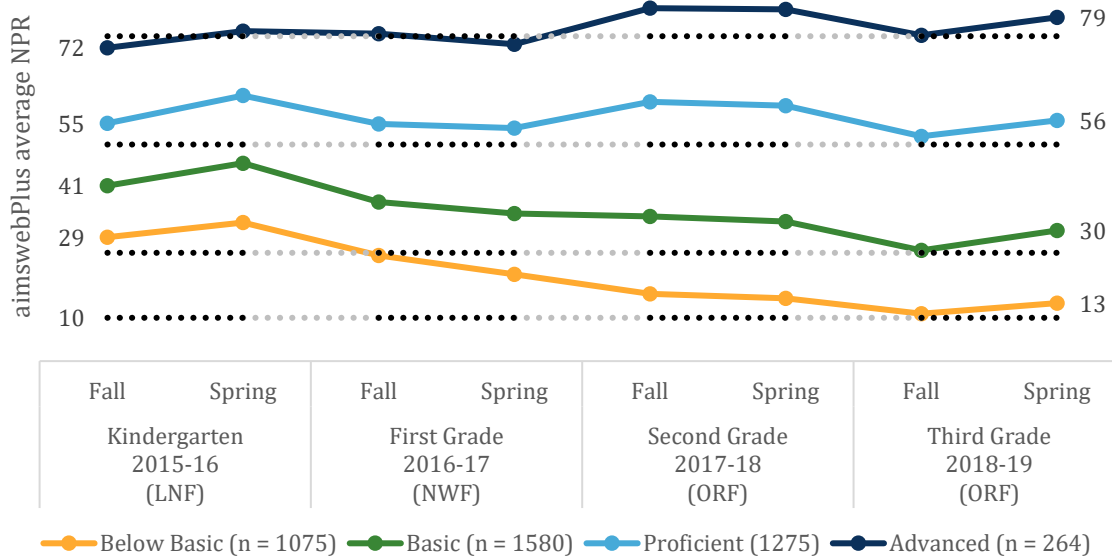
Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0] and PSSA State Accountability File, data accessed on January 15, 2021

How to read this figure: This figure presents the average NPR of the four third grade ELA PSSA performance levels in fall and spring of kindergarten through third grade on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points for the aimswebPlus performance group ranges (Table 4). For example, students who scored Advanced on the third grade ELA PSSA (the dark blue line at the top) had an average NPR of 79 by the spring of their third-grade year, and an average NPR of 72 in the fall of their kindergarten year, demonstrating an overall increase in average NPR from kindergarten to third grade.

Conclusions

This report examined the four-year longitudinal trends of a cohort of kindergarten students in 2015-16 through their third-grade year (2018-19) based on their aimswebPlus core reading assessment in spring and fall of the four years.

More than 12,000 kindergarten students were enrolled in SDP schools in 2015-16; however, only a quarter (4,194) of these students met the criteria for inclusion in the analytical sample: remaining enrolled in SDP schools in the third grade, taking the aimswebPlus core reading assessment in fall and spring between 2015-16 and 2018-19, and taking the ELA PSSA in 2018-19. The demographics of students who met these criteria did not differ much from year to year. All three groups (kindergarten students enrolled in 2015-16, students who were still enrolled in third grade in SDP schools in 2018-19, and students who took aimswebPlus between 2015-16 and 2018-19 and the ELA PSSA in 2018-19) were about 46-47% female, 43-50% Black/African American, 19-24% Hispanic/Latinx, 69-76% economically disadvantaged, 11-13% ELs, and 8-17% Special Education students. [13]

The average student generally fell in the same aimswebPlus performance group in spring of third grade (2018-19) as where they started in fall of kindergarten (2015-16).

Although average female students performed higher than the average male students, both groups, on average, experienced National Percentile Rank (NPR) decreases across the four years. Average Black/African American and Multi-Racial/Other students experienced the largest decreases in average NPR, while average Asian/Pacific Islander, White, and Hispanic/Latinx students experienced little change from fall of kindergarten to spring of third grade.

Although average non-ELs performed higher than average ELs, ELs had the same average NPR across the four years, whereas non-ELs experienced a decline over time. Students with IEPs experienced larger decreases in average NPR than students without IEPs from fall of kindergarten to spring of third grade, ranging from 10 to 20 NPR percentiles from kindergarten to third grade.

Although average non-economically disadvantaged students performed higher than economically disadvantaged students, both groups experienced similar declines in average NPR over time.

Students who performed in the Advanced and Proficient categories on the ELA PSSA experienced an increase in average NPR—having higher average NPR in spring of third grade when the ELA PSSA is administered—than in fall of kindergarten. Students who performed in the Basic and Below Basic categories experienced decreases in average NPR from kindergarten to third grade, and had lower average NPRs in spring of third grade than in fall of kindergarten.

---

[13] The larger variation for Special Education between the analytic sample and other samples is not surprising, as students with IEPs do not always take progress monitoring or state assessments based on their IEPs. The increase from 8.4% to 16.7% from 2015-16 to 2018-19 may be due to more students being identified by third grade as benefiting from an IEP than while enrolled in kindergarten. The decrease from 16.7% in third grade to the 7.8% in the analytic sample may represent the group of students who were not required to take the aimswebPlus core reading assessments and/or the ELA PSSA based on their IEP accommodations after they received their IEP, meaning they did not have the data required to be included in the analytic sample.

Appendix A

Table A1. Average NPR on core assessment over 4 years, disaggregated by gender

Gender	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Female	2049	45	50	42	39	42	41	35	39
Male	2145	43	48	41	39	38	37	31	34

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores by student between fall of kindergarten and spring of third grade by gender. The average NPR scores correspond to the markers in Figure 4. For example, female students in the analytic sample had an average NPR of 45 in the fall of kindergarten and an average NPR of 39 by the spring of third grade.

Table A2. Average NPR on core assessment over 4 years disaggregated by race/ethnicity student group

Race/ Ethnicity	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Asian/ Pacific Islander	366	50	55	48	48	55	53	47	51
Black/African American	1817	48	51	41	37	38	36	30	34
Hispanic/ Latinx	999	32	40	36	34	32	32	26	31
Multi-Racial/ Other	412	42	45	40	38	38	37	33	36
White	600	48	56	50	49	50	51	44	47

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores by race/ethnicity student group between fall of kindergarten and spring of third grade. The NPR scores correspond to the markers in Figure 5. For example, Asian/Pacific Islander students in the analytic sample had an average NPR of 50 in fall of kindergarten and 51 in spring of third grade.

Table A3. Average NPR on core assessment over 4 years, disaggregated by English Learner (EL) status

EL Status	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Non-ELs	3665	45	50	43	40	41	40	33	37
ELs	529	33	42	34	35	34	34	29	33

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores between fall of kindergarten (2015-16) and spring of third grade (2018-19) by EL status. The average NPR scores correspond to the markers in Figure 6. For example, ELs in the analytic sample had an average NPR of 33 in the fall of kindergarten and in spring of third grade.

Table A4. Average NPR on core assessment over 4 years, disaggregated by Special Education status

Special Education Status	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Students w/o IEPs	3867	46	51	44	41	43	42	35	39
Students with IEPs	327	30	30	23	20	19	18	15	18

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores between fall of kindergarten and spring of third grade by Special Education status. The NPR scores correspond to the markers in Figure 8. For example, students with IEPs in the analytic sample had an average NPR of 30 in the fall of kindergarten, and an average NPR of 18 by the spring of third grade.

Table A5. Average NPR on core assessment over 4 years, disaggregated by economic disadvantage status

Economically Disadvantaged Status	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Non-Economically Disadvantaged	1214	45	51	43	43	44	43	37	40
Economically Disadvantaged	2980	43	48	41	37	38	37	31	35

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR by student between fall of kindergarten (2015-16) and spring of third grade (2018-19) by Special Education status. The average NPR scores correspond to the markers in Figure 9. For example, economically disadvantaged students in the analytic sample had an average NPR of 43 in the fall of kindergarten, and an average NPR of 35 by the spring of third grade.

Table A6. Average NPR on core assessment over 4 years, disaggregated by spring third grade (2018-19) ELA PSSA performance

ELA PSSA in Spring of Third Grade	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Advanced	1075	72	76	76	73	81	81	75	79
Proficient	1580	55	61	55	54	60	59	52	56
Basic	1275	41	46	37	34	33	32	26	30
Below Basic	264	29	32	24	20	15	14	11	13

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR between fall of kindergarten (2015-16) and spring of third grade (2018-19) by ELA PSSA performance level. The average NPR scores correspond to the markers in Figure 11. For example, student who performed at Advanced on the third grade ELA PSSA in the analytic sample had an average NPR of 79 in spring of third grade and an average NPR of 72 in fall of kindergarten.

Appendix B

Average NPR trajectories for race/ethnicity student group by gender.

Table B1. Gender and Race/Ethnicity group data

Variable	Female		Male	
	# of students	% of students	# of students	% of students
Asian/Pacific Islander	167	8%	199	9%
Black/African American	895	44%	922	43%
Hispanic/ Latinx	509	25%	490	23%
Multi-Racial/Other	190	9%	222	10%
White	288	14%	312	15%

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

Figure B1. Average NPR on core assessment over 4 years for female students, disaggregated by race/ethnicity
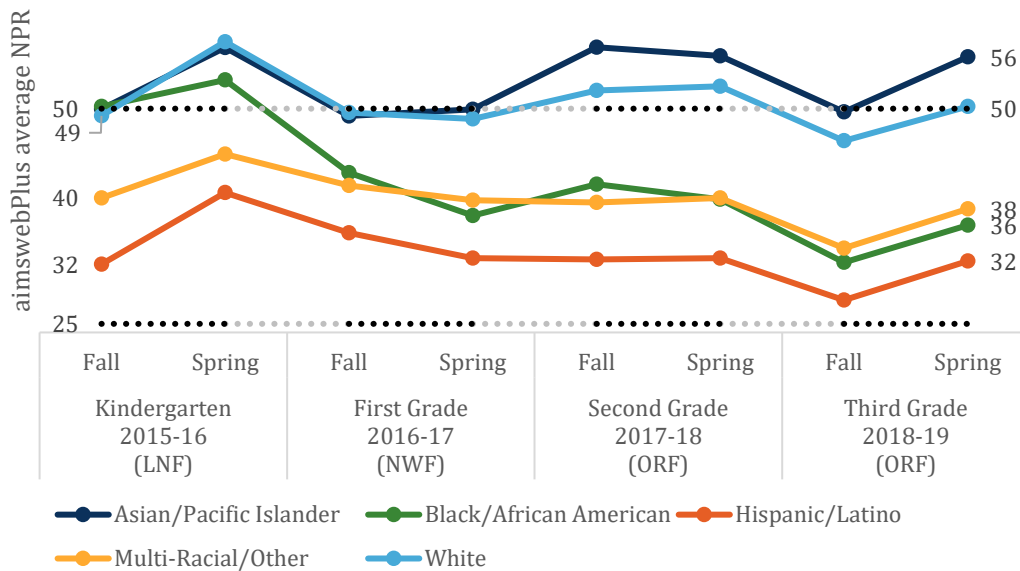
Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

Table B2. Average NPR on core assessment over 4 years for female students, disaggregated by race/ethnicity

Race/Ethnicity	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Asian/Pacific Islander	50	57	49	50	57	56	50	56
Black/African American	50	53	43	38	41	39	32	36
Hispanic/Latinx	32	40	36	33	32	33	28	32
Multi-Racial/Other	40	45	41	39	39	40	34	38
White	49	58	50	49	52	53	46	50

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

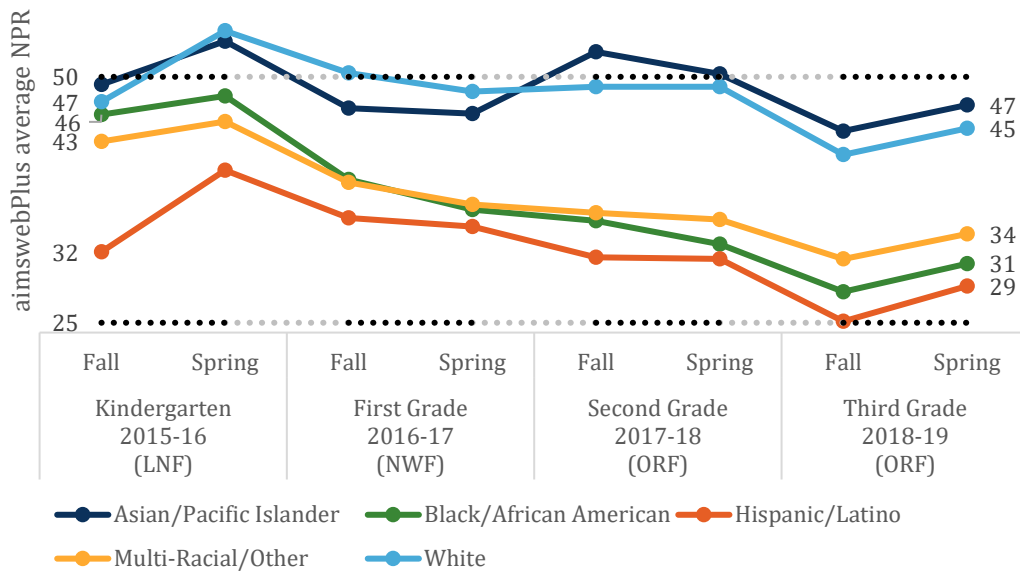Figure B2. Average NPR on core assessment over 4 years for male students disaggregated by race/ethnicity

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR of the five race/ethnicity student group for male students in fall and spring of the kindergarten through third grade years in the markers on each of the solid lines. The dotted lines represent the aimswebPlus performance group cut points.

Table B3. Average NPR on core assessment over 4 years for male students, disaggregated by race/ethnicity

Race/Ethnicity	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Asian/Pacific Islander	49	54	47	46	53	50	44	47
Black/African American	46	48	40	36	35	33	28	31
Hispanic/Latinx	32	41	36	35	32	32	25	29
Multi-Racial/Other	43	45	39	37	36	35	31	34
White	47	55	50	49	49	49	42	45

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

23

Appendix C

Of the 529 ELs in the analytic sample, 49% spoke Spanish, 14% spoke Chinese, 6% spoke Arabic, 5% spoke Vietnamese, 4% spoke Khmer, and 22% spoke all other languages (Table C1).

Table C1. Home Language data for ELs

Home Language	Arabic	Chinese (Mandarin)	Khmer	Spanish	Vietnamese	All Other Languages
# of ELs	33	75	21	258	27	115
% of ELs	6%	14%	4%	49%	5%	22%

Source: Data from RL- Total Student Yearly Enrollment [v1.0.0], data pulled on January 15, 2021

Vietnamese, Arabic, and Khmer speakers had higher average NPR in third grade than in kindergarten (Figure C1 and Table C2). Spanish speakers had the same average NPR percentile at both timepoints. Chinese speakers and speakers of all other languages had lower average NPR percentiles in third grade than in kindergarten. In comparison, Non-ELs had lower average NPR in third grade than in kindergarten, which was lower than the average NPR for Vietnamese, Chinese and Khmer speakers, and speakers of all other languages.

Figure C1. Average NPR on core assessment over 4 years, disaggregated by Home Language for ELs
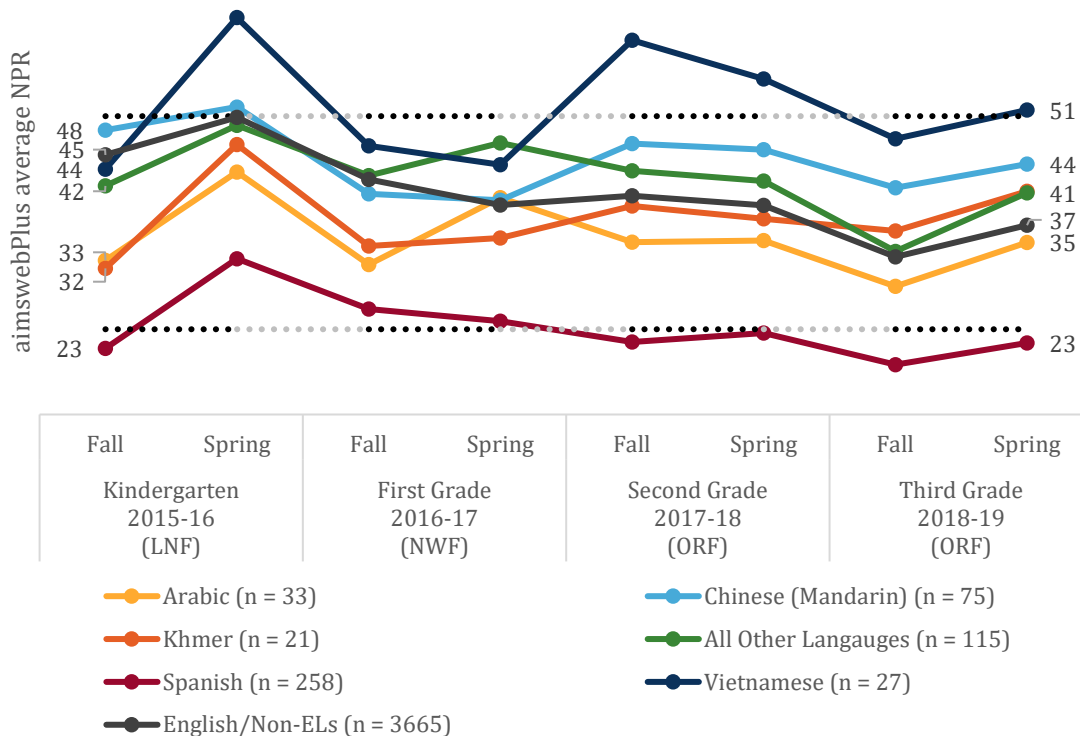
Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR of the home language EL groups in fall and spring of the kindergarten through third grade years in the markers on the solid lines.

24

Table C2. Average NPR on core assessment over 4 years, disaggregated by Home Language for ELs

EL Home Languages	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Arabic	33	33	43	33	40	35	35	30	35
Chinese (Mandarin)	75	48	51	41	40	47	46	42	44
Khmer	21	32	47	35	36	39	38	37	41
Spanish	115	23	33	27	26	24	25	21	23
Vietnamese	258	44	62	47	44	59	54	47	51
All Other Languages	27	42	49	43	47	44	42	34	41

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores between fall of kindergarten and spring of third grade by EL home language. The average NPR scores correspond to the markers in Figure 7. For example, Vietnamese speakers in the analytic sample had an average NPR of 44 in fall of kindergarten and 51 by spring of third grade.

Students in each primary disability group experienced a decline in average NPR of 10 to 20 points from kindergarten to third grade.

In addition to looking NPR across the four years by IEP status, we also looked at performance changes by the primary disability of students with IEPs.

Average NPR change by primary disability

Of the 481 students with IEPs in the sample, 42% had a Specific Learning Disability, 31% had Speech or Language Impairment, 14% had Other Health Impairment, and 5% had, and 7% of the sample is categorized into "All Other Primary Disabilities" (Table C3).

Table C3. Primary disability data for students with IEPs

Variable	Number of students	Percent of students
Autism	26	5%
Other Health Impairment	68	14%
Specific Learning Disability	204	42%
Speech or Language Impairment	151	31%
All Other Primary Disabilities	32	7%

Source: Data from RL- Total Student Yearly Enrollment [v1.0.0], data pulled on January 15, 2021

All groups experienced lower average NPRs in third grade than in kindergarten (Figure C2).

Figure C2. Average NPR on core assessment over 4 years by primary disability for students with IEPs
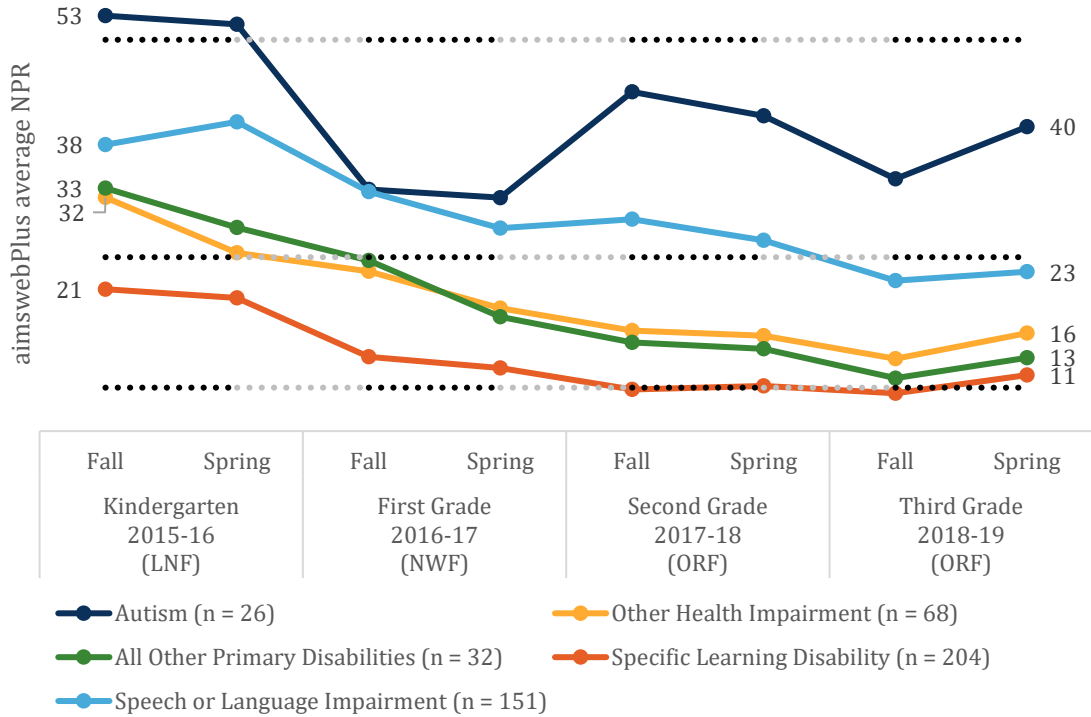
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
	Kindergarten 2015-16 (LNF)		First Grade 2016-17 (NWF)		Second Grade 2017-18 (ORF)		Third Grade 2018-19 (ORF)	

- Autism (n = 26)
- Other Health Impairment (n = 68)
- All Other Primary Disabilities (n = 32)
- Specific Learning Disability (n = 204)
- Speech or Language Impairment (n = 151)

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this figure: This figure presents the average NPR of the primary disability groups in fall and spring of kindergarten through third grade in the markers on the solid lines. The dotted lines represent the aimswebPlus performance cut points. For example, students with Autism (the dark blue line) had an average NPR of 53 in the fall of kindergarten, and an average NPR of 40 by the spring of third grade.

Table C4 presents average NPR scores by primary disability between fall of kindergarten and spring of third grade. The average NPR scores correspond to the markers in Figure B1. For example, students with Other Health Impairments (the yellow line) had an average NPR of 32 in the fall of Kindergarten and an average NPR of 16 in the spring of third grade.

Table C4. Average NPR on core assessment over 4 years, disaggregated by primary disability

Primary Disability	Number of students	Kindergarten (LNF) 2015-16		First Grade (NWF) 2016-17		Second Grade (ORF) 2017-18		Third Grade (ORF) 2018-19	
		Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Autism	26	53	52	33	32	44	41	34	40
Other Health Impairment	68	32	26	23	19	17	16	13	16
Specific Learning Disability	151	21	20	14	12	10	10	9	11
Speech or Lang.Impairment	204	38	41	33	28	29	27	22	23
All Other Primary Disabilities	32	33	28	25	18	15	14	11	13

Source: Data from Qlik L3_AIMSWEBSTAR [v1.0.0], data pulled on January 15, 2021

How to read this table: This table presents average NPR scores between fall of kindergarten and spring of third grade by primary disability. The average NPR scores correspond to the markers in Figure 8.

From kindergarten to third grade, Black/African American students who were economically disadvantaged experienced the largest decline in average NPR compared to all other economically disadvantaged race/ethnicity groups.

In addition to looking Economic Disadvantage status, we also examined average NPR by the race/ethnicity of economically disadvantaged students. Of the 2,980 economically disadvantaged students in the analytic sample, 46% were Black/African American, 24% were Hispanic/Latinx, 12% were White, 10% were Multi-Racial/Other, and 8% were Asian/Pacific Islander (Table C5).

Table C5. Race/Ethnicity data for Economically Disadvantaged students

Race/Ethnicity	Asian/ Pacific Islander	Black/ African American	Hispanic/ Latinx	Multi- Racial/ Other	White
Number of Economically Disadvantaged students	236	1359	725	288	372
Percentage of Economically Disadvantaged students	8%	46%	24%	10%	12%

Source: Data from RL- Total Student Yearly Enrollment [v1.0.0], data pulled on January 15, 2021

Students in all groups stayed in Low-Average-NPR group throughout the four years (Figure C3 and C6). Asian/Pacific Islander, Hispanic/Latinx, and white students experiences the least average NPR change, while Black/African American and Multi-Racial/Other students experience larger declines from kindergarten to spring of third grade. For example, Multi-Racial/Other students (the yellow line that starts with the 40th percentile in the fall of 2015) had an average NPR of 40th percentile in the fall of kindergarten and an average NPR of 34th percentile in the spring of third-grade. Whereas, Asian/Pacific Islander students experienced a one-percentile increase in their NPR from the 49th percentile in fall of kindergarten to the 50th percentile in spring of third grade.

Appendix D

National Percentile Rank was used as the measure for student performance in this report. However, when the scores that the students get in an assessment are normally distributed, which is often the case, NPR is not an interval measure. That is, the range of scores that students receive in the lower and upper ends of the NPR spectrum (around the 1st and the 99th percentiles) will be much larger than the range of scores that students receive around 50th percentile.

NPR is an easily accessible measure and therefore we used average NPR for different student groups' performance in this report without any conversions. The usual practice is converting NPR into a "Normal Curve Equivalent (NCE),"[14] averaging over the NCE scores and then converting those averages back to the percentile equivalents rather than averaging over an ordinal scale such as NPR. There were no changes in performance when we converted NPR to NCE, averaged the NCE and converted the average NCE back to NPR.

---

[14] For a more detailed explanation see: "Making Sense of NCEs and Standard Errors" by Pennsylvania Department of Education available at: https://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PVAAS/Methodology/MakingSenseOfNCEsAndStandardErrors.pdf