

# Using kindergarten entry assessments to measure whether Philadelphia's students are on-track for reading proficiently

Jessica F. Harding, Mariesa Herrmann, Elías S. Hanno, Christine Ross

October 2019



# Using kindergarten entry assessments to measure whether Philadelphia’s students are on-track for reading proficiently

Jessica F. Harding, Mariesa Herrmann, Elías S. Hanno, Christine Ross

October 2019

The School District of Philadelphia (SDP) is a partner in the Free Library of Philadelphia’s *Read by 4th*, a citywide initiative that aims to substantially increase the percentage of students who can read on grade level when they enter grade 4. To measure the percentage of entering kindergarteners on-track to meet this goal, SDP partnered with the Regional Educational Laboratory Mid-Atlantic to identify an indicator based on the Pennsylvania Kindergarten Entry Inventory (KEI). Students with higher scores on the KEI in 2014/15 had higher English language arts scores on Pennsylvania’s grade 3 state assessment in 2017/18. The study established the threshold of 6 on a 1-10 index of the two KEI dimensions (Emerging Academic Competencies and Learning Engagement Competencies) to predict whether students would be proficient in English Language Arts in grade 3. The percentage of students above this threshold matched the actual 37 percent proficiency rate of the study cohort in grade 3. Accurate prediction of the proficiency rate of the full cohort of students was possible despite a large number of errors in predicting the proficiency of individual students. The threshold correctly predicted the proficiency status of 53 percent of the students who were proficient on the grade 3 state assessment and 73 percent of those who were not proficient. Students with higher scores on SDP’s annual spring reading assessments in kindergarten through grade 3 also had higher KEI scores and English language arts scores on the grade 3 state assessment.

## Why this study?

In the 2017/18 school year, nearly two-thirds of students in the School District of Philadelphia (SDP) were not proficient in reading by the end of grade 3 (SDP, 2019). Research has shown that students who fail to read at grade level by the end of grade 3 have a higher risk of negative long-term outcomes, such as failing to reach proficiency on sixth-grade reading and math assessments, grade retention, and dropping out of high school (Connor et al., 2011; Massachusetts Department of Elementary and Secondary Education and American Institutes for Research, 2013; Phillips, Norris, Osmond, & Maynard, 2002; Hernandez, 2012).

SDP is a partner in the Free Library of Philadelphia’s [Read by 4th Campaign](#), a citywide initiative that, by 2020, aims to substantially increase the percentage of students who can read on grade level when they enter grade 4. More than 100 partner organizations help support Philadelphia’s families and their children from birth through grade 3 in developing early language and literacy skills through family activities, pre-kindergarten experiences, and school and out-of-school experiences in kindergarten through grade 3. SDP has also made early literacy a focus from kindergarten through grade 3, setting as one of its four strategic “anchor” goals: “100 percent of 8-year-olds will read on or above grade level” (City of Philadelphia, 2017). To meet this goal, SDP has strengthened its early literacy curriculum and assessment systems, provided a literacy coach in every elementary school, strengthened classroom literacy environments, and distributed one million new books across the district’s kindergarten through grade 3 classrooms.

To help track progress toward the goal of proficient reading by the end of grade 3, SDP requested assistance from the Regional Educational Laboratory (REL) Mid-Atlantic to identify an indicator measuring the percentage of entering kindergarteners on-track for proficiency in English Language Arts (ELA) on the grade 3 Pennsylvania System of School Assessment (PSSA). The indicator can provide useful information on the progress of families and organizations across the city to promote children’s early development before kindergarten entry and inform the

city’s continued investments in early education. This indicator would be based on Pennsylvania’s Kindergarten Entry Inventory (KEI), an assessment that measures children’s academic, socioemotional, and physical competencies that has been used in the district since 2014.

SDP also wanted to explore the accuracy of using this indicator or alternative thresholds on the KEI, along with its other measures of early reading skills, to predict which individual students may need support in reading. In addition, SDP wanted to understand the relationships between its different measures of early reading skills—the KEI, PSSA, and AIMSweb assessments (see box 1 for key terms). SDP currently uses the thresholds set by the developer on the AIMSweb to identify individual students who need additional support. Therefore, the study examines the accuracy of developer-selected thresholds on the AIMSweb for predicting which individual students are at risk of not being proficient in reading on the grade 3 PSSA, the accuracy of thresholds set on the KEI for such predictions, and the relationships between the different measures of early reading skills. This will inform SDP’s decisions about the measures used to predict which students may need support in reading and possible future work on other indicators of being on-track for reading proficiency.

---

### Box 1. Key terms

**Accuracy.** The study examines four measures of accuracy (see box A1 in appendix A for equations):

- *Absolute difference between the actual proficiency rate and predicted proficiency rate based on the threshold.* This measures how accurately the threshold predicts the percentage of entering kindergarteners in the study cohort who are on-track for reading proficiency.
- *Percentage of proficient students whom the threshold correctly predicted.* This measures how accurately the threshold made predictions for individual students who are proficient.
- *Percentage of not proficient students whom the threshold correctly predicted.* This measures how accurately the threshold made predictions for individual students who are not proficient.
- *Percentage of students whom the threshold correctly predicted.* This measures how accurately the threshold made predictions for individual students (both those who were proficient and those who were not).

**AIMSweb Reading.** The AIMSweb Reading assessment measures children’s progress in developing reading skills between kindergarten and grade 3. The AIMSweb includes subtests that measure early literacy skills, such as letter naming fluency (naming upper and lower case letter names) in kindergarten, nonsense word fluency (associating letters with sounds and blending letter sounds into words) in grades 1 and 2, and oral reading fluency in grade 3.

Teachers administer the AIMSweb three times (fall, winter, and spring) during each school year, and each subtest takes about 1 to 4 minutes. This study focuses on the spring scores as a measure of students’ reading skills by the end of each grade. This study was not able to use the AIMSweb as an indicator for the percentage of entering kindergarteners on-track for reading proficiency because the district did not require the fall AIMSweb until 2015.

AIMSweb scores divide students into three tiers (1, 2, and 3) that are consistent with Response to Intervention, a widely used approach for identifying and supporting struggling students. In Tier 1, students are considered “at target” and receive typical instruction and periodic screening. In Tier 2, struggling students receive strategic interventions and more frequent progress monitoring. In Tier 3, students receive intensive intervention.

**Kindergarten Entry Inventory (KEI).** Since 2014, SDP has required schools to administer the KEI. The KEI measures children’s academic, socioemotional, and physical competencies at kindergarten entry using 30 indicators. A previous validation study identified two distinct dimensions on the KEI with evidence of validity (Howard et al., 2017). This study focuses on these two dimensions, as well as an index of the two dimensions constructed by the study team:

- *Emerging Academic Competencies*, which includes 12 indicators that measure early literacy and mathematics skills, such as letter recognition and counting.
- *Learning Engagement Competencies*, which includes nine indicators that measure behavioral and motor skills, such as task persistence, emotion regulation, and fine and gross motor skills.

- *Index of Emerging Academic Competencies and Learning Engagement Competencies*, which is a weighted average of these dimensions, with greater weight given to Emerging Academic Competencies, which has a stronger relationship with ELA scores on the grade 3 PSSA than Learning Engagement Competencies (equation A2 in appendix A).

Teachers administer the KEI by observing children in a variety of classroom activities during the first 45 days of instruction. Teachers give children a rating on each indicator using four rating levels: (1) Not yet evident, (2) Emerging, (3) Evident, and (4) Exceeds. For each dimension, the score is the average rating across the indicators and ranges from 1 to 4. The study team put the index on a 1 to 10 scale, in consultation with SDP, to make this scale more distinct from those of the individual dimensions. On the index, 10 indicates students scored perfectly (were rated as 4, Exceeds) on all items on the two KEI dimensions used.

**Pennsylvania System of School Assessment (PSSA) in English language arts.** This state standardized assessment measures students' reading skills, including decoding and comprehension, in the spring of grade 3. On the PSSA, performance levels can be classified as proficient (proficient or advanced) or not proficient (below basic or basic). More details about the indicators in each KEI dimension and the subtests included in the AIMSweb is in appendix A.

---

## Research questions

To identify an indicator of being on-track to read proficiently by the end of grade 3, this study addressed several research questions for the cohort of students who entered kindergarten in SDP in the 2014/15 school year. This cohort was the first to be administered the current version of the KEI assessment in fall 2014 and the only cohort for whom grade 3 PSSA scores were available when the study began:

### *Primary research questions:*

1. What is the relationship between KEI scores and scores on the grade 3 PSSA in ELA?
2. What threshold score on the KEI most accurately predicts reading proficiency on the grade 3 PSSA in ELA for the cohort? What threshold score on the KEI has a correct prediction rate of at least 90 percent for the individual students who were not proficient in ELA on the grade 3 PSSA?

### *Exploratory research question:*

3. How do AIMSweb Reading scores in the spring of kindergarten through grade 3 relate to children's scores on the KEI and grade 3 PSSA in ELA?

The primary research questions inform the development of an indicator of progress towards grade-level reading based on KEI scores. The relationship between KEI scores and the scores on the grade 3 PSSA provides information about the potential for using a KEI threshold for predicting reading proficiency in grade 3. The most accurate KEI threshold score for predicting the percentage of students that are on track for grade-level reading—SDP's primary use for the indicator—is identified.

Should SDP decide to use the KEI, in combination with other measures, to predict which individual children are at risk of not meeting grade level reading expectations, a threshold that accurately predicted more of the students who are not proficient might be more appropriate. Studies that aim to identify students at risk of not being proficient in reading often try to set thresholds on screening instruments that correctly identify 90 percent of those who are not proficient (Compton, Fuchs, Fuchs, & Bryant, 2010; Jenkins, Hudson, & Johnson, 2007). Early identification could allow these children to receive further assessment and potentially additional services to support reading growth. Yet, such a threshold should also minimize the risk of spending limited resources supporting students who may not need extensive supports. Therefore, the accuracy of alternative thresholds to predict which individual students are most likely to need support in reading is explored.

The final research question explores the relationship between the different measures of early reading skills used in SDP—the KEI, AIMSweb, and PSSA. Because SDP was not using AIMSweb scores across the district in fall 2014,

the study focused on spring AIMSweb scores between kindergarten and grade 3. The relationships between KEI scores, AIMSweb scores from the spring of kindergarten through grade 3, and grade 3 PSSA scores in ELA provide information about how students' reading proficiency changes across grades and how it is classified by different measures. For example, the AIMSweb could be more accurate at predicting grade 3 reading proficiency because it measures more specific reading skills than the validated KEI dimensions, which measure literacy, math, and behavioral skills. The AIMSweb may also help identify the specific reading skills that reading interventions should target. Therefore, information about the relationship between AIMSweb and the PSSA can also inform SDP's decisions about measures to identify students at risk of not being proficient in reading. However, these analyses were exploratory because SDP's primary goal was to develop an indicator based on the KEI that could track entering kindergarteners' progress towards reading proficiency.

The data sources, sample, and methods used to answer these research questions are described in box 2.

---

## Box 2. Data sources, sample, and methods

**Data sources.** This study used administrative student records from the School District of Philadelphia (SDP) that included information about students' school enrollments and demographic characteristics, Kindergarten Entry Inventory (KEI) scores on the two dimensions, grade 3 Pennsylvania System of School Assessment (PSSA) in English Language Arts (ELA) scores, and AIMSweb Reading scores for each spring semester from kindergarten to grade 3.

**Sample.** The study sample was 13,335 kindergartners who entered SDP by December of the 2014/15 school year. The sample did not include students in charter or private schools in Philadelphia. Of these 13,335 students, 58 to 68 percent had scores on one or both KEI dimensions. Students with both KEI dimension scores were more likely to be White than all kindergartners (table A7 in appendix A). In addition, of students with KEI scores, 65 percent also had ELA scores on the grade 3 PSSA (table A4 in appendix A).

For research questions 1 and 2, the sample was restricted to a random sample of 70 percent of the students who had scores on both KEI dimensions in 2014 and took the grade 3 PSSA on-time in 2018 (3,521 students; see the methods below for explanation of why analyses used a random sample of 70 percent of students). An analysis for research question 2 also used the remaining 30 percent random sample of students (1,508 students). The samples for research question 3 were restricted to a 70 percent random sample and included enrolled kindergarteners who completed the spring AIMSweb and either the KEI or PSSA in each of the years that the 2014/15 kindergarteners were expected to complete them if they remained on grade level (3,689 to 5,446 students).

**Methods.** Following the typical approach for predictive analyses, the study team split the sample randomly into two parts, using a method that ensured that each part contained the same percentages of students who were and were not proficient on the PSSA. The study team conducted most analyses on 70 percent of the sample to select the threshold (described as the sample used to set the threshold) and then checked the accuracy of the predictions using the remaining 30 percent.

The study team used descriptive statistics and regression analyses to analyze the relationships between KEI and PSSA scores (research question 1). To assess the relationship between each KEI dimension score and students' PSSA score, the study team used pairwise correlations. Positive correlations were classified as follows: weak (0.1–0.39), moderate (0.4–0.69), strong (0.7–0.99), and perfect (1; Dancey & Reidy, 2007). The study team used logistic regression analyses to predict the likelihood that a student was proficient on the grade 3 PSSA based on their scores on the KEI dimensions (both separately and combined). To make the sample used in the analysis (the sample of kindergarteners with both KEI and PSSA scores) representative of the sample to whom the threshold will be applied (the sample of kindergarteners with KEI scores), the study team used weights for the analyses for research questions 1 and 2. These weights adjust for the fact that some students with KEI scores were missing PSSA scores, and thus, are excluded from the predictions. Students with PSSA scores had higher KEI scores on both dimensions, were more likely to be Hispanic/Latino, and were less likely to receive special education services than students without PSSA scores (table A8 in appendix A). The analysis adjusts for the exclusion of these students by giving greater weight to students with similar KEI scores and characteristics who did have PSSA scores (findings using alternative weights are shown in appendix B).

To select a threshold and assess its accuracy (research question 2), the study team and SDP first considered three potential indicators on which the threshold could be set: Emerging Academic Competencies dimension only, Learning Engagement Competencies dimension only, or an index of both Emerging Academic and Learning Engagement Competencies (see appendix A for detail on the construction of the index). To provide information about deciding between these three potential indicators, the study team calculated two statistics that provide information about the accuracy of thresholds for predicting individual students' proficiency: the percentages of proficient students and students who were not proficient whom the threshold correctly predicted. Over the course of the study, the study team and SDP staff discussed these different indicators, and SDP decided to use an index based on both dimensions as the indicator of whether students were on-track for reading proficiently in grade 3. The study team then identified the threshold for this index that had the smallest absolute difference between the actual and predicted proficient rates and rounded to the nearest whole number to more easily communicate the threshold.

To help shed light on the potential usefulness of the overall threshold for predicting reading proficiency for future cohorts of students, the accuracy of the threshold was also assessed for the remaining 30 percent sample of students (who had not been used to set the threshold), based on these three statistics and the percentage of all students whom the threshold correctly predicted. As an exploratory analysis, the study team also used this method to explore the accuracy of alternative thresholds on the KEI index that made correct predictions for at least 90 percent of individual students who were at risk of not being proficient in ELA on the grade 3 PSSA.

Descriptive statistics were used to explore the relationships between AIMSweb and KEI and PSSA scores (research question 3). As with research question 1, pairwise correlations were calculated between students' scores on each assessment. Weights were not used in the analyses that included the AIMSweb scores because the analyses were exploratory and not intended to set an indicator that could be applied to a representative population of students (such as all students with AIMSweb scores in a particular grade). Measures of accuracy (correctly predicting students who were and were not proficient on the grade 3 PSSA in ELA) were calculated for the developer-selected AIMSweb thresholds, using the equations shown in box A1 in appendix A.

More information about the study data sources, sample, and methods is in appendix A.

---

## Findings

This section presents findings to address the three study research questions. Additional findings are available in appendix B.

### ***Students with higher Kindergarten Entry Inventory scores had higher scores on the grade 3 Pennsylvania System of School Assessment in English language arts***

Students' KEI scores were weakly to moderately related to ELA scores on the grade 3 PSSA. Emerging Academic Competencies had a correlation of 0.42 with ELA scores on the grade 3 PSSA, and Learning Engagement Competencies had a correlation of 0.32 with the same grade 3 test scores (table 1). The stronger correlation between Emerging Academic Competencies and grade 3 PSSA in ELA scores, relative to that for Learning Engagement Competencies, may reflect the fact that the Emerging Academic Competencies dimension includes measures of early reading skills (as well as math skills). In contrast, the Learning Engagement Competencies dimension includes measures of behavioral and motor skills that are less likely to be strongly correlated with grade 3 reading skills.

**Table 1. Both Kindergarten Entry Inventory dimensions had weak to moderate correlations with grade 3 Pennsylvania System of School Assessment scores**

KEI dimension	Correlation with grade 3 PSSA scores
Emerging Academic Competencies	0.42***
Learning Engagement Competencies	0.32***

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

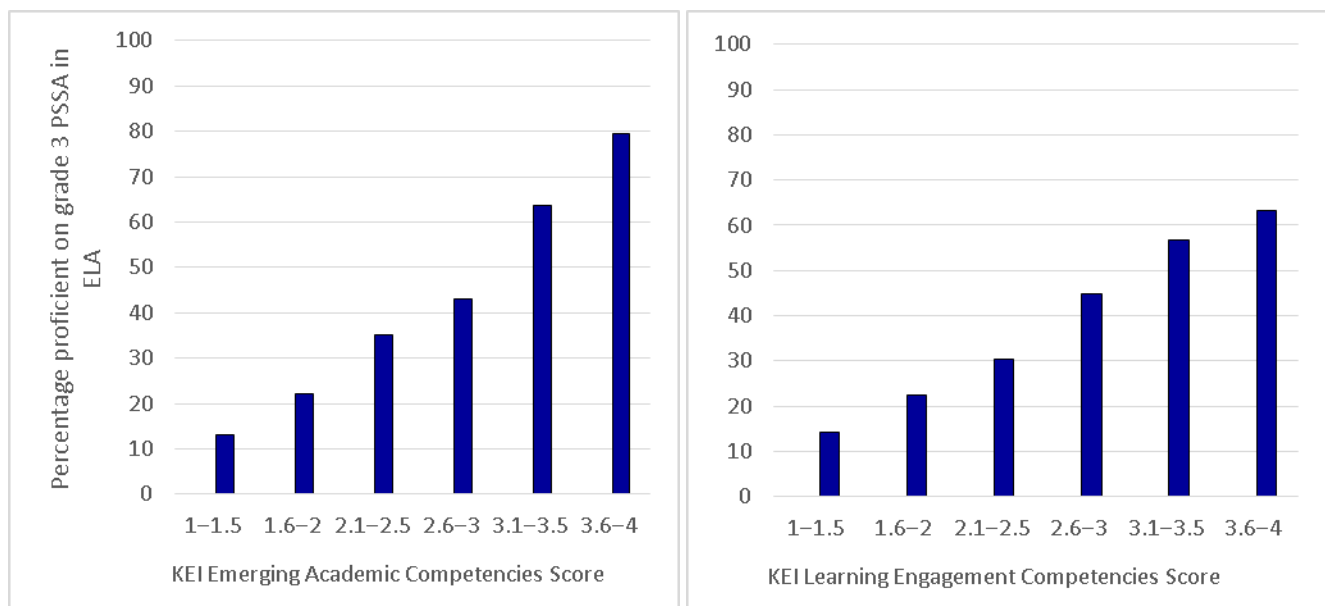
\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*Students with higher KEI scores were more likely to receive a score in the proficient level on the grade 3 PSSA in ELA.* For both Emerging Academic Competencies and Learning Engagement Competencies, PSSA proficiency rates ranged from less than 15 percent for students with the lowest scores on each KEI dimension (1–1.5) to over 60 percent for students with the highest scores on each dimension (3.6–4; figure 1). However, PSSA proficiency rates increased more steeply with scores on Emerging Academic Competencies (left graph) than with those on Learning Engagement Competencies (right graph), consistent with the higher correlation found for Emerging Academic Competencies.

**Figure 1. Students with higher Kindergarten Entry Inventory scores were more likely to be proficient on the grade 3 Pennsylvania System of School Assessment in English Language Arts**



ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*After accounting for Emerging Academic Competencies, Learning Engagement Competencies had little relationship with ELA scores on the grade 3 PSSA.* When examined separately, both Emerging Academic Competencies and Learning Engagement Competencies had statistically significant relationships with students' grade 3 PSSA proficiency in ELA (model 1 and model 2, table 2). On average, students who scored 1-point higher (on the 1-to-4 point scale) on Emerging Academic Competencies were 25 percent more likely to be proficient on the grade 3 PSSA (model 1), and students who scored 1-point higher on Learning Engagement Competencies were 24 percent more likely to be proficient on the grade 3 PSSA (model 2). When both dimensions were included (model 3), only Emerging Academic Competencies had a statistically significant relationship with grade 3 PSSA proficiency in ELA, which may be because Learning Engagement Competencies was strongly correlated

with Emerging Academic Competencies (correlation = 0.73) (not shown). In addition, the model that included both KEI dimensions had the same ability to explain proficiency as the model that only included Emerging Academic Competencies (as evidenced by these models having the same Pseudo R-squared statistic) (table 2). This suggests that after accounting for students' early reading and math skills (as measured by Emerging Academic Competencies), the behavioral and motor skills measured by Learning Engagement Competencies do not further explain students' grade 3 PSSA proficiency.

**Table 2. Emerging Academic Competencies were more predictive of students' probability of being proficient on the grade 3 Pennsylvania System of School Assessment than were Learning Engagement Competencies**

KEI dimensions	Relationship between KEI scores and the probability of being proficient on the grade 3 PSSA in ELA		
	Model 1: Emerging Academic Competencies	Model 2: Learning Engagement Competencies	Model 3: Both dimensions
Emerging Academic Competencies	0.25***		0.24***
Learning Engagement Competencies		0.24***	0.02
<b>Pseudo R-squared</b>	<b>0.09</b>	<b>0.05</b>	<b>0.09</b>

ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA. Reported values are average marginal effects, or the average increase in the predicted probability of PSSA proficiency associated with a 1-point increase in a KEI score. That is, in model 1, for every 1-point increase in Emerging Academic Competencies, the probability of proficiency increases on average 25 percent. Pseudo R-squared is a measure of how well the model fits the data compared to a model with no predictors.

Source: Authors' analysis based on SDP 2014/15 and 2017/18 data described in appendix A.

***A threshold of 6 out of 10 on an index of Emerging Academic Competencies and Learning Engagement Competencies accurately predicted the actual grade 3 proficiency rate of students (37 percent).***

Thresholds on Emerging Academic Competencies and the index based on both dimensions had similar rates of correctly predicting students' proficiency. The threshold of 6 on the index of both dimensions correctly predicted the proficiency status of 66 percent of students.

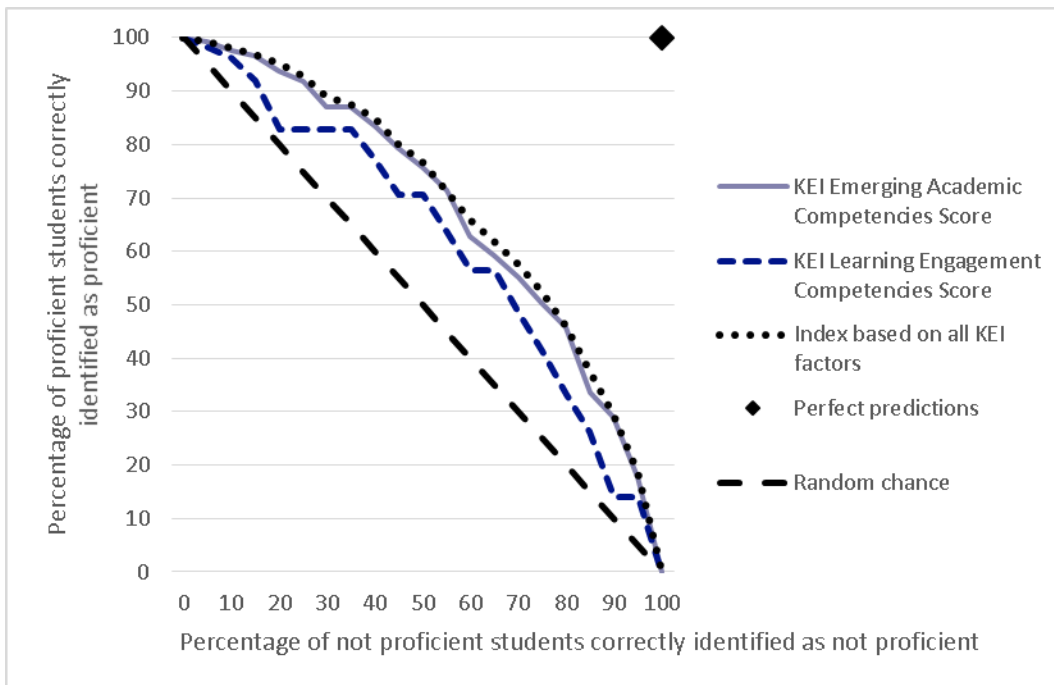
*Thresholds based on the Emerging Academic Competencies score or the index that included this score more accurately predicted grade 3 PSSA proficiency in ELA than the Learning Engagement Competencies score alone.* For example, the thresholds on Emerging Academic Competencies and the index of both dimensions that correctly predicted about 50 percent of proficient students also correctly predicted about 75 percent of students who were not proficient (figure 2). There is no threshold on Learning Engagement Competencies that performs better.

When identifying a threshold, there are tradeoffs between correctly predicting the proficiency status of students who were proficient and those who were not (as shown in figure 2). Thresholds that are set closer to the lowest scores will increase correct predictions for proficient students at the cost of decreasing correct predictions for students who were not proficient; setting a threshold closer to the highest possible score will have the opposite effect.

To decide whether to use Emerging Academic Competencies or the index of both dimensions as the indicator, SDP considered how much information each one included. SDP chose to set a threshold on the index of Emerging Academic Competencies and Learning Engagement Competencies for the indicator because it used more of the information available on children's kindergarten competencies.



**Figure 2. Thresholds based on indices that included Emerging Academic Competencies correctly predicted more proficient students and more students who were not proficient than those based on only Learning Engagement Competencies**



ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment. Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA.

Figure reads: On Learning Engagement Competencies, the threshold that correctly predicted 50 percent of students who were not proficient also correctly predicted about 71 percent of proficient students.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*At a threshold of 6 out of 10 on the index using both KEI dimensions, the percentage of students above the threshold closely matched the actual proficiency rate of students.* This was true for both the sample used to select the threshold and the sample used to test the threshold. The selected threshold accurately predicted the percentage of proficient students for the sample of kindergarteners from 2014/15 that had not been used to select the threshold (table 3). This evidence supports the threshold's potential accuracy when applied to new data for future cohorts of kindergarteners.

The sample used to create the threshold is based on students who entered SDP in kindergarten, had KEI scores, made normal grade progress, and took the grade 3 PSSA. However, the predicted proficiency rates are also similar to the proficiency for all SDP grade 3 students in 2018 (36 percent), providing additional evidence of the validity of this indicator and threshold (SDP, 2019).

**Table 3. The threshold had similar rates of accuracy in samples that were and were not used to select the threshold**

Sample	Actual proficiency (percentage)	Predicted proficiency (percentage)	Absolute difference between actual and predicted proficiency (percentage points)	Correct predictions for			Number of students
				Proficient students (percentage)	Not proficient students (percentage)	All students (percentage)	
<b>Used to set the threshold</b>	37	37	1	53	73	66	3,521
<b>Not used to set the threshold</b>	37	37	0	58	76	69	1,508

KEI is Kindergarten Entry Inventory.

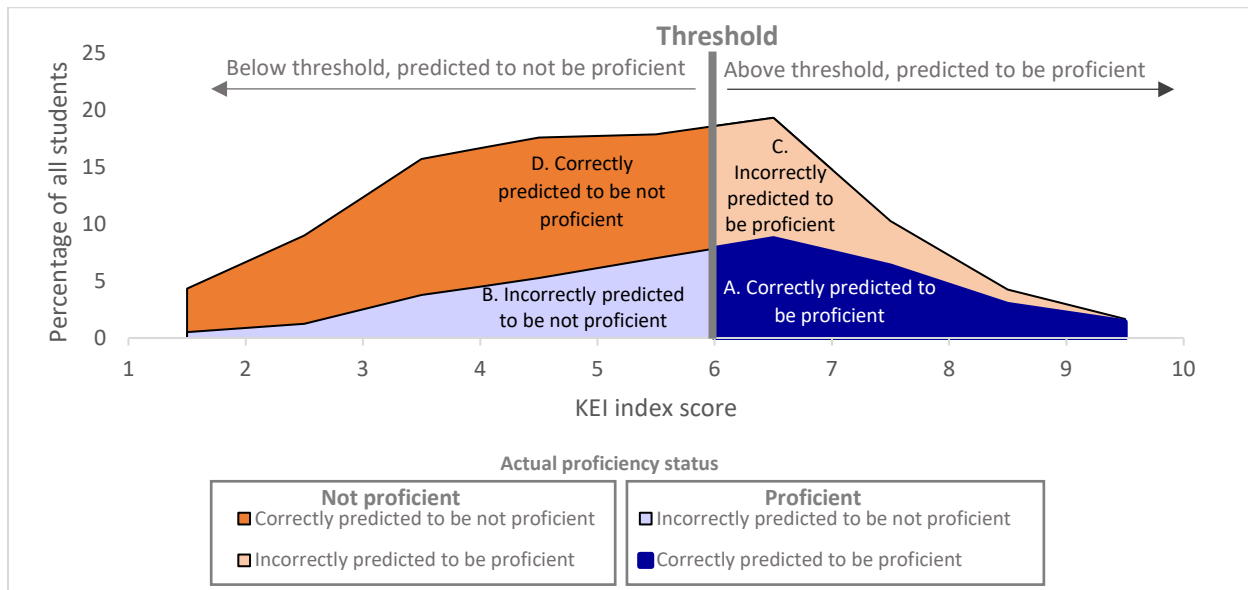
Note: Models were weighted to represent all students with KEI scores.

The difference between actual and predicted proficiency may not equal the difference between the actual and predicted proficiency columns in the table due to rounding.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*The threshold selected to make predictions for the cohort of kindergarten students made correct predictions for 53 percent of proficient students and 73 percent of students who were not proficient.* In the sample used to select the threshold, 37 percent of students were above the threshold, and thus predicted to be proficient (table 3; areas A and C in figure 3), matching the percentage who were actually proficient (table 3; areas A and B in figure 3). However, this does not mean that the threshold correctly predicted the proficiency of every individual student (figure 3). It means that the incorrect predictions for individual students cancel each other out in the aggregate (that is, areas B and C in figure 3 are equal). Overall, the threshold correctly predicted the proficiency status of 66 percent of students, which is higher than the 50 percent expected to be correctly predicted by chance, but still results in some errors for individual students. This illustrates the fact that an indicator can be very accurate at the aggregate level even while making errors at the individual level.

**Figure 3. A threshold of 6 resulted in the percentage of students above the threshold matching the actual proficiency rate**



ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA. Data were aggregated to smooth out the distribution, so no data are displayed at exactly 1 or 10.

Figure reads: At this threshold, students who were actually proficient (areas A and B) were either correctly predicted to be proficient (area A) or incorrectly predicted to be not proficient (area B). Students who were actually not proficient (areas D and C) were either correctly predicted to be not proficient (area D) or incorrectly predicted to be proficient (area C).

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*A threshold of 7 on the index of both dimensions made correct predictions for 91 percent of students who were not proficient on the PSSA and only 29 percent of students who were proficient. Compared with the threshold of 6, the higher threshold of 7 correctly predicted more of the students who were not proficient and fewer of the proficient students (table 4; figure B1 in appendix B; see table B4 in appendix B for statistics for additional thresholds).*

**Table 4. A threshold of 7 on the index of both Kindergarten Entry Inventory dimensions correctly predicted 91 percent of students who were not proficient and 29 percent of proficient students**

Threshold on the index using both KEI dimensions	Predicted proficiency (percentage)	Absolute difference between actual and predicted proficiency (percentage points)	Correct predictions for		
			Proficient students (percentage)	Not proficient students (percentage)	All students (percentage)
6.00	37	1	53	73	66
7.00	17	21	29	91	68

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores.

Source: Authors' analysis based on SDP 2014/15 and 2017/18 data described in appendix A.

***Students with higher AIMSweb scores in kindergarten through grade 3 had generally higher Kindergarten Entry Inventory scores and grade 3 ELA scores on the Pennsylvania System of School Assessment***

Each measure of students' reading skills from kindergarten through grade 3 were positively correlated with one another. Findings suggest that thresholds on the KEI could also be used to predict whether student cohorts are

on-track for meeting the grade-level target on the grade 3 AIMSweb.<sup>1</sup> The results also suggest that AIMSweb scores could be used, along with additional measures, to predict which students are at risk for not being proficient in reading by grade 3. However, the KEI score threshold resulted in more accurate predictions of the overall percentage of students proficient in reading by grade 3 than the annual AIMSweb grade-level targets.

*Students’ KEI scores on Emerging Academic Competencies generally had moderate positive correlations with their AIMSweb scores from spring of kindergarten through grade 3.* Emerging Academic Competencies scores had generally moderate associations with AIMSweb scores (0.39 to 0.51), as opposed to the weak associations between AIMSweb and Learning Engagement Competencies scores (0.29 to 0.39; table 5). This likely reflects Emerging Academic Competencies being a better measure of students’ early reading skills than Learning Engagement Competencies.

The KEI scores on both Learning Engagement Competencies and Emerging Academic Competencies had the strongest associations with spring kindergarten AIMSweb scores (0.39 and 0.51 respectively). The associations between KEI and AIMSweb scores generally decreased as students moved from kindergarten through grade 3. This is consistent with students’ early reading proficiency changing over time, and past research that shows stronger correlations between measures at closer points in time (Duncan et al., 2007). Associations between the KEI and the AIMSweb in the spring of grade 3 were identical to those between the KEI and ELA scores on the PSSA in grade 3. This suggests that the KEI could also be used to predict whether student cohorts are on-track for meeting their grade-level target on the spring AIMSweb assessments in grade 3 or earlier.

**Table 5. Kindergarten through grade 3 AIMSweb scores had weak to moderate correlations with Emerging Academic Competencies and Learning Engagement Competencies**

KEI (fall of kindergarten)	Spring kindergarten AIMSweb	Spring grade 1 AIMSweb	Spring grade 2 AIMSweb	Spring grade 3 AIMSweb	PSSA ELA grade 3
<b>Correlations</b>					
Emerging Academic Competencies	0.51***	0.39***	0.46***	0.42***	0.42***
Learning Engagement Competencies	0.39***	0.29***	0.33***	0.32***	0.32***
<b>Number of students</b>					
Emerging Academic Competencies	4,930	3,689	3,948	3,707	3,697
Learning Engagement Competencies	5,446	4,200	4,460	4,177	4,169

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to students who completed the KEI and either the AIMSweb or PSSA. Correlations for AIMSweb scores were not weighted.

Source: Authors’ analysis based on SDP 2014–2018 data described in appendix A.

*Students’ AIMSweb scores from kindergarten through grade 3 were moderately to strongly related to their ELA scores on the grade 3 PSSA.* The strength of the associations between AIMSweb scores and grade 3 PSSA scores generally increased from kindergarten to grade 3 (0.50 to 0.73; table 6). Correlations were stronger between the AIMSweb and PSSA than between both KEI dimensions and the PSSA. This could be because the AIMSweb was measured closer in time to the PSSA or could indicate that the skills measured by the AIMSweb are a stronger predictor of the skills measured by the PSSA. These findings suggest that the AIMSweb in grades K through 3 could be used to develop an indicator of students’ likelihood of being proficient in ELA on the grade 3 PSSA.

<sup>1</sup> For more information about the relationship between the KEI and AIMSweb tiers, see figures B1-B4 in appendix B.

**Table 6. Kindergarten through grade 3 AIMSweb scores were moderately to strongly correlated with grade 3 Pennsylvania System of School Assessment English Language Arts scores**

Assessment	Grade 3 PSSA ELA	
	Correlation	Number of students
<b>KEI</b>		
Emerging Academic Competencies (fall kindergarten)	0.42***	3,697
Learning Engagement Competencies (fall kindergarten)	0.32***	4,169
<b>AIMSweb</b>		
Spring kindergarten AIMSweb	0.50***	4,696
Spring grade 1 AIMSweb	0.53***	4,174
Spring grade 2 AIMSweb	0.73***	5,027
Spring grade 3 AIMSweb	0.72***	5,266

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

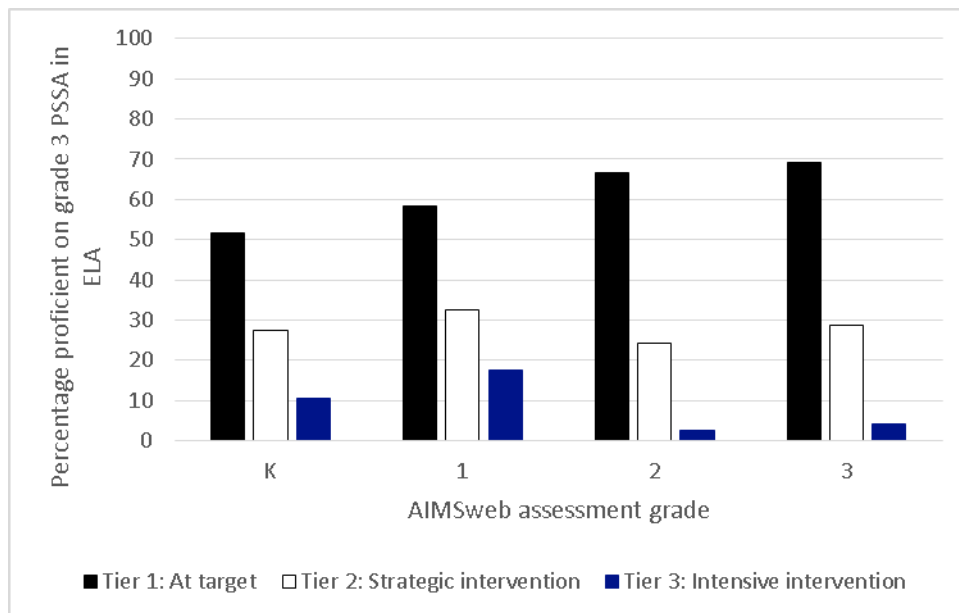
\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analysis was based on the random sample of students used to set the threshold and is restricted to students who completed the PSSA and either the KEI or AIMSweb. Correlations for AIMSweb and PSSA scores were not weighted. For comparison, unweighted correlations for the KEI and PSSA are shown.

Source: Authors' analysis based on School District of Philadelphia 2014–2018 data described in appendix A.

*Students in higher performing AIMSweb tiers in the spring of each year were more likely to be proficient on the grade 3 PSSA in ELA.* SDP uses cutoffs on the AIMSweb to identify individual students for support in reading. More than half of the students who meet the AIMSweb's grade-level target (Tier 1) in spring of kindergarten were proficient in reading on the PSSA by grade 3 (figure 4). In contrast, less than 25 percent of students identified for intensive support (Tier 2) and less than 10 percent of those identified for the most intensive support (Tier 3) in kindergarten were proficient in reading on the PSSA by grade 3 (table B5 in appendix B). These classifications become increasingly accurate as assessments become closer to the grade 3 PSSA.

**Figure 4. Students in higher performing AIMSweb tiers were more likely to be proficient on the grade 3 Pennsylvania System of School Assessment**



ELA is English language arts. PSSA is Pennsylvania System of School Assessment.

Note: Analysis was based on random sample of students used to set the threshold and was restricted to students who completed the PSSA and AIMSweb: 4,174 to 5,266 students had information on the grade 3 PSSA and AIMSweb in grades K to 3. Analysis was not weighted.

Source: Authors' analysis based on SDP 2014–2018 data described in appendix A.

On the spring AIMSweb in grades K through 3, the cutoff for the Tier 1 grade-level target correctly predicted 52 to 78 percent of students who were not proficient in ELA on the grade 3 PSSA and 65 to 83 percent of students who were proficient. The Tier 1 cutoff on the AIMSweb in the spring of kindergarten correctly predicted fewer students who were not proficient on the grade 3 PSSA than the study-selected KEI threshold of 6 (52 versus 73 percent) but predicted more of the students who were proficient (83 versus 53 percent) (table 7).

In contrast, by spring of grade 2, the Tier 1 cutoff correctly predicted 76 percent of students who were not proficient on the grade 3 PSSA and 81 percent of students who were proficient (table 7). This further demonstrates the increased accuracy of the AIMSweb for predicting reading proficiency assessed at closer points in time.

**Table 7. The AIMSweb Tier 1 cutoff in later grades correctly predicted the proficiency status of more students than the AIMSweb Tier 1 cutoff in kindergarten**

Assessment and threshold	Actual proficiency (percentage)	Predicted proficiency (percentage)	Absolute difference between actual and predicted proficiency (percentage points)	Correct predictions for:			Number of students
				Proficient students (percentage)	Not proficient students (percentage)	All students (percentage)	
<b>KEI index threshold (fall of kindergarten)</b>							
6.00	37	37	1	53	73	66	3,521
7.00	37	17	21	29	91	68	3,521
<b>AIMSweb tier 1 cutoff (spring of each grade)</b>							
Kindergarten	37	61	23	83	52	64	4,696
Grade 1	36	42	6	65	71	69	4,174
Grade 2	37	45	8	81	76	77	5,027
Grade 3	37	42	5	76	78	77	5,266

Note: The table is based on the sample used to select the threshold and the AIMSweb data was not weighted. The KEI data was weighted to represent all students with KEI scores. The difference between actual and predicted proficiency may not equal the difference between the actual and predicted proficiency columns in the table due to rounding.

Source: Authors' analysis based on School District of Philadelphia data 2014–2018 data described in appendix A.

The AIMSweb tier cutoffs in the spring of kindergarten through grade 3 were not designed as an indicator of the overall percentage of students who are on-track for reading proficiency and were less accurate than the study-selected KEI threshold of six from the fall of kindergarten for predicting the overall percentage of children proficient in reading by grade 3 (table 7).

### Limitations

Three limitations should be kept in mind. First, the study focuses on a single cohort of students in SDP—those who entered kindergarten in the 2014/15 school year—two-thirds of whom had KEI scores. The threshold might or might not yield the same accuracy when predicting grade 3 reading proficiency of future cohorts of students in SDP, who may have more complete data on KEI scores and will take the PSSA in a different school year. The study also does not include students in Philadelphia charter and private schools, or students from other districts in Pennsylvania, so the accuracy of the threshold might also differ for these groups of students. Also, changes in KEI administration or instruction and supports for students between kindergarten and grade 3 could generate different relationships than those described. SDP has improved teacher training on the KEI administration and has created initiatives intended to increase reading supports to struggling students between kindergarten and grade 3, such as providing instructional coaches focused on early literacy and strengthening classroom literacy

environments. These additional supports could alter the relationship between KEI and PSSA scores in ELA. Likewise, changes in the AIMSweb assessment (such as the change to AIMSwebPLUS in the 2018/19 school year) could alter the relationships between the KEI and PSSA and those scores.

Second, to account for students who were missing scores on the PSSA in ELA when calculating whether kindergarteners are on-track, the study reweighted the analyses based on students' KEI scores and demographic characteristics. This assumes that students who were missing PSSA scores in grade 3 would have received the same scores as those with similar KEI scores and background characteristics in kindergarten. However, students who were missing PSSA scores in grade 3 might have had different scores than assumed. For example, students who moved out of SDP might have characteristics that were not measured in the data (such as family instability), which could reduce their achievement by grade 3.<sup>2</sup> Students who were missing KEI scores were less likely to be White than all kindergarteners. For students who were missing KEI scores, limited data (only background characteristics) were available for predicting their likely proficiency in reading in grade 3. Thus, the study's main analyses did not make predictions for students who are missing KEI scores.

Third, this study only analyzed the two dimensions of the Pennsylvania KEI—Emerging Academic Competencies and Learning Engagement Competencies—for which the study team received data. These dimensions were validated in a study of the KEI (Howard et al., 2017). However, it is possible that particular KEI items within these dimensions (for example, early reading skills in the Emerging Academic Competencies dimension) or items outside of these dimensions (for example, language comprehension skills) might have even stronger correlations with grade 3 PSSA scores in ELA. Alternative assessments, such as the fall kindergarten AIMSweb used by SDP or kindergarten entry assessments used by other states, might have different relationships with reading proficiency in grade 3. For example, other states' kindergarten entry assessments could have stronger relationships with later reading proficiency if they more accurately measure skills that are more predictive of reading proficiently or if states train teachers to rate students' skills more accurately (Jenkins, Hudson, & Johnson, 2007).

## Implications

SDP can apply the study-established threshold of six on a 1–10 index based on two KEI dimensions (Emerging Academic Competencies and Learning Engagement Competencies) to the KEI scores of future kindergarten cohorts, allowing SDP to examine changes over time in the percentages of students who are on-track for reading proficiently in grade 3. Differences in this indicator for future cohorts might reflect the efforts of Philadelphia's Read by 4th partnership's citywide investments in early childhood education from birth to age 5, such as parent education about reading to children and the expansion of prekindergarten in the city. Differences in this indicator could also reflect changing composition of the student population, such as changing demographics.

SDP could also use data on more recent cohorts of kindergarteners to reassess the accuracy of and potentially update the threshold for the indicator of whether entering kindergarteners are on-track for reading proficiently. The study's analyses were based on approximately two-thirds of kindergarteners in SDP who had KEI scores in 2014, the first year the district implemented a new version of KEI. SDP staff indicated that the KEI was more consistently administered in recent years. Moreover, should KEI data become available for children in charter and private schools, the analyses could be extended to those groups.

SDP may also wish to use a different threshold on the KEI, AIMSweb scores, other measures, or some combination to predict which individual students are at risk of not being proficient in reading, as opposed to predicting cohort-level proficiency rates. Previous research has suggested selecting thresholds that correctly identify 90 percent of students who are at risk of not being proficient (Compton et al. 2010, Jenkins et al. 2007). Yet, in early grades,

---

<sup>2</sup> Students might also be missing PSSA scores because they were absent on the assessment date or were not promoted to grade 3. However, SDP staff indicated that few students are retained in earlier grades.

such thresholds often have low rates of correctly identifying students who will be proficient, requiring decisions about the appropriate tradeoff between correctly identifying students who will and will not be proficient (Hintze, Ryan, & Stoner, 2003; O'Connor & Jenkins, 1999). Consistent with findings from previous research, the threshold of 7 on the 1–10 KEI index correctly predicted over 90 percent of students who will not be proficient on the grade 3 PSSA in ELA while correctly predicting 29 percent of students who will be proficient. The Tier 1 grade-level targets on the spring AIMSweb in kindergarten correctly predicted 52 percent of students who will not be proficient, but 83 percent of students who will be proficient. In later grades, the Tier 1 grade-level targets correctly predicted between 71 and 78 percent of not proficient students, but had overall correct prediction rates above both KEI thresholds. SDP's decisions about which threshold or measures to use will depend on how the district balances the costs of identifying students who truly are not at risk versus the costs of not identifying students who truly are at risk (Elliot, Hai, & Roach, 2007).

The study also highlights the importance of SDP frequently reassessing students' reading skills at early ages to determine which students need additional reading supports. Students' reading skills change markedly between entering kindergarten and grade 3, as students acclimate to school and learn from their teachers, parents, and peers over multiple school years (O'Connor & Jenkins, 1999). In fact, KEI scores from the beginning of kindergarten had weaker correlations with grade 3 PSSA scores than AIMSweb scores at the end of kindergarten, and the strength of correlations between AIMSweb scores and grade 3 PSSA scores generally increased from kindergarten to grade 3.

SDP could consider developing other indicators to identify whether individual students need support in reading or to measure the percentages of students who are on-track for reading proficiently at the end of kindergarten through grade 3. Including multiple measures from the AIMSweb or KEI could reduce errors in the predictions for identifying individual students who need supports.

Future research could also explore creating indicators of the percentages of students who are on-track for reading proficiently based on the AIMSweb. For example, SDP could use successive spring AIMSweb scores to predict the percentages of students at the end of kindergarten through grade 2 who are on-track for grade 3 PSSA proficiency. Or, SDP could use AIMSweb scores to create indicators that measure cohorts' changes in reading skills from year to year. For example, they could measure the percentage of students in grade 1 who are on-track to be at the AIMSweb grade-level target by the end of grade 2. That research would further inform ongoing district efforts to support children's reading achievement growth from kindergarten through grade 3 by identifying specific times in students' academic careers that may disproportionately predict their likelihood of reading proficiency.

This study provides a basis for the use of kindergarten entry assessments that other states could consider. Many other states are in the process of understanding how to use kindergarten assessments to inform instruction, identify students in need of additional support, and measure children's school readiness and progress in early grades (Shields, Cook, & Greller, 2016). The moderate correlations between KEI Emerging Academic Competencies and ELA scores on the grade 3 PSSA (0.43) are in line with past literature on the relationship between children's early skills and later academic achievement (Duncan et al., 2007; LaParo & Pianta, 2000). These links between teacher-reported early learning skills and children's reading proficiency in later grades support the use of kindergarten entry assessments as an aggregate indicator of children's later skills. They also potentially support the use of these assessments as an early indicator for identifying individual students who need additional support, though this may be more error-prone than an indicator of the overall proportion of students expected to be proficient. In combination with additional assessments, an early indicator could bring earlier reading intervention services to improve students' chances of becoming proficient readers.



## References

- City of Philadelphia. (2017). Ensuring Quality Schools for Every Philadelphia Child. Retrieved from <https://whyy.org/wp-content/uploads/2017/11/Ensuring-Quality-Schools-for-Every-Philadelphia-Child.pdf>
- Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2010). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98*(2), 394. Retrieved from <https://eric.ed.gov/?q=%22selecting+at+risk+readers%22&id=EJ884854>
- Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J. R., Lundblom, E., Crowe, E. C., & Fishman, B. (2011). Effective classroom instruction: Implications of child characteristics by reading instruction interactions on first graders' word reading achievement. *Journal of Research on Educational Effectiveness, 4*(3), 173–207. Retrieved from <https://eric.ed.gov/?q=%22effective+classroom+instruction%22+AND+%22child+characteristics%22&id=EJ932552>
- Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Harlow, UK: Pearson Education.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. Retrieved from <https://eric.ed.gov/?q=%22school+readiness+and+later%22&id=EJ779938>
- Elliot, S. N., Haiu, N., & Roach, A. T. (2007). Universal and early screening for educational difficulties: Current and future approaches. *Journal of School Psychology, 45*, 27–51. Retrieved from <https://eric.ed.gov/?q=%22universal+and+early+screening%22&id=EJ751874>
- Hernandez, D.J. (2012). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: Annie E. Casey Foundation. Retrieved from <http://www.aecf.org/m/resourcedoc/AECF-DoubleJeopardy-2012-Full.pdf>
- Hintze, J., Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the dynamic indicators of basic early literacy skills and the comprehensive test of phonological processing. *School Psychology Review, 32*, 541–556. Retrieved from <https://eric.ed.gov/?q=%22concurrent+validity+and+diagnostic%22&id=EJ823574>
- Howard, E.C., Dahlke, K., Tucker, N., Liu, F., Weinber, E., Williams, R., ... Brumley, B. (2017). *Evidence-based kindergarten entry inventory for the Commonwealth: A journey of ongoing improvement*. Washington, DC: American Institutes for Research.
- Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a Response to Intervention framework. *School Psychology Review, 36*(4), 582–600. Retrieved from <https://eric.ed.gov/?q=%22screening+for+at-risk+readers%22&id=EJ788363>
- Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review, 37*(3), 374. Retrieved from <https://eric.ed.gov/?q=%22long-term+diagnostic+accuracy%22&id=EJ817300>
- LaParo, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research, 70*(4), 443–484. Retrieved from <https://eric.ed.gov/?q=%22predicting+childrens%27+competence%22&id=EJ642216>
- Massachusetts Department of Elementary and Secondary Education and American Institutes for Research (2013). Technical descriptions of risk model development: Early and late elementary age groupings (grades 1-6). Author. Retrieved from <http://www.doe.mass.edu/ccr/ewi/EWISrisk-EE-LE.pdf>
- National Assessment of Education Progress. (2017). NAEP Technical Documentation: NAEP Assessment Weighting Procedures. Retrieved from: <https://nces.ed.gov/nationsreportcard/tdw/weighting/>
- O'Connor, R. E., & Jenkins, J. R. (1999). The prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3*(2), 159–197. Retrieved from <https://eric.ed.gov/?q=%22prediction+of+reading+disabilities%22&id=EJ585457>
- Pearson. (2019). AIMSwebPlus Overview Brochure. Retrieved from: <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/aimswebPlus-overview.pdf>

- Pennsylvania Department of Education. (2018). *2018 Pennsylvania System of School Assessment technical report: mathematics, English Language Arts, and Science*. Harrisburg, PA: Pennsylvania Department of Education. Retrieved from <https://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PSSA/Technical%20Reports/2018%20PSSA%20Technical%20Report.pdf>
- Pennsylvania Office of Child Development and Early Learning (2017). *A guide to using the Pennsylvania Kindergarten Entry Inventory*. Harrisburg, PA: Pennsylvania Office of Child Development and Early Learning. Retrieved from <https://www.philasd.org/earlychildhood/wp-content/uploads/sites/835/2018/07/2018-KEI-Manual-Updated.pdf>
- Phillips, L. M., Norris, S. P., Osmond, W.E., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology*, 94(1), 3–13. Retrieved from <https://eric.ed.gov/?q=%22relative+reading+achievement%22&id=EJ644659>
- Quirk, M., Dowdy, E., Goldstein, A., & Carnazzo, K. (2017). School readiness as a longitudinal predictor of social-emotional and reading performance across the elementary grades. *Assessment for Effective Intervention*, 42(4), 248–253. Retrieved from <https://eric.ed.gov/?q=%22school+readiness+as+a+longitudinal%22&id=EJ1151222>
- School District of Philadelphia. (2017, September 27). *Philadelphia's students make academic gains* [Press release]. Retrieved from <https://www.philasd.org/blog/2017/09/27/philadelphias-students-make-academic-gains/>
- School District of Philadelphia. (2019). District score card. Retrieved from <https://www.philasd.org/performance/programs-services/school-progress-reports/district-scorecard/>
- Shields, K. A., Cook, K. D., & Greller, S. (2016). How kindergarten entry assessments are used in public schools and how they correlate with spring assessments (REL 2017–182). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>

## Appendix A. Methods

This appendix describes the data, samples, weights, and analysis methods for the Regional Educational Laboratory (REL) Mid-Atlantic study to identify a school readiness indicator for the School District of Philadelphia (SDP).

### Data

The data used for this study came from SDP’s administrative database and included (1) students’ school enrollment records and demographic characteristics from the 2014/15 school year, (2) Kindergarten Entry Inventory (KEI) scores for students who entered kindergarten in 2014, (3) Pennsylvania System of School Assessment (PSSA) in English language arts (ELA) scores for students who took the grade 3 PSSA assessment in 2018, and (4) AIMSweb Reading scores for each spring semester from 2015 (kindergarten) to 2018 (grade 3).

### School enrollment records and demographic characteristics

This study used enrollment records for students enrolled in SDP in kindergarten at any point between September and December 2014, the period when KEI assessments were conducted for the 2014/15 school year. The characteristics used to describe the study samples and construct weights for the analysis are listed in table A1.

**Table A1. Student characteristics used in the study**

Characteristic	Description
Female	Whether student was female
Race/ethnicity	Whether student was <ul style="list-style-type: none"><li>• Black/African American</li><li>• Hispanic/Latino</li><li>• White</li><li>• Multiracial/other</li></ul>
Economically disadvantaged	Whether student participates in Supplemental Nutrition Assistance Program, Temporary Assistance for Needy Families, Medicaid, or other social service programs and is therefore eligible for free lunch without additional verification
English learner	Whether student was an English learner or had limited English proficiency
Special education	Whether student received special education services

Source: Authors’ analysis based on School District of Philadelphia 2014/15 data.

### Kindergarten Entry Inventory

The KEI is a teacher-administered assessment created by Pennsylvania’s Office of Child Development and Early Learning to understand and track students’ skills at kindergarten entry (Pennsylvania Office of Child Development and Early Learning, 2017). The full KEI includes 30 indicators across five areas: social and emotional development; English language arts; mathematics; approaches to learning; and health, wellness, and physical development. On each of these indicators, teachers give students a rating of 1 (Not yet evident), 2 (Emerging), 3 (Evident), or 4 (Exceeds) based on their observations of students in the classroom.

The study analyzed two KEI dimensions identified as having evidence of validity by a previous study (Howard et al., 2017). For each dimension, the score is the average rating (on a 1- to 4-point scale) across the indicators:

- Emerging Academic Competencies score is the average rating across 12 indicators of early literacy and early mathematics skills.
- Learning Engagement Competencies score is the average rating across nine indicators of behavioral and motor skills.

The KEI indicators included in each dimension and those excluded from the study because they were not part of a validated dimension are listed in table A2.

**Table A2. Kindergarten Entry Inventory indicators, by dimension**

Dimension	Included in this study	Indicators
Emerging Academic Competencies	Yes	<ul style="list-style-type: none"> <li>○ Counting</li> <li>○ Identifying shapes</li> <li>○ Naming numbers</li> <li>○ Operations and algebraic thinking</li> <li>○ Phonics</li> <li>○ Phonological awareness</li> <li>○ Print concepts/letters</li> <li>○ Print concepts/words</li> <li>○ Stages of writing</li> <li>○ Text analysis</li> <li>○ Text structure</li> <li>○ Writing process</li> </ul>
Learning Engagement Competencies	Yes	<ul style="list-style-type: none"> <li>○ Behavior regulation</li> <li>○ Conflict resolution</li> <li>○ Control and coordination: fine motor</li> <li>○ Control and coordination: gross motor</li> <li>○ Emotional regulation</li> <li>○ Engagement, attention, and persistence</li> <li>○ Reasoning and problem solving</li> <li>○ Self-awareness</li> <li>○ Task analysis</li> </ul>
Not part of a validated dimension	No	<ul style="list-style-type: none"> <li>○ Collaborative communication</li> <li>○ Conventions of English language</li> <li>○ Expressive language</li> <li>○ Receptive language</li> <li>○ Positional words</li> <li>○ Curiosity and initiative</li> <li>○ Data</li> <li>○ Measurement</li> <li>○ Stages of play</li> </ul>

Source: Adapted from Howard et al., 2017.

### ***Pennsylvania System of School Assessment***

The PSSA is Pennsylvania’s annual statewide standardized assessment, which covers both math and ELA in grades 3 through 8. This study focused on the grade 3 PSSA in ELA. PSSA scale scores ranged from 600 to 1551, and the performance levels can be classified as proficient (proficient or advanced) and not proficient (basic or below basic) (Pennsylvania Department of Education, 2018).

### ***AIMSweb***

AIMSweb is a widely used progress-monitoring system (Pearson, 2019). The system allows districts to select from a range of established assessments (or develop their own assessments) and track student performance against benchmarks. SDP used the grade-appropriate literacy-related AIMSweb assessments for each semester from spring 2015 through spring 2018 (table A3).

**Table A3. School District of Philadelphia’s use of AIMSweb spring reading assessments, by grade**

	Kindergarten (2014–15)	Grade 1 (2015–2016)	Grade 2 (2016–2017)	Grade 3 (2017–2018)
Letter naming fluency	X			
Nonsense word fluency		X	X	
Reading curriculum-based measurement of oral reading fluency				X

Source: School District of Philadelphia 2014–2018 data.

Table reads: School District of Philadelphia used the AIMSweb letter naming fluency assessment to assess students in kindergarten.

The AIMSweb data included two measures of student performance. Raw scores are the number of correct answers. Performance tiers describe students’ performance relative to the assessments’ benchmarks: at target (Tier 1), requiring strategic intervention (Tier 2), or requiring intensive intervention (Tier 3).

### Samples

The study sample was the cohort of kindergarteners who entered SDP by December of the 2014/15 school year (13,335 students). The samples used to answer the research questions are smaller because students needed to have data on the relevant assessments. As discussed in the report, 2014 was the first year that the entire district was required to use the KEI, and 58 to 68 percent of kindergarteners had KEI scores for one or both dimensions (table A4). District staff provided several explanations for missing KEI scores.<sup>3</sup> First, the 2014/15 school year was the first year that the entire district was required to use the KEI, so some teachers may not have complied with the requirement. Second, students must be rated on all individual indicators included in each KEI dimension to receive a score, and some students were missing data on particular indicators. Third, SDP experienced some issues with matching students based on the identifiers included in the data collected by the state (KEI and PSSA) and SDP (enrollment records and AIMSweb). By grade 3, about 40 percent of the full sample of kindergarteners and 65 percent of those with KEI scores had ELA scores on the grade 3 PSSA. Missing PSSA scores could be due to students moving out of the district, transferring to non-district schools (Philadelphia has a large charter sector), being absent on the assessment date, or repeating a grade. The sample of students with KEI and PSSA scores includes 43 percent of all students enrolled in an SDP school for grade 3 during the 2017/18 school year.

**Table A4. Two-thirds of kindergarteners had Kindergarten Entry Inventory scores and about 40 percent had both Kindergarten Entry Inventory and Pennsylvania System of School Assessment scores**

KEI dimension	With scores on the KEI dimension		With scores on the KEI dimension and PSSA scores		
	Number of students	Percentage of kindergarteners	Number of students	Percentage of kindergarteners	As percentage of students with scores on the KEI dimension
Emerging Academic Competencies	8,089	61	5,267	39	65
Learning Engagement Competencies	9,089	68	5,944	45	65
Both KEI dimensions	7,707	58	5,029	38	65

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: 13,335 kindergarteners entered School District of Philadelphia by December of the 2014/15 school year.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2017/18 data.

Table reads: 8,089 kindergarteners (61 percent of all kindergarteners) had scores on Emerging Academic Competencies dimension. 5,267 kindergarteners (39 percent of all kindergarteners and 65 percent of those with scores on the Emerging Academic Competencies dimension) had scores on both the Emerging Academic Competencies dimension and grade 3 PSSA.

<sup>3</sup> Personal communication on April 22, 2019.

The sample sizes for analyses used to answer each research question that are presented in the main text (table A5) differ for two key reasons:

- **Students missing data on one or more assessments.** Some students were missing data on one or both of the KEI dimensions, grade 3 PSSA in ELA, or spring AIMSweb assessments in grades K through 3.
- **Analyses that focus on a random 70 or 30 percent of the sample.** Following the typical approach for predictive analyses, the study team split the sample randomly into two parts, conducting analyses on the first part to select the threshold (70 percent of the full sample) and then checking the accuracy of the predictions using the second part (30 percent of the full sample). To ensure that the two parts would be similar in terms of students' PSSA proficiency, the data were stratified into three groups based on student's PSSA proficiency level (proficient, not proficient, and missing PSSA data) before students were randomly split into the two parts. Consistent with random sampling, students in the two parts did not differ on any characteristics by more than 5 percent of a standard deviation for all students (table A6). For comparison with findings from research question 1 (which establish the feasibility of setting a threshold in a subsample of data), research question 3 also used a 70 percent sample (so its findings can be used to inform the feasibility of setting a threshold based on different assessments in future work).

**Table A5. Sample sizes by research question and analysis sample for the analyses presented in the main text**

Research question (RQ)	Analysis sample	Number of students
RQ1. Relationship between KEI scores and grade 3 PSSA in ELA scores (in sample used to select threshold)	Random 70 percent sample of students who took the KEI in 2014/15 and the grade 3 PSSA in ELA in 2017/18	3,521
RQ2. Accuracy of threshold (in sample used to select it)	Random 70 percent sample of students who took the KEI in 2014/15 and the grade 3 PSSA in ELA in 2017/18	3,521
RQ2. Accuracy of threshold (in sample not used to select it)	Random 30 percent sample of students who took the KEI in 2014/15 and the grade 3 PSSA in ELA in 2017/18	1,508
RQ3. Relationship between KEI scores, spring AIMSweb Reading scores in grades K-3, and grade 3 PSSA in ELA	Random 70 percent sample of students who took both of the assessments being analyzed in each year (that is, AIMSweb and KEI or AIMSweb and PSSA)	3,689–5,446

ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: The number of students varied based on the KEI dimension (Emerging Academic Competencies, Learning Engagement Competencies, and both dimensions) and other assessments (AIMSweb Reading or grade 3 PSSA in ELA) included in the analyses.

Source: Authors' analysis based on School District of Philadelphia 2014–2018 data.

**Table A6. Students whose data were used to set the threshold and those whose data were used to assess the threshold’s accuracy had similar characteristics**

Characteristic	Students whose data were used to set the threshold (mean <sup>a</sup> or percentage <sup>b</sup> )	Students whose data were used to assess the threshold’s accuracy (mean <sup>a</sup> or percentage <sup>b</sup> )	Standard deviation for all students	Difference (in standard deviation units)
Female	49	48	50	0.020
Black/African American	49	51	50	-0.040
Hispanic/Latino	21	20	40	0.025
White	14	14	35	0.000
Multiracial or other race/ethnicity	16	15	36	0.028
Economically disadvantaged program participant	67	66	47	0.021
English learner student	10	10	30	0.000
Special education student	8	8	27	0.000
Emerging Academic Competencies score	2.41	2.39	0.66	0.030
Learning Engagement Competencies score	2.50	2.51	0.54	-0.019
Kindergarten spring AIMSweb score	47.8	47.5	20.7	0.014
Grade 1 spring AIMSweb score	57.1	56.4	34.3	0.020
Grade 2 spring AIMSweb score	84.4	84.8	43.7	-0.009
Grade 3 spring AIMSweb score	104.6	103.9	49	0.014
Grade 3 PSSA score	973	974	99	-0.010
Grade 3 PSSA proficiency	34	34	48	0.000
<b>Number of students</b>	<b>5,435–9,317</b>	<b>2,322–4,018</b>		

aThe values for Emerging Academic Competencies scores, Learning Engagement Competencies scores, AIMSweb scores, PSSA scores are mean values.

bThe values for female, Black/African American, Hispanic/Latino, White, Multiracial or other race/ethnicity, economically disadvantaged program participant, English language student, special education student, grade 3 PSSA proficiency are percentages.

PSSA is Pennsylvania System of School Assessment.

Note: Descriptive statistics were not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014–2018 data described in appendix A.

## Weights

This section describes the weights used in the analyses to account for missing data and the differences between the characteristics of students included in the analysis and those in the population of interest. It also describes how the study assessed differences between the analytic sample and the population of interest.

### Primary weights

SDP intends to apply the threshold to all entering kindergarteners with KEI scores to estimate the proportion of students on-track to achieve reading proficiency in grade 3. However, not all of these students will take the grade 3 PSSA in ELA, even though the analyses that set the threshold are based on students with both KEI and grade 3 PSSA scores. Thus, the analyses for research questions 1 and 2 used weights to make the findings based on students who have KEI and grade 3 PSSA scores representative for all kindergarteners with KEI scores.

The analyses for research questions 1 and 2 used trimmed inverse probability weights, which were constructed following the National Assessment of Education Progress (2017) procedure. First, a logistic regression model estimated each student's probability of taking the grade 3 PSSA in ELA. This model included students' KEI scores on the two domains; students' characteristics (gender, race/ethnicity, economic disadvantage, English learner, and special education); and indicators for missing KEI scores or student characteristics. It also included interactions between race/ethnicity indicators and KEI scores (and indicators for missing KEI scores) because multiple unweighted analytic samples had differences on these characteristics.

Second, the initial weights were set equal to the inverse of students' predicted probability of taking the PSSA from the logistic regression. Therefore, students who represent the groups that were least likely to have PSSA data received greater weight in the analysis.

Finally, the initial weights were "trimmed" to ensure that no single student would have too much influence over final estimates. The study team used a maximum weight of 3.5 times the median weight, following the National Assessment of Education Progress (2017) guidelines. Students with calculated weights from the logistic model that were larger than 3.5 were given a weight of 3.5.

For research question 3, the analysis was exploratory and not intended to set an indicator that could be applied to a representative population of students (such as all students with AIMSweb scores in a particular grade). Thus, to facilitate comparisons across the analyses in research question 3, these analyses did not use weights.

#### *Alternative weights*

To test the sensitivity of the findings to the weighting strategy, the analyses to address research question 2 were repeated using two types of alternative weighting approaches: one that weighted the sample to represent all entering kindergarteners in SDP and one that was not weighted.

SDP would like to understand the proportion of all entering kindergarteners who are on-track to achieve reading proficiency in grade 3, but 30 to 40 percent of kindergarteners were missing KEI scores, a key predictor of grade 3 PSSA scores in ELA. Therefore, as a sensitivity test, another weight was constructed to calculate a predicted probability of proficiency that was reflective of all kindergarteners. This weight used a two step-approach, which incorporated the available KEI data. First, a logistic regression model was used to estimate each student's probability of having a KEI score in fall 2014. This model included students' demographic characteristics and indicators for missing demographic characteristics. Second, the predicted probability of having a KEI score was multiplied by the predicted probability of having a grade 3 PSSA in ELA score, conditional on KEI scores and other characteristics (as calculated in the logistic regression model described above). Next, as above, the initial weight was set to the inverse of the product of the two predicted probabilities and trimmed to a maximum of 3.5 times the median weight.

#### *Assessing differences between samples*

This section presents analyses that assess the extent to which weighted and unweighted analytic samples were similar to other samples and populations of interest, based on the characteristics available in the SDP data. It also provides analyses that examine how differences in characteristics relate to the primary outcome of interest—students' grade 3 PSSA in ELA scores. These analyses provide context about the extent to which the findings for a given analytic sample might generalize to other groups of students. Analytic samples were defined in terms of whether students (1) were enrolled in SDP in kindergarten in the fall of 2014, (2) had KEI scores, and (3) had scores on the grade 3 PSSA in ELA. Samples were considered similar on a given characteristic if, on average, differences between groups did not exceed 5 percent of a standard deviation for all students. However, differences above 5 percent of a standard deviation could still represent small percentage point differences.



*Differences between unweighted samples of all kindergarteners and students with KEI scores.* For the indicator to accurately measure the percentage of a cohort of entering kindergarteners who are on track for reading proficiently, ideally, students with scores on the KEI must not be substantially different from those without scores. Without weighting, the students enrolled in kindergarten by December 2014 and the subsamples of students with KEI and KEI and PSSA in ELA data were similar on most characteristics (table A7). However, there were a few differences that exceeded 5 percent of a standard deviation. For example, students with both KEI dimension scores were more likely to be White than all kindergarteners. In addition, students with both KEI and PSSA in ELA scores were less likely to be Black or receiving special education services than all kindergarteners. The analyses only examined the characteristics available in the data for all kindergarteners, so they do not provide information about differences in students' achievement.

**Table A7. Among students whose data was used to set the threshold, those with Kindergarten Entry Inventory data and those with Kindergarten Entry Inventory and Pennsylvania System of School Assessment data differed from all kindergarteners on some demographic characteristics**

Characteristic	All kindergarteners		Subsample with both KEI dimension scores		Subsample with PSSA and both KEI dimension scores	
	Percentage	Standard deviation (SD)	Percentage	Difference from all kindergarteners (in SD units)	Percentage	Difference from all kindergarteners (in SD units)
Female	48	50	48	0.011	49	0.031
Black/African American	49	50	48	-0.035	46	-0.072†
Hispanic/Latino	21	41	20	-0.023	22	0.036
White	14	35	16	0.053†	15	0.026
Multiracial or other race/ethnicity	16	36	17	0.023	17	0.035
Economically disadvantaged program participant	66	47	68	0.044	68	0.039
English language learner	10	30	10	0.013	11	0.045
Special education student	8	27	8	-0.015	5	-0.105†
<b>Number of students</b>	<b>9,317</b>		<b>5,398</b>		<b>3,521</b>	

† Larger than 5 percent of a standard deviation.

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: Descriptive statistics were not weighted.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*Differences between the characteristics of unweighted samples of students with KEI scores and students with both KEI and grade 3 PSSA ELA scores.* The analytic sample of students with KEI data had several characteristics that differed by more than 5 percent of a standard deviation from those of students with both KEI and grade 3 PSSA in ELA scores (table A8). For example, compared with students with KEI scores, students with both KEI and PSSA scores had higher KEI scores on both dimensions, were more likely to be Hispanic/Latino, and were less likely to receive special education services.

**Table A8. Among students whose data were used to set the threshold, Kindergarten Entry Inventory scores and other student characteristics differed for students with Kindergarten Entry Inventory and grade 3 Pennsylvania System of School Assessment in English Language Arts scores and students with just Kindergarten Entry Inventory scores**

Characteristic	Kindergarteners with both KEI dimension scores		Subsample with PSSA and both KEI dimension scores	
	Mean <sup>a</sup> or percentage <sup>b</sup>	Standard deviation (SD)	Mean <sup>a</sup> or percentage <sup>b</sup>	Difference from kindergarteners with KEI scores (in SD units)
Female	48	50	49	0.020
Black/African American	48	50	46	-0.038
Hispanic/Latino	20	40	22	0.060 <sup>†</sup>
White	16	37	15	-0.026
Multi-racial or other race/ethnicity	17	37	17	0.012
Economically disadvantaged program participant	68	47	68	-0.006
English learner student	10	30	11	0.031
Special education student	8	26	5	-0.093 <sup>†</sup>
Emerging Academic Competencies score	2.395	66	2.440	0.068 <sup>†</sup>
Learning Engagement Competencies score	2.515	54	2.552	0.069 <sup>†</sup>
<b>Number of students</b>	<b>5,398</b>		<b>3,521</b>	

<sup>a</sup>The values for Emerging Academic Competencies scores and Learning Engagement Competencies scores are mean values.

<sup>b</sup>The values for female, Black/African American, Hispanic/Latino, White, Multiracial or other race/ethnicity, economically disadvantaged program participant, English learner student and special education student are percentages.

<sup>†</sup> Larger than 5 percent of a standard deviation.

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: Descriptive statistics were not weighted.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*Relationships between student characteristics and grade 3 PSSA in ELA scores.* The threshold was set using data on students with both KEI and grade 3 PSSA in ELA scores, but it will be ultimately applied to students with KEI scores (some of whom will not have PSSA scores in grade 3). Findings for students with both KEI and grade 3 PSSA scores might not generalize to students with KEI scores if there are substantial differences between the two groups of students on characteristics that are correlated with grade 3 PSSA scores that are not adjusted using weights.

KEI scores had weak to moderate correlations with grade 3 PSSA scores (table A9). Other student characteristics (such as economic disadvantage) had weak but statistically significant correlations with grade 3 PSSA scores. Therefore, the analyses used weights to make the sample of students with both KEI and grade 3 PSSA scores representative of students with KEI scores.

**Table A9. Among students whose data were used to set the threshold, Kindergarten Entry Inventory scores had weak to moderate correlations with grade 3 Pennsylvania System of School Assessment scores, while other student characteristics had weak correlations with grade 3 Pennsylvania System of School Assessment scores**

Characteristic	Correlation with grade 3 PSSA scores
Female	0.11***
Black/African American	-0.21***
Hispanic/Latino	-0.11***
White	0.24***
Multiracial or other race/ethnicity	0.20***
Economically disadvantaged program participant	-0.20***
English learner student	-0.04**
Special education student	-0.12***
Emerging Academic Competencies score	0.42***
Learning Engagement Competencies score	0.33***
<b>Number of students</b>	<b>3,729–5,311</b>

\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

PSSA is Pennsylvania System of School Assessment.

Note: Correlations were not weighted.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*Differences between the characteristics of the weighted samples of students with KEI scores and students with both KEI and grade 3 PSSA in ELA scores.* Weighting reduced the differences between students with both KEI and grade 3 PSSA in ELA scores and students with KEI scores (table A10). Between the weighted analytic samples, no characteristics had a difference greater than 5 percent of a standard deviation.

**Table A10. Among students whose data were used to set the threshold, the weights used in the main analyses make the students with Kindergarten Entry Inventory and Pennsylvania System of School Assessment scores similar to students with just Kindergarten Entry Inventory scores**

Characteristic	Kindergarteners with both KEI dimension scores		Weighted subsample with PSSA and both KEI dimension scores	
	Mean <sup>a</sup> or percentage <sup>b</sup>	Standard deviation (SD)	Mean <sup>a</sup> or percentage <sup>b</sup>	Difference from students with both KEI scores (in SD units)
Female	48	0.500	48	0.001
Black/African American	48	0.499	48	0.002
Hispanic/Latino	20	0.399	20	0.001
White	16	0.367	16	0.000
Multiracial or other race/ ethnicity	17	0.371	16	-0.003
Economically disadvantaged program participant	68	0.467	68	-0.001
English learner student	10	0.304	10	0.001
Special education student	8	0.264	8	-0.002
Emerging Academics Competencies score	2.395	0.659	2.398	0.005
Learning Engagement Competencies score	2.515	0.536	2.519	0.008
<b>Number of students</b>	<b>5,398</b>		<b>3,521</b>	

<sup>a</sup>The values for Emerging Academic Competencies scores and Learning Engagement Competencies scores are mean values.

<sup>b</sup>The values for female, Black/African American, Hispanic/Latino, White, Multiracial or other race/ethnicity, economically disadvantaged program participant, English learner student and special education student are percentages.

KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: None of the differences were larger than 5 percent of a standard deviation.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

## Methods

This section describes the methods used to answer each of the study's research questions.

### *Research question 1: What is the relationship between KEI scores and scores on the grade 3 PSSA in ELA?*

Descriptive statistics and pairwise correlations were used to assess the relationship between each KEI dimension score and students' PSSA scaled score among students whose data were used to set the threshold. Positive correlations were classified as follows: weak (0.1–0.39), moderate (0.4–0.69), strong (0.7–0.99), and perfect (1; Dancy & Reidy, 2007). Statistical significance was evaluated at and below the 0.05 level.

Logistic regressions were used to examine the relationship between the KEI dimensions and grade 3 PSSA in ELA proficiency (equation A1). These regressions analyzed KEI dimensions together (as in equation A1) or separately (by excluding the other dimension). Logistic regressions are designed to analyze binary (yes/no) outcomes like proficiency, and their results can be translated into readily interpretable findings (odds ratios or average marginal effects). For example, for Emerging Academic Competencies, the coefficient ( $\beta_{EAC}$ ) in the logistic equation is the log odds, and the odds ratio is  $e^{\beta_{EAC}}$ . The coefficients and odds ratios represent the average relationship between ELA proficiency and KEI scores. Average marginal effects, which can be interpreted as the average increase in the probability of proficiency for every 1-point increase, were calculated for the report. McFadden's Pseudo R Squared is a measure of how well the model fits the data compared to a model with no predictors. SDP wanted to set a single threshold for all groups of students so the regressions did not include student characteristics as covariates.

**Equation A1. Logistic regression model for relationship between Kindergarten Entry Inventory dimensions and grade 3 Pennsylvania System of School Assessment proficiency**

$$\Pr(Y_i = 1) = \frac{e^{\beta_0 + \beta_{EAC}EAC_i + \beta_{LEC}LEC_i}}{1 + e^{\beta_0 + \beta_{EAC}EAC_i + \beta_{LEC}LEC_i}}$$

$Y_i$  = Student is proficient on the grade 3 PSSA in ELA

$EAC_i$  = Student’s Emerging Academic Competencies score

$LEC_i$  = Student’s Learning Engagement Competencies score

*Research question 2: What threshold scores on the KEI most accurately predict reading proficiency on the grade 3 PSSA in ELA for the cohort? What threshold score on the KEI has a correct prediction rate of at least 90 percent for the individual students who were not proficient in ELA on the grade 3 PSSA?*

Answering research question 2 involved creating an index based on the two KEI dimensions, setting the threshold based on a subset of data, and assessing the accuracy of the threshold in data that were not used to set the threshold.

**Creating an index based on the two KEI dimensions.** The index of the two KEI dimensions (equation A2) was a weighted average based on the coefficients from the logistic regression shown in equation A1. This index is scaled to range from 1 to 10, such that if students performed perfectly (were rated as “Exceeds” on all indicators on both dimensions) they would receive a 10. This scaling was used because SDP wanted to prevent the index from being confused with a raw KEI score (range: 1 to 4) or a percentile rank (range: 1 to 100).

**Equation A2. Formula for calculating a Kindergarten Entry Inventory index score**

$$Y_i = 1 + 9 \frac{(\widehat{\beta}_{EAC}EAC_i + \widehat{\beta}_{LEC}LEC_i - \widehat{Min})}{\widehat{Max} - \widehat{Min}} = 1 + \frac{9(1.147EAC_i + 0.108LEC_i - 1.255)}{5.020 - 1.255} = \frac{1 + 9(1.147EAC_i + 108LEC_i - 1.255)}{3.765}$$

$Y_i$  = Student’s Kindergarten Entry Inventory index score

$EAC_i$  = Student’s Emerging Academic Competencies score

$LEC_i$  = Student’s Learning Engagement Competencies score

$\widehat{\beta}_{EAC}$  = Estimated coefficient on Emerging Academic Competencies score from equation A1 = 1.147

$\widehat{\beta}_{LEC}$  = Estimated coefficient on Learning Engagement Competencies scores from equation A1 = 0.108

$\widehat{Min}$  = Minimum estimated value of  $\widehat{\beta}_{EAC}EAC_i + \widehat{\beta}_{LEC}LEC_i$ , which corresponds to  $EAC_i = 1$  and  $LEC_i = 1$ .

$$\widehat{Min} = 1.147 + 0.108 = 1.255$$

$\widehat{Max}$  = Maximum estimated value of  $\widehat{\beta}_{EAC}EAC_i + \widehat{\beta}_{LEC}LEC_i$ , which corresponds to  $EAC_i = 4$  and  $LEC_i = 4$ .

$$\widehat{Max} = (1.147 * 4 + 0.108 * 4) = 5.020$$

**Setting the threshold based on a subset of data.** Several statistics were calculated to assess the accuracy of potential KEI thresholds and select a threshold (Box A1). In the sample used to select the threshold, the percentage of students correctly predicted to be proficient and the percentage of students correctly predicted to be not proficient were calculated for all possible thresholds on (1) Emerging Academic Competencies, (2) Learning Engagement Competencies, and (3) the index of Emerging Academic Competencies and Learning Engagement

Competencies. Then, the percentage of correct predictions for students who were and were not proficient based on these thresholds across the different potential indicators were compared to decide which dimensions to use for setting the threshold (shown in figure 2 in the report).

**Box A1. Statistics used to assess the accuracy of potential Kindergarten Entry Inventory thresholds**

Absolute difference between the actual proficiency rate and predicted proficiency rate based on the threshold = |Percentage of proficient students - Percentage of students above threshold|

Percentage of proficient students whom the threshold correctly predicted = 
$$\frac{\text{Number of proficient students above threshold}}{\text{Total number of proficient students}}$$

Percentage of not proficient students whom the threshold correctly predicted = 
$$\frac{\text{Number of not proficient students below threshold}}{\text{Total number of not proficient students}}$$

Percentage of students whom the threshold correctly predicted = 
$$(\text{Percentage of proficient students} \times \text{Percentage of proficient students whom the threshold correctly predicted} + \text{Percentage of not proficient students} \times \text{Percentage of not proficient students whom the threshold correctly predicted}) / 100$$

Emerging Academic Competencies and the index of the two dimensions had similar percentages of correct predictions for proficient students and students who were not proficient, and these percentages were higher than those for Learning Engagement Competencies. Thus, the study team asked SDP to choose between using Emerging Academic Competencies and the index of the two dimensions, and SDP decided to use the index of the two dimensions.

Because the indicator will be used to estimate the overall proportion of students who are on-track to achieve grade 3 reading proficiency, the study then calculated the threshold on the index of the two dimensions that came closest to matching the actual grade 3 reading proficiency rate. That is, the selected threshold minimized the absolute difference between actual and predicted proficiency.

To select the threshold with the smallest difference between actual and predicted proficiency, students were ranked by their KEI index scores. The study team then attempted to identify the KEI index score at exactly the 63rd percentile of the distribution because the percentage of students with KEI index scores above this value (37 percent) would match the percentage of students in the data who were actually proficient on the grade 3 PSSA in ELA. In practice, because students cluster at discrete KEI index scores, the study identified two KEI index scores—one directly above and one directly below the 63rd percentile of the distribution. The study team then selected the number with the smallest difference between actual and predicted proficiency (5.97) and rounded it up to 6 to make it easier to communicate. The difference between actual and predicted proficiency for a selection of possible thresholds (between 1 and 10) on the KEI index score are shown in table B4. SDP can use the equation to create the index (equation A2) and the threshold of 6 to calculate the percentages of future cohorts of students with KEI scores who are on-track to be proficient in reading in grade 3.

Different measures of accuracy (such as correctly predicting proficient students or students who were not proficient) could be given priority depending on the intended use of the threshold. Most predictive analyses aim to achieve acceptable levels of predicting proficient students or students who were not proficient, depending on the analyses’ goals. For example, for referring students for additional risk assessment, there may be a large benefit to identifying students at risk of not being proficient and a low cost to misclassifying some students as being at risk. This application might target more accurate identification of students who were not proficient than an

application that aims to provide resource-intensive supports to at risk students. Therefore, the same methodology described above was used to identify a cut point where at least 90 percent of students who were not proficient were correctly identified. Ninety percent was selected as a threshold to match existing research (see Compton, Fuchs, Fuchs, & Bryant, 2010; Jenkins, Hudson, & Johnson, 2007).

**Assessing the accuracy of the threshold using data that were not used to set the threshold.** All four measures of accuracy described in box 1 were calculated using the cohort threshold of 6 for the 30 percent sample that was not used to set the threshold. These measures were compared to their values in the sample used to set the threshold.

*Research question 3: How do AIMSweb Reading scores in the spring of kindergarten through grade 3 relate to children's scores on the KEI and grade 3 PSSA in ELA?*

Descriptive statistics and pairwise correlations were used to explore the relationships between spring AIMSweb scores from kindergarten through grade 3 and KEI and PSSA ELA scores, following the approach used in research question 1. Measures of accuracy (correctly predicting proficient students or students who were not proficient on the grade 3 PSSA in ELA) were calculated for the developer-selected AIMSweb thresholds, using the equations shown in box A1.

## Appendix B. Supporting analyses

This appendix presents supporting analyses for each research question.

### Relationships between KEI scores and grade 3 PSSA scores

For research question 1, students with higher KEI scores on the Emerging Academic Competencies and Learning Engagement Competencies dimensions had higher grade 3 PSSA in ELA scores (figure 1). The numeric values for this figure are shown in table B1.

**Table B1. Students with higher Kindergarten Entry Inventory scores were more likely to be proficient on the grade 3 Pennsylvania System of School Assessment in English Language Arts**

KEI score	Percentage of students who were proficient on	
	Emerging Academic Competencies	Learning Engagement Competencies
1—1.5	13	14
1.6—2	22	22
2.1—2.5	35	30
2.6—3	43	45
3.1—3.5	63	57
3.6—4	79	63

Note: The table is based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA.

Source: Authors' analysis based on School District of Philadelphia data 2014/15 and 2017/18 data described in appendix A.

### Regression analyses using the sample with one KEI dimension

For the logistic regression analyses predicting PSSA proficiency (table 2), findings were restricted to students with both KEI dimensions. Findings were similar when all students with the relevant KEI dimension were included in the analyses (table B2).

**Table B2. Emerging Academic Competencies were more predictive of students' probability of being proficient on the grade 3 Pennsylvania System of School Assessment than were Learning Engagement Competencies**

KEI dimensions	Relationship between KEI scores and the probability of being proficient on the grade 3 PSSA in ELA		
	Model 1: Emerging Academic Competencies	Model 2: Learning Engagement Competencies	Model 3: Both dimensions
Emerging Academic Competencies	0.25***		0.24***
Learning Engagement Competencies		0.25***	0.02
<b>Pseudo R-squared</b>	<b>0.09</b>	<b>0.06</b>	<b>0.09</b>
<b>Number of students</b>	<b>3,697</b>	<b>4,169</b>	<b>3,521</b>

ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analysis based on the random sample of students with KEI dimension scores and grade 3 PSSA scores that was used to set the threshold. Analyses were weighted to represent all students with the KEI dimensions scores included in the analysis. Reported values are average marginal effects, or the proportional increase in the average probability of PSSA proficiency associated with a 1-point increase in a KEI score. That is, in model 1, for every 1-point increase in Emerging Academic Competencies, the probability of proficiency increases on average 25 percent. Pseudo R-squared is a measure of how well the model fits the data compared to a model with no predictors.

Source: Authors' analysis based on SDP 2014/15 and 2017/18 data described in appendix A.



### Alternative weights

For research question 2, analyses of the accuracy of the selected threshold were repeated using alternative weighting approaches: no weights and weights based on student characteristics to represent all kindergarteners. The similarity of the findings indicate that the study’s results are not overly sensitive to the weights used (table B3).

**Table B3. The threshold of 6 had similar rates of accuracy in samples that were and were not used to select the threshold using alternative weights**

Sample	Actual proficiency (percentage)	Predicted proficiency (percentage)	Absolute difference between actual and predicted proficiency (percentage points)	Correct predictions for			Number of students
				Proficient students (percentage)	Not proficient students (percentage)	All students (percentage)	
<b>Primary analysis, weighted to represent all students with KEI scores</b>							
Used to set the threshold	37	37	1	53	73	66	3,521
Not used to set the threshold	37	37	0	58	76	69	1,508
<b>Alternative analysis, weighted to represent all kindergarteners</b>							
Used to set the threshold	37	36	0	53	73	66	3,521
Not used to set the threshold	36	37	1	58	76	69	1,508
<b>Alternative analysis, not weighted</b>							
Used to set the threshold	38	39	1	55	71	65	3,521
Not used to set the threshold	34	37	3	61	74	69	1,508

KEI is Kindergarten Entry Inventory.

Note: The primary analysis weighted the sample based on KEI scores and student characteristics. The alternative weighted analysis weighted the sample based on student characteristics.

The difference between actual and predicted proficiency may not equal the difference between the actual and predicted proficiency columns in the table due to rounding.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

### Alternative thresholds on the index of both Kindergarten Entry Inventory dimensions

This section presents the accuracy of alternative thresholds on the index of both KEI dimensions (table B4). The selected threshold for predicting the proportion of students on-track for reading proficiently by grade 3 minimizes the difference between actual and predicted proficiency. Another threshold that accurately predicted more of the students who were not proficient might be more appropriate for predicting which students need additional supports. For example, the threshold of 7 correctly predicted a higher percentage of students who were not proficient than the threshold of 6 (table B4 and figure B1).

**Table B4. Alternative thresholds on the index of both Kindergarten Entry Inventory dimensions had different rates of accuracy**

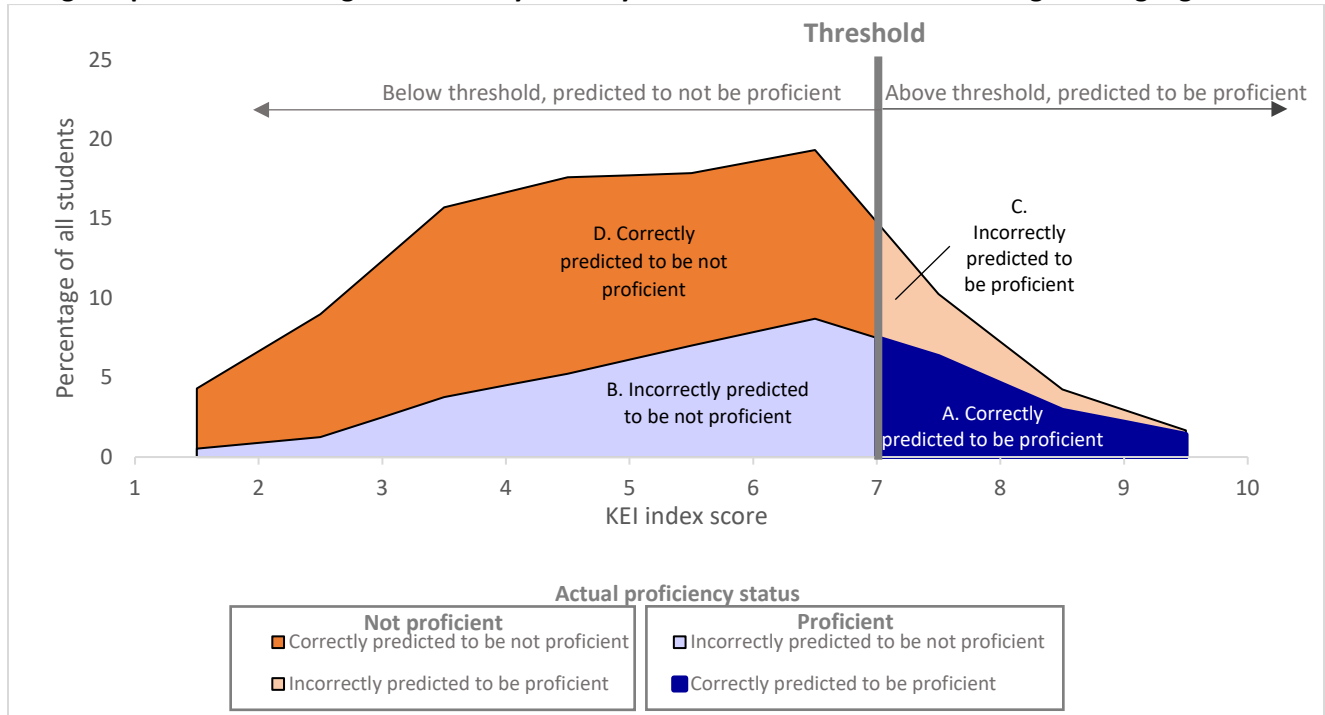
Threshold on the index using both KEI dimensions	Predicted proficiency (percentage)	Absolute difference between actual and predicted proficiency (percentage points)	Correct predictions for		
			Proficient students (percentage)	Not proficient students (percentage)	All students (percentage)
1.00	100	63	100	0	37
1.25	99	62	100	2	38
1.50	98	61	99	3	39
1.75	99	60	99	4	39
2.00	96	59	99	6	40
2.25	94	57	98	8	42
2.50	92	55	97	11	43
2.75	89	52	96	15	45
3.00	87	50	95	18	47
3.25	84	47	94	22	49
3.50	81	44	92	25	50
3.75	78	41	90	29	52
4.00	72	34	85	34	53
4.25	67	30	84	42	57
4.50	63	26	80	47	59
4.75	58	21	76	52	61
5.00	54	17	71	57	62
5.25	50	13	67	61	63
5.50	46	8	62	64	64
5.75	40	3	57	70	65
<b>6.00</b>	<b>37</b>	<b>1</b>	<b>53</b>	<b>73</b>	<b>66</b>
6.25	32	5	49	78	67
6.50	28	9	44	81	68
6.75	23	14	39	86	68
7.00	17	21	29	91	68
7.25	14	24	25	93	68
7.50	11	27	21	95	67
7.75	8	29	16	97	67
8.00	6	31	12	97	66
8.25	5	32	10	98	65
8.50	4	34	7	99	65
8.75	3	35	5	99	64
9.00	2	35	4	100	64
9.25	1	36	3	100	64
9.50	1	36	3	100	64
9.75	1	37	2	100	63
10.00	0	37	0	100	63

KEI is Kindergarten Entry Inventory.

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA. The threshold selected for estimating the proportion of students on track to achieve reading proficiency is highlighted. The difference between actual and predicted proficiency may not equal the difference between the actual and predicted proficiency columns in the table due to rounding.

Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

**Figure B1. A higher threshold of 7 increased the percentage of students correctly predicted to be at risk of being not proficient on the grade 3 Pennsylvania System of School Assessment in English Language Arts**



ELA is English language arts. KEI is Kindergarten Entry Inventory. PSSA is Pennsylvania System of School Assessment.

Note: Analysis based on the random sample of students with both KEI and grade 3 PSSA scores that was used to set the threshold and was weighted to represent all students with KEI scores. 3,521 students had information on both KEI competencies scores and grade 3 PSSA in ELA. Data are aggregated to smooth out the distribution, so no data are displayed at exactly 1 or 10.

Figure reads: At this threshold, students who were actually proficient (areas A and B) were either correctly predicted to be proficient (area A) or incorrectly predicted to be not proficient (area B). Students who were actually not proficient (areas D and C) were either correctly predicted to be not proficient (area D) or incorrectly predicted to be proficient (area C).

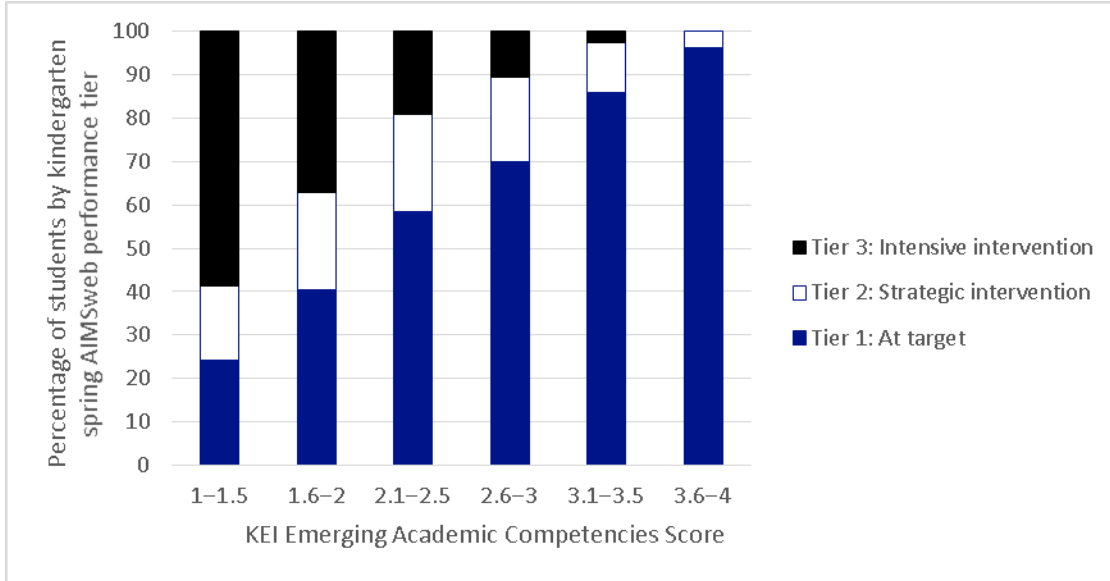
Source: Authors' analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

## Exploratory analyses of AIMSweb scores

The final section presents additional information on the exploratory research question about the relationship between children's reading skills across time (measured by the AIMSweb), their KEI scores, and their scores on the grade 3 PSSA in ELA. This study also presents sensitivity analyses that use a consistent sample of students who had scores on the KEI index, spring AIMSweb in each year from K through grade 3, and PSSA.

*Students with higher Emerging Academic Competencies and Learning Engagement Competencies scores were more likely to be identified as on target (Tier 1) on the AIMSweb.* Consistent with the patterns from the correlations presented in the report, students with higher KEI scores were more likely to be classified in AIMSweb Tier 1, in which students meet the grade-level target. For example, over 95 percent of those with the highest Emerging Academic Competencies scores (3.6-4) were in AIMSweb Tier 1 by the spring of kindergarten, compared with less than 25 percent of students with the lowest scores on Emerging Academic Competencies (1-1.5) (figure B2). By grade 3, about 80 percent of students in the highest range (3.6-4) were classified in Tier 1 (figure B5). Students with higher Learning Engagement Competencies scores were also more likely to be identified as on target (Tier 1) on the AIMSweb from kindergarten through grade 3 (figures B6 through B9).

**Figure B2. Students with higher Emerging Academic Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of kindergarten**

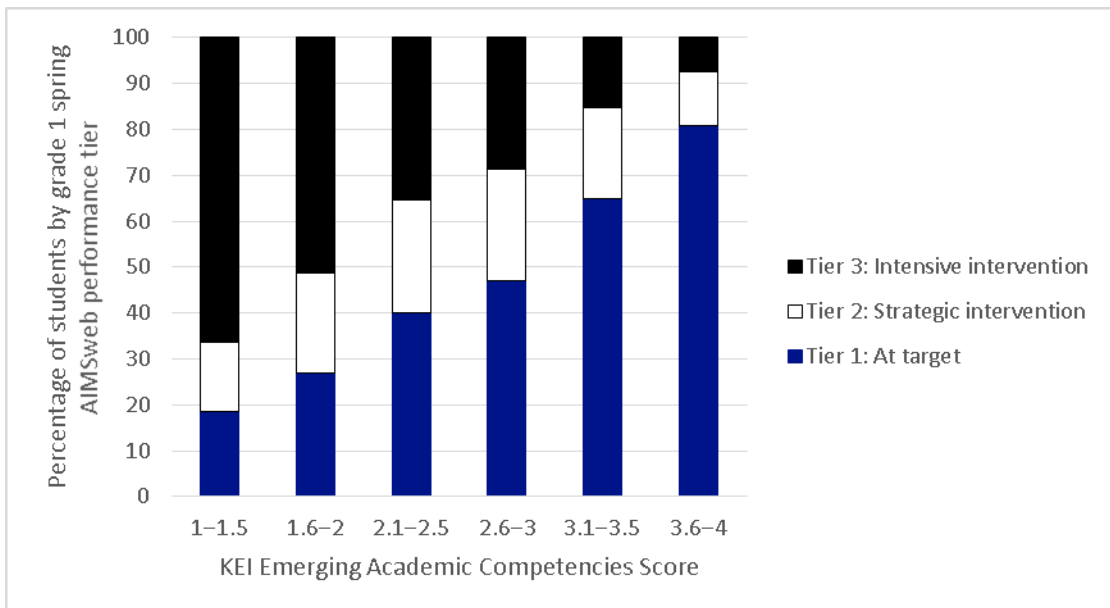


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 4,930 students that had information on KEI Emerging Academic Competencies and kindergarten spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 data described in appendix A.

**Figure B3. Students with higher Emerging Academic Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of grade 1**

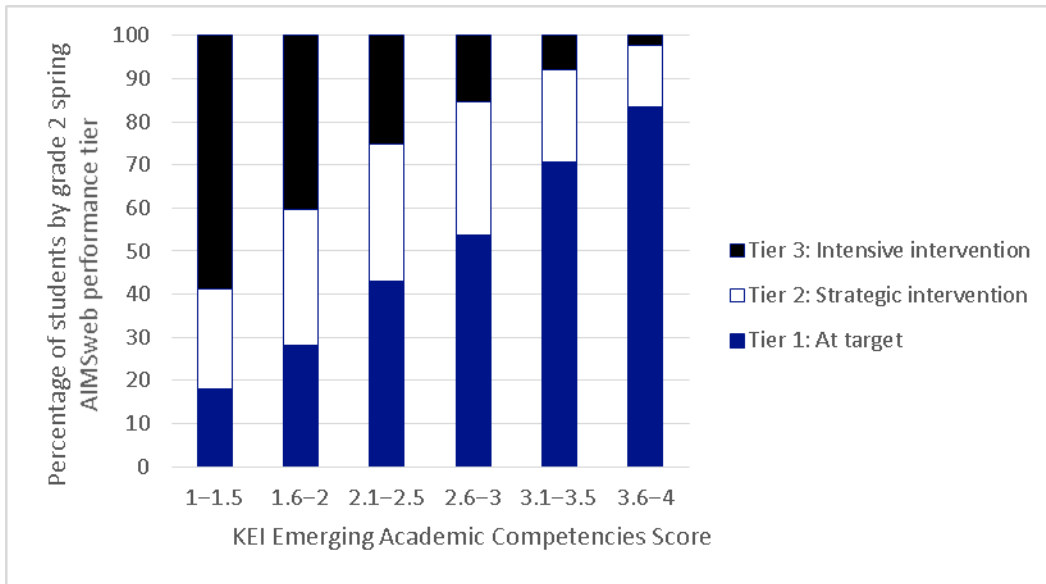


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 3,689 students that had information on KEI Emerging Academic Competencies and grade 1 spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2015/16 data described in appendix A.

**Figure B4. Students with higher Emerging Academic Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of grade 2**

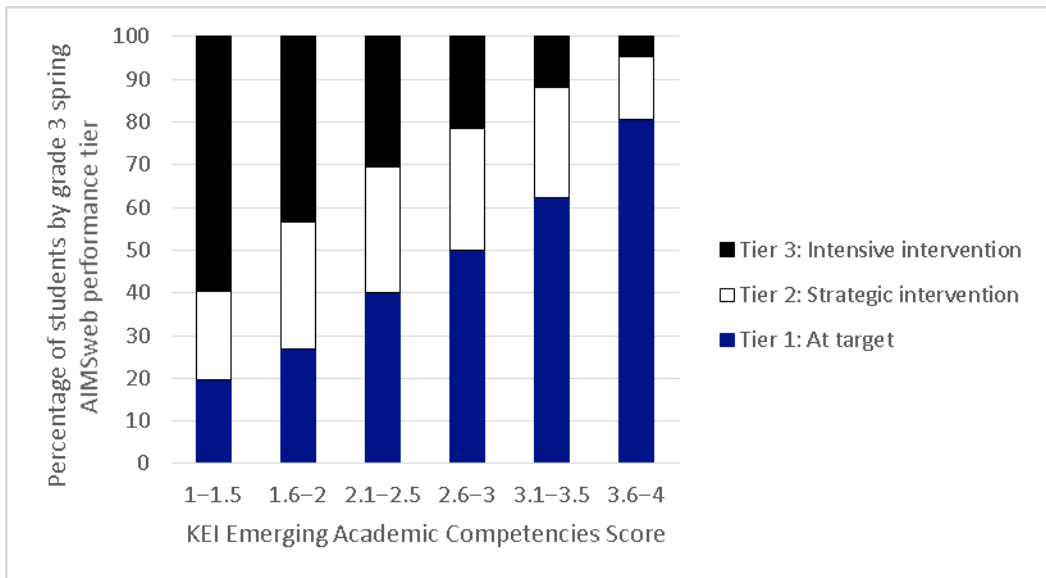


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 3,948 students that had information on KEI Emerging Academic Competencies and grade 2 spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2016/17 data described in appendix A.

**Figure B5. Students with higher Emerging Academic Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of grade 3**

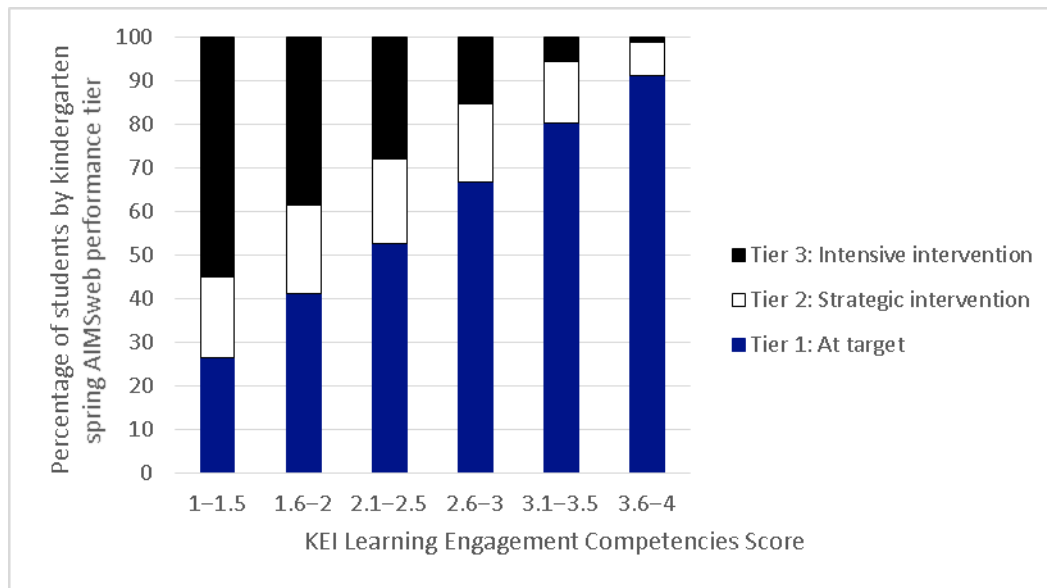


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 3,707 students that had information on KEI Emerging Academic Competencies and grade 3 spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

**Figure B6. Students with higher Learning Engagement Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of kindergarten**

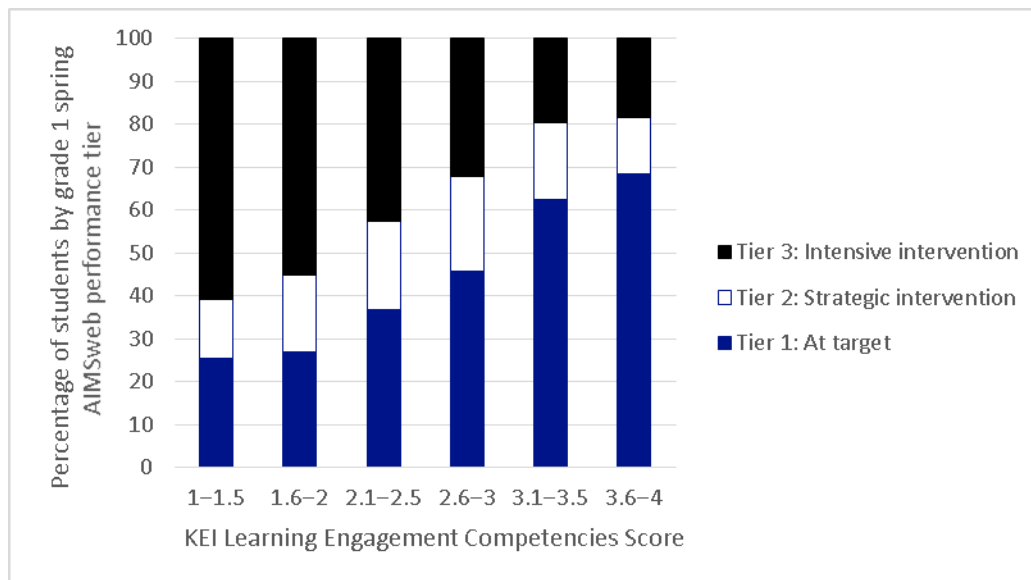


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 5,446 students that had information on KEI Learning Engagement Competencies and kindergarten spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 data described in appendix A.

**Figure B7. Students with higher Learning Engagement Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of grade 1**

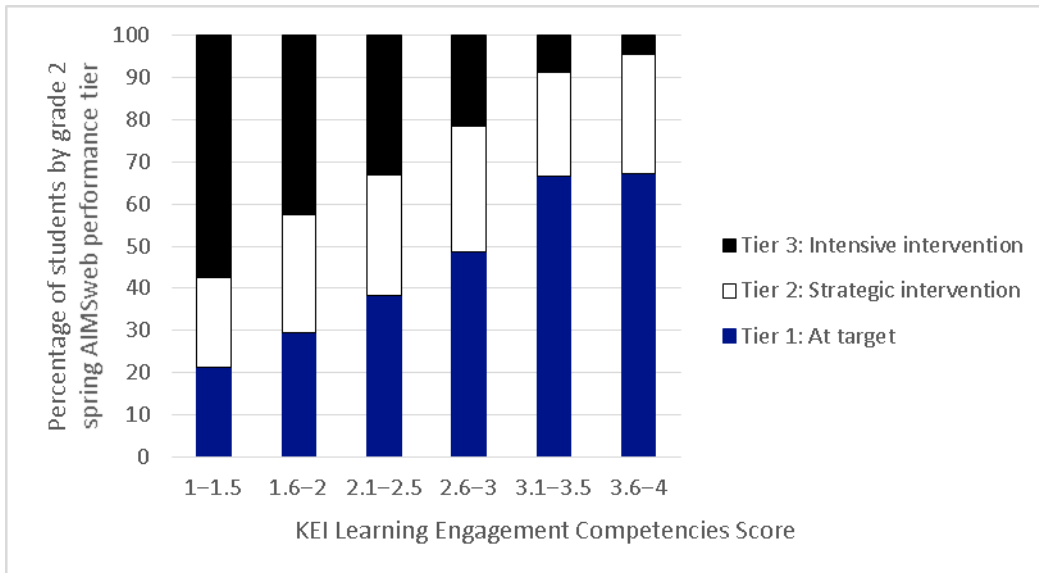


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 4,200 students that had information on KEI Learning Engagement Competencies and grade 1 spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2015/16 data described in appendix A.

**Figure B8. Students with higher Learning Engagement Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of grade 2**

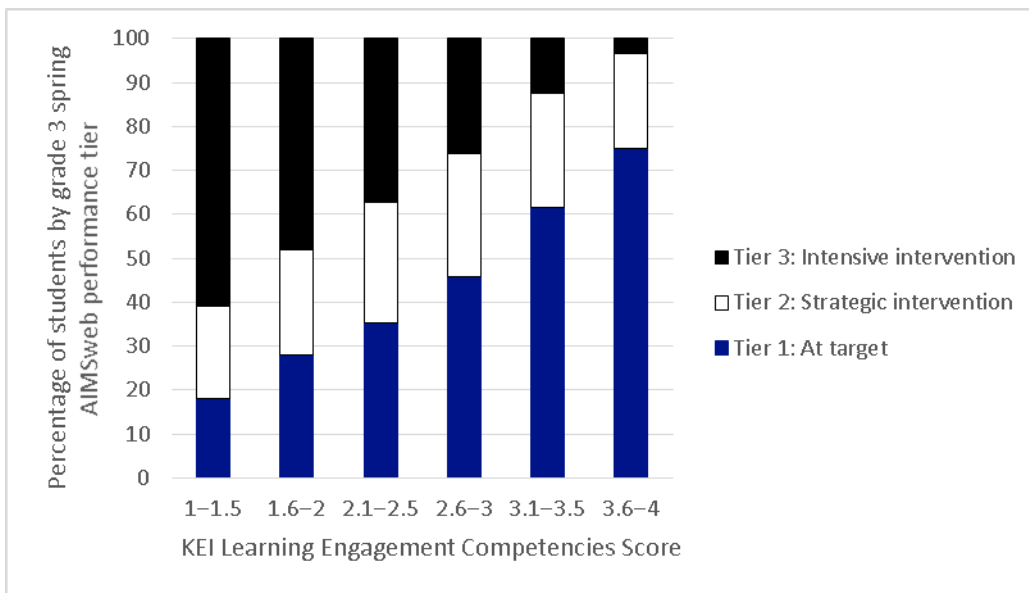


KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 4,460 students that had information on KEI Learning Engagement Competencies and grade 2 spring AIMSweb performance (. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2016/17 data described in appendix A.

**Figure B9. Students with higher Learning Engagement Competencies scores were more likely to be in Tier 1 (“at target”) on the AIMSweb in spring of grade 3**



KEI is Kindergarten Entry Inventory.

Note: Analysis was based on the random sample of students used to set the threshold and was restricted to the 4,177 students that had information on KEI Learning Engagement Competencies and grade 3 spring AIMSweb performance. The data the graph is based on was not weighted.

Source: Authors’ analysis based on School District of Philadelphia 2014/15 and 2017/18 data described in appendix A.

*Students in higher performing AIMSweb tiers in the spring of each year were more likely to be proficient on the grade 3 PSSA in ELA. The numeric values for figure 4, which illustrates this finding, are shown in table B5.*

**Table B5. Students in higher performing AIMSweb tiers in the spring of each year were more likely to be proficient on the grade 3 Pennsylvania System of School Assessment**

AIMSweb assessment	Proficient students tier 3 “intensive intervention” (percentage)	Proficient students in tier 2 “strategic intervention” (percentage)	Proficient students in tier 1 “at-target”(percentage)	Number of students
Spring kindergarten	10	24	51	4,696
Spring grade 1	16	31	56	4,174
Spring grade 2	2	23	66	5,027
Spring grade 3	3	28	67	5,266

Note: The table is based on the random sample of students used to select the threshold and the data is not weighted.  
 Source: Authors’ analysis based on School District of Philadelphia data 2014–2018 data described in appendix A.

*Results for the AIMSweb analyses were similar when based on a sample of students who had scores on all of the assessments used in the analyses.* The samples of students used in the AIMSweb analyses presented in the report vary across grades because they used all of the students who took each pair of assessments. To examine whether the differences in samples contributed to differences in findings across years, the study team conducted all of the analyses for a restricted sample of 2,358 students who had scores on both KEI dimensions, the AIMSweb in the spring of each year from kindergarten through grade 3, and the grade 3 PSSA in ELA. The patterns of correlations between the AIMSweb and KEI (table B6) and the AIMSweb and grade 3 PSSA (table B7) for this restricted sample were similar to those for the larger sample of students who took each pair of assessments. Likewise, the findings were similar across samples for the analyses examining the accuracy of the AIMSweb tier 1 cutoff in predicting grade 3 PSSA proficiency in ELA (table B8).

**Table B6. Kindergarten through grade 3 AIMSweb scores had weak to moderate correlations with Emerging Academic Competencies and Learning Engagement Competencies in the sample with assessments in each year**

KEI (fall of kindergarten)	Spring kindergarten AIMSweb	Spring grade 1 AIMSweb	Spring grade 2 AIMSweb	Spring grade 3 AIMSweb
Emerging Academic Competencies	0.45***	0.37***	0.44***	0.38***
Learning Engagement Competencies	0.32***	0.31***	0.27***	0.30***

KEI is Kindergarten Entry Inventory.

\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analyses were based on the random sample of students used to set the threshold and were not weighted. 2,358 students completed the KEI, each spring AIMSweb assessment from kindergarten through grade 3, and the PSSA.

Source: Authors’ analysis based on SDP 2014–2018 data described in appendix A.

**Table B7. Kindergarten through grade 3 AIMSweb scores were moderately to strongly correlated with grade 3 Pennsylvania System of School Assessment English Language Arts scores**

AIMSweb assessment	Grade 3 PSSA ELA Correlation
Spring kindergarten	0.49***
Spring grade 1	0.53***
Spring grade 2	0.70***
Spring grade 3	0.72***

PSSA is Pennsylvania System of School Assessment.

\* Significant at  $p = .05$ ; \*\* significant at  $p = .01$ ; \*\*\* significant at  $p = .001$ .

Note: Analyses were based on the random sample of students used to set the threshold and were not weighted. 2,358 students completed the KEI, each spring AIMSweb assessment from kindergarten through grade 3, and the PSSA.

Source: Authors’ analysis based on School District of Philadelphia 2014–2018 data described in appendix A.



**Table B8. The AIMSweb Tier 1 cutoff correctly predicted more than 49 percent of the students who were not proficient and more than 65 percent of the proficient students in English Language Arts on the grade 3 Pennsylvania System of School Assessment**

AIMSweb assessment and threshold	Correct predictions for:		
	Proficient students (percentage)	Not proficient students (percentage)	All students (percentage)
Spring kindergarten AIMSweb tier 1	83	49	63
Spring grade 1 AIMSweb tier 1	65	71	68
Spring grade 2 AIMSweb tier 1	81	73	76
Spring grade 3 AIMSweb tier 1	76	77	77

Note: Analyses were based on the random sample of students used to set the threshold and were not weighted. 2,358 students completed the KEI, each spring AIMSweb assessment from kindergarten through grade 3, and the PSSA.

Source: Authors' analysis based on School District of Philadelphia data 2014—2018 data described in appendix A.

#### October 2019

This report was prepared under Contract ED-IES-17-C-0006 by Regional Educational Laboratory Mid-Atlantic administered by Mathematica. The report has been peer reviewed and approved for publication by the U.S. Department of Education's Institute of Education Sciences. The content of the publication does not necessarily reflect the views or policies of the Institute of Education Sciences 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:

Harding, J. F., Herrmann, M. A., Hanno, E. S., & Ross, C. (2019). *Using kindergarten entry assessments to measure whether Philadelphia's students are on-track for reading proficiently*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic.