

Understanding the Relationship between Early Literacy Domains and Reading in Later Grades

Deanne A. Crone, Ph.D., Molly K. Duncan, Ph.D, and Julia Febiger, M.Ed.

Curriculum Associates Research Report | January 2023

Introduction

Educators and researchers have long recognized that low reading proficiency in elementary school is negatively related to later outcomes (Hernandez, 2011; World Literacy Foundation, 2015) and that students can be reliably identified as at risk in reading as early as kindergarten and first grade (Scheffel et al., 2012; Utchell et al., 2016). The academic achievement of students in early elementary grades was severely impacted by the COVID-19 pandemic and loss of school time during the critically formative years for establishing foundational reading skills (Curriculum Associates, 2022; Kuhfeld & Lewis, 2022). Research suggests that low literacy rates are correlated with low grades, high absences, and behavioral issues (Barnett, 1995). Moreover, when students aren't able to read well, they may be held back by third-grade retention policies, which may harm their confidence, interrupt their social structures, and set them further back in their growth (Bakken et al., 2017; Dickinson & Neuman, 2007; Hernandez, 2011; McCoy et al., 2017). For individual students and society, these can have long-lasting effects (Heckman et al., 2013; Levin et al., 2007; Reynolds et al., 2001; Weiss et al., 1991). Educators should have insight into both students' foundational reading skills as well as higher-level reading ability so they can intervene in order to improve reading proficiency for equitable outcomes and long-term success.

The purpose of this study was to understand the combined and unique contributions of how Reading domain performance in early grades is related to later overall Reading performance. The *i-Ready Diagnostic* for Reading provides domain level data for six domains that are essential for developing reading ability: Phonics, Phonological Awareness, High-Frequency Words, Vocabulary, Comprehension: Literature, and Comprehension: Informational Text. In this paper, we use the term “foundational reading skills domains” to refer to Phonological Awareness, Phonics, and High-Frequency Words. We use the term “language domains” to refer to Vocabulary, Comprehension: Literature, and Comprehension: Informational Text.

Student reading performance was tracked using data from the Diagnostic over two years for three cohorts of students in Grades K–2. Using linear regression supplemented by a descriptive exploration of student placement patterns, this longitudinal analysis found that foundational Reading domain placements in Grades K and 1 were relatively more important for predicting later reading performance, while language domain placements in Grade 2 were more important for predicting later reading performance. Furthermore, the results in this paper delineate how we can use students' domain placements in early elementary grades to predict students' later reading performance, which may be helpful for the early identification of students who would benefit from additional supports. Equipped with this information, educators can select the appropriate evidence-based reading instruction method necessary to put students on a path toward grade-level proficiency during the early elementary school years. Possible classroom implications are considered in the discussion section.

Methodology

This study was designed to address the following research question:

How does domain-level performance in Reading in Grade K, 1, or 2 predict overall Reading performance two years later, in Grade 2, 3, or 4?

Students who were in Grade K, 1, or 2 during the 2016–2017 school year were eligible for inclusion in this study. To be included in the analysis, students had to complete an *i-Ready Diagnostic* for Reading in winter and spring of the 2016–2017 school year as well as two years later during spring of the 2017–2018 and 2018–2019 school years. Although this study utilized only the winter 2016–

2017 (i.e., Year 1) and spring 2018–2019 (i.e., Year 3) time points, we required the additional time points for comparability of findings in future longitudinal analyses. Students were excluded from the final sample if: 1) their chronological grade level at any time point did not match the expected grade level (e.g., if the student was retained in a grade), or 2) the student’s Diagnostic was flagged with a red Rush flag, indicating that the student spent so little time on the assessment that they were likely “rushing” through the assessment with little effort. The final sample included 425,480 students from 45 states. The state with the most students was Florida, with 175,440 students, followed by California (41,065 students) and New York (35,080 students). Thirty-one additional states were represented by at least 1,000 students each. Table 1 shows students’ grade level at the beginning and end of the study and the final sample size for each grade-level cohort. School districts are not required to report demographic information for their students to Curriculum Associates. Therefore, reliable demographic data about this sample was not available.

Table 1: Cohorts by Grade Level and Sample Size

Name of Cohort	Year 1 Winter	→	Year 3 Spring	N Sample
Grade K Cohort	Grade K	→	Grade 2	113,945
Grade 1 Cohort	Grade 1	→	Grade 3	152,618
Grade 2 Cohort	Grade 2	→	Grade 4	158,917

Spring overall Reading scale scores during the 2018–2019 school year were predicted using winter domain placements from the 2016–2017 school year using linear regression. Linear regression was chosen as the analysis method because it allows the use of all domains simultaneously to predict the outcome. Within the construct of reading, each domain is related to other domains, so regression allowed us to isolate the unique contribution of a single domain, over and above the contribution of other domains. Regression also allowed us to accommodate the discontinuous nature of foundational skills domains, in which students can test out or earn a maximum score.

The Diagnostic classifies students into criterion-referenced placement levels based on a scale score for both overall Reading achievement and domain achievement. For the purposes of the exploratory analyses, students were placed into risk categories reflecting time of year. See Table 2 for a crosswalk between the Diagnostic’s criterion-referenced placement levels and the categories used in this analysis.

Table 2: Crosswalk of Analysis Categories and Diagnostic Grade-Level Placements

Analysis Category	Year 1 (Winter)	Year 3 (Spring)
Two or More Grade Levels Below	Two Grade Levels Below Three or More Grade Levels Below	Two Grade Levels Below Three or More Grade Levels Below
Approaching Grade Level	One Grade Level Below	One Grade Level Below Early On Grade Level
Grade Level	Early On Grade Level Mid On Grade Level Late On Grade Level Above Grade Level	Mid On Grade Level Late On Grade Level Above Grade Level

Due to the nature of the Diagnostic’s test flow, a student in Grade 2 who is on grade level can test out of Phonological Awareness. As such, we were not able to differentiate between Early On Grade Level, Mid On Grade Level, and Late On Grade Level placements for that domain and grade level. Additional information on the Diagnostic’s test flow can be found in Appendix A, and achievement-level descriptors can be found in Appendix B.

For each cohort, the following model was fitted:

$$Y_i = \beta_0 + \Sigma\beta_1(Ph) + \Sigma\beta_2(PA) + \Sigma\beta_3(HFW) + \Sigma\beta_4(Voc) + \Sigma\beta_5(Comp: Lit) + \Sigma\beta_6(Comp: Inf. Text)$$

In this equation, Y_i represents the predicted overall Reading score in Year 3 for student i , which is predicted as a function of the intercept (i.e., the predicted Year 3 score when the Year 1 placement for all domains is Early On Grade Level, represented by β_0) and the point estimate associated with each of the student's domain placements. The terms $\Sigma\beta_1$ through $\Sigma\beta_6$ each represent a vector of dummy-coded point estimates for the possible placements in that domain for the relevant cohort. Each vector excluded the placement associated with Early On Grade Level expectations because it was used as the reference. Therefore, the point estimate for the Early On Grade Level placement was zero. As such, the intercept can be interpreted as the estimated Year 3 reading score for a student who scored Early On Grade Level in all domains assessed in Year 1. The estimated Year 3 reading score for a student whose domain-specific placement in Year 1 was anything other than Early On Grade Level can be calculated by adding the point estimate associated with that domain-specific placement to the intercept. The models also allow the calculation of R^2 , the squared multiple correlation, which provides a measure of the proportion of variance in the outcome that is explained by the predictors (Pedhazur, 1997).

All analyses were conducted in R version 4.1.3 (R Core Team, 2022). Visual inspections of the models and data confirmed there were no major causes for concern about violation of the assumptions of linear regression (i.e., linearity, normality, and homoscedasticity of residuals). Assumptions were also not violated regarding multicollinearity or outlier effects.

Results

Overall, this analysis found that foundational Reading domain placements in Grades K and 1 were more strongly related to overall Reading scores two years later (at the end of Grades 2 and 3, respectively) compared to the language domains. Conversely, for students in Grade 2, the language domain placements were more strongly related to overall Reading scores two years later (at the end of Grade 4) than the foundational reading domains.

Examining Observed Placement Levels

As a preliminary analysis, descriptive patterns of domain placements in Year 1 compared to overall Reading placements in Year 3 were examined. Across cohorts and domains, the majority of students tended to achieve an overall Reading placement in the same category in Year 3 that they had achieved in Year 1 for the domain in question. In the following section, we will describe the patterns observed across all three cohorts in the Phonics domain.

For these tables, placements were grouped into three categories: Two or More Grade Levels Below, Approaching Grade Level, and Grade Level. These different categorizations are based on instructional “views” available in the platform for educators and reflect an expectation that students should place at a higher level at the end of the school year than in the beginning or middle of the school year. See the Methodology section for more detail.

For example, in the Grade K cohort, 61% of students who placed in the Approaching Grade Level category in Phonics in Year 1 also placed in the Approaching Grade Level category for overall Reading two years later. Of the Grade K students who were Grade Level in Phonics in Year 1, 64% were Grade Level in overall Reading two years later. Staying with the Grade K Phonics example, although 61% of students who were Approaching Grade Level in Year 1 were also Approaching Grade Level in overall Reading two years later, nearly one-third (i.e., 29%) of those students placed higher (i.e., Grade Level) in overall Reading two years later. Ten percent of these students Approaching Grade Level placed lower (i.e., Two or More Grade Levels Below) in overall Reading two years later. Among the students who started in the Grade Level category, approximately one-third (i.e., 34%) of the students placed one category lower in overall Reading (i.e., Approaching Grade Level), while 2% of students placed two categories lower in overall Reading (i.e., Two or More Grade Levels Below.) The percentage of students whose Year 3 overall Reading was two categories lower than Year 1 Phonics was quite small, representing just 1% of the overall sample of students. Table 3 includes results for the Grade K cohort for each domain.

Table 3: Grade K Cohort Placement by Domain in Year 1 and Placement by Overall Reading in Year 3

Cohort	Year 1 (Winter)			Year 3 (Spring) Overall Reading Placement Distribution		
	Domain	Domain Placement	Number of Students	Two or More Grade Levels Below	Approaching Grade Level	Grade Level
Grades K → 2	Phonics	Two or More Grade Levels Below	0	N/A	N/A	N/A
		Approaching Grade Level	45,148	10%	61%	29%
		Grade Level	68,797	2%	34%	64%
	Phonological Awareness	Two or More Grade Levels Below	0	N/A	N/A	N/A
		Approaching Grade Level	35,895	11%	62%	27%
		Grade Level	78,050	2%	36%	61%
	High-Frequency Words	Two or More Grade Levels Below	0	N/A	N/A	N/A
		Approaching Grade Level	46,849	10%	58%	32%
		Grade Level	67,096	1%	35%	63%
	Vocabulary	Two or More Grade Levels Below	0	N/A	N/A	N/A
		Approaching Grade Level	45,281	9%	59%	33%
		Grade Level	68,664	3%	35%	62%
	Comprehension: Literature	Two or More Grade Levels Below	0	N/A	N/A	N/A
		Approaching Grade Level	31,749	11%	61%	29%
		Grade Level	82,196	3%	38%	59%
	Comprehension: Informational Text	Two or More Grade Levels Below	0	N/A	N/A	N/A
		Approaching Grade Level	35,022	10%	60%	29%
		Grade Level	78,923	3%	37%	60%

Note: The lowest placement level that a Grade K student can receive is Emerging Grade K, which is considered Approaching Grade Level at all time points. For Grade K, there is no placement level that is Two or More Grade Levels Below. Thus, the corresponding row of data for the Grade K cohort reads N/A (i.e., not applicable).

For the Grade 1 cohort in Phonics, of the students who were Two or More Grade Levels Below in Year 1, just more than half (i.e., 53%) placed in the Two or More Grade Levels Below category for overall Reading in Year 3. Thirty-nine percent of these students performed one category higher (i.e., Approaching Grade Level) in overall reading in Year 3, and 8% placed two categories higher (i.e., Grade Level). For the students who started in the Approaching Grade Level category, 58% were in the Approaching

Grade Level category for overall Reading two years later, while about a quarter (i.e., 26%) of the students placed one category higher (i.e., Grade Level) and 16% one category lower (i.e., Two or More Grade Levels Below). Among the students in the Grade 1 cohort who started in the Grade Level category, two-thirds (i.e., 66%) also placed in the Grade Level category in Year 3, while approximately one-third (i.e., 32%) of the students placed one category lower in the Approaching Grade Level category. In the group of Grade 1 students who placed Grade Level in Phonics, only 2% were Two or More Grade Levels Below in overall Reading two years later. As with the Grade K cohort, the percentage of students who placed two categories lower was quite small, representing just 1% of the overall sample of students. Table 4 includes results for the Grade 1 cohort for each domain.

Table 4: Grade 1 Cohort Placement by Domain in Year 1 and Placement by Overall Reading in Year 3

Cohort	Year 1 (Winter)			Year 3 (Spring) Overall Reading Placement Distribution		
	Domain	Domain Placement	Number of Students	Two or More Grade Levels Below	Approaching Grade Level	Grade Level
Grades 1 → 3	Phonics	Two or More Grade Levels Below	4,490	53%	39%	8%
		Approaching Grade Level	66,959	16%	58%	26%
		Grade Level	81,169	2%	32%	66%
	Phonological Awareness	Two or More Grade Levels Below	5,361	45%	46%	9%
		Approaching Grade Level	59,694	17%	57%	26%
		Grade Level	87,563	3%	34%	63%
	High-Frequency Words	Two or More Grade Levels Below	5,894	49%	41%	10%
		Approaching Grade Level	47,557	20%	59%	21%
		Grade Level	99,167	2%	36%	61%
	Vocabulary	Two or More Grade Levels Below	7,506	40%	50%	10%
		Approaching Grade Level	78,459	14%	57%	29%
		Grade Level	66,653	2%	26%	72%
	Comprehension: Literature	Two or More Grade Levels Below	5,047	40%	49%	11%
		Approaching Grade Level	75,275	15%	58%	27%
		Grade Level	72,296	2%	28%	70%
	Comprehension: Informational Text	Two or More Grade Levels Below	6,459	38%	51%	11%
		Approaching Grade Level	76,010	14%	58%	27%
		Grade Level	70,149	2%	27%	71%

The results for the Grade 2 cohort show that of the students who were Two or More Grade Levels Below in Year 1, 38% were also in the Two or More Grade Levels Below category for overall reading in Year 3, while 56% placed one category higher in overall reading (i.e., Approaching Grade Level) and 6% placed two categories higher in Year 3 (i.e., Grade Level). For the students who started in the Approaching Grade Level category, more than two-thirds (i.e., 69%) were in the Approaching Grade Level category in overall Reading two years later, while 21% of the students placed one category higher (i.e., Grade Level) and 10%

placed one category lower (i.e., Two or More Grade Levels Below) in overall Reading. Among the students in the Grade 2 cohort who started in the Grade Level category, more than half (i.e., 56%) also placed in the Grade Level category in overall Reading, while approximately 42% of the students placed one category lower (i.e., Approaching Grade Level) and 2% of students placed two categories lower (i.e., Two or More Grade Levels Below). As with the Grades K and 1 cohorts, the percentage of students who placed two categories lower was quite small, representing just 1% of the overall sample of students. Table 5 includes results for the Grade 2 cohort for each domain.

Table 5: Grade 2 Cohort Placement by Domain in Year 1 and Placement by Overall Reading in Year 3

Cohort	Year 1 (Winter)			Year 3 (Spring) Overall Reading Placement Distribution		
	Domain	Domain Placement	Number of Students	Two or More Grade Levels Below	Approaching Grade Level	Grade Level
Grades 2 → 4	Phonics	Two or More Grade Levels Below	22,817	38%	56%	6%
		Approaching Grade Level	53,828	10%	69%	21%
		Grade Level	82,272	2%	42%	56%
	Phonological Awareness	Two or More Grade Levels Below	10,418	52%	44%	4%
		Approaching Grade Level	4,345	32%	62%	6%
		Grade Level	144,154	6%	54%	40%
	High-Frequency Words	Two or More Grade Levels Below	10,711	49%	45%	6%
		Approaching Grade Level	17,008	24%	64%	11%
		Grade Level	131,198	5%	53%	43%
	Vocabulary	Two or More Grade Levels Below	18,887	41%	55%	4%
		Approaching Grade Level	67,697	10%	71%	19%
		Grade Level	72,333	1%	37%	62%
	Comprehension : Literature	Two or More Grade Levels Below	17,293	42%	54%	5%
		Approaching Grade Level	53,788	13%	72%	16%
		Grade Level	87,836	1%	42%	56%
	Comprehension : Informational Text	Two or More Grade Levels Below	19,660	40%	55%	5%
		Approaching Grade Level	53,960	12%	72%	17%
		Grade Level	85,297	1%	42%	57%

These data provide preliminary evidence for the predicted pattern of relationships between students' performance in early Reading domains and subsequent overall Reading achievement. In particular, the domains for which more students had similar Year 1 domain placements and Year 3 overall placements were expected to be stronger predictors than the domains for which relatively fewer students had similar placements in Year 1 and Year 3. However, because each of the Reading domains are related, these descriptive statistics cannot isolate the relationship between an individual domain and the later overall score.

Predicting Overall Reading Scale Scores

In order to examine the relationship of individual domains with later overall reading scores, we used linear regression to predict overall Reading scale scores in Year 3 from the Year 1 domain placements for each cohort. Across all cohorts, lower placement levels within each domain in Year 1 of the study predicted a lower overall Reading score in Year 3 of the study, while higher placement levels predicted higher overall Reading scores. There were differences by cohort and by domain in the magnitude of the point estimates and in the corresponding prediction of overall Reading scores. For students in the Grades K and 1 cohorts, results showed that the foundational skills domains had a stronger predictive relationship with overall Reading in Grades 2 and 3, respectively, compared to the language domain placements. Detailed results of the regression analyses for each cohort are provided in the tables in Appendix C.

Within the Grade K cohort, students who placed Early On Grade Level in all six domains in winter of Year 1 were predicted to have an overall Reading score of 510 by the end of Year 3, i.e., spring of Grade 2. As such, a score of 510 is the intercept and can be interpreted as the baseline for the Grade K cohort in this model. For Grade K students who placed Early On Grade Level in all domains except Phonics, where they placed One Grade Level Below, the predicted overall score in Grade 2 was 498, or 12 scale score points lower than the baseline. For each of the other domains, the predicted scale score associated with a placement of One Grade Level Below was also lower, ranging from 3 points lower (i.e., Vocabulary) to 11 points lower (i.e., Phonological Awareness). See Table 6.

The model also predicted scores when multiple domain placements were different from Early On Grade Level. Performing One Grade Level Below in all three foundational skills domains (i.e., Phonics, Phonological Awareness, and High-Frequency Words) in Year 1 predicted a much lower scale score in Year 3 (i.e., 476) compared to performing One Grade Level Below in all three language domains (i.e., Vocabulary, Comprehension: Literature, and Comprehension: Informational Text), which predicted a scale score of 497 in Year 3. That is, students were predicted to perform 34 points lower than baseline when performance in the foundational skills domains in Year 1 was One Grade Level Below, compared to 13 points lower than baseline when performance in the language domains in Year 1 was One Grade Level Below. See Table 6. The full model R^2 for this model was .36, indicating that the model explained about 36% of the variance in Grade 2 scores.

Table 6: Predicting Grade 2 Overall Reading Score Based on Grade K Reading Domain Placements

Grade K		Predicted Overall Reading Score in Grade 2	Difference from Baseline
Early On Grade Level	One Grade Level Below		
Baseline—All domains (i.e., Phonics, Phonological Awareness, High-Frequency Words, Vocabulary, Comprehension: Literature, Comprehension: Informational Text)	None	510	N/A
All domains except Phonics	Phonics	498	-12
All domains except Phonological Awareness	Phonological Awareness	499	-11
All domains except High-Frequency Words	High-Frequency Words	500	-10
All domains except Vocabulary	Vocabulary	507	-3
All domains except Comprehension: Literature	Comprehension: Literature	505	-5
All domains except Comprehension: Informational Text	Comprehension: Informational Text	505	-5
All language domains (i.e., Vocabulary, Comprehension: Literature, Comprehension: Informational Text)	Phonics, Phonological Awareness, High-Frequency Words	476	-34

All foundational skills domains (i.e., Phonics, Phonological Awareness, High-Frequency Words)	Vocabulary, Comprehension: Literature, Comprehension: Informational Text	497	-13
---	--	-----	-----

Note: Difference from baseline when multiple domains are One Grade Level Below may differ slightly from the sum of the differences reported for individual domains due to rounding. More precise point estimates can be found in Appendix C.

Table 7 includes results for the Grade 1 cohort. Within the Grade 1 cohort, students who placed Early On Grade Level in all six domains in winter of Year 1 were predicted to have an overall score of 551 by the end of Year 3, i.e., spring of Grade 3. This score of 551 is the intercept and can be interpreted as the baseline for the Grade 1 cohort in this model. For Grade 1 students, Phonological Awareness and High-Frequency Words were the domains associated with the largest differences in later overall score. For Grade 1 students who placed Early On Grade Level in all domains except Phonological Awareness, in which they placed One Grade Level Below, the predicted overall score in Grade 3 was 541, or 10 scale score points lower than the baseline. For Grade 1 students who placed Early On Grade Level in all domains except High-Frequency Words, in which they placed One Grade Level Below, the predicted overall score in Grade 3 was 538, or 13 points lower than the baseline. For each of the other domains, the predicted overall scale score was also lower if the Year 1 domain was One Grade Level Below instead of Early On Grade Level, ranging from 5 points lower (i.e., Comprehension: Literature and Comprehension: Informational Text) to 9 points lower (i.e., Vocabulary).

The model also predicted scores when multiple domain placements were different from Early On Grade Level. Performing One Grade Level Below in all three foundational skills domains in Year 1 predicted a lower overall Reading scale score in Year 3 (i.e., 520) compared to performing One Grade Level Below in all three language domains, which predicted a scale score of 533 in Year 3. That is, students are predicted to perform 31 points lower than baseline when performance in the foundational skills domains in Year 1 was One Grade Level Below, compared to 18 points lower than baseline when performance in the language domains in Year 1 was One Grade Level Below. The full model R^2 for this model was .48, indicating that the model explained about 48% of the variance in Grade 3 scores.

Table 7: Predicting Grade 3 Overall Reading Score Based on Grade 1 Reading Domain Placements

Grade 1		Predicted Overall Reading Score in Grade 3	Difference from Baseline
Early On Grade Level	One Grade Level Below		
Baseline—All domains (i.e., Phonics, Phonological Awareness, High-Frequency Words, Vocabulary, Comprehension: Literature, Comprehension: Informational Text)	None	551	N/A
All domains except Phonics	Phonics	543	-8
All domains except Phonological Awareness	Phonological Awareness	541	-10
All domains except High-Frequency Words	High-Frequency Words	538	-13
All domains except Vocabulary	Vocabulary	542	-9
All domains except Comprehension: Literature	Comprehension: Literature	546	-5
All domains except Comprehension: Informational Text	Comprehension: Informational Text	546	-5
All language domains (i.e., Vocabulary, Comprehension: Literature, Comprehension: Informational Text)	Phonics, Phonological Awareness, High-Frequency Words	520	-31

All foundational skills domains (i.e., Phonics, Phonological Awareness, High-Frequency Words)	Vocabulary, Comprehension: Literature, Comprehension: Informational Text	533	-18
---	--	-----	-----

Note: Difference from baseline when multiple domains are One Grade Level Below may differ slightly from the sum of the differences reported for individual domains due to rounding. More precise point estimates can be found in Appendix C.

For students in the Grade 2 cohort, results showed that the language domains had a stronger predictive relationship with overall reading in Grade 4 compared to the foundational skills domain placements. However, it is important to note that this finding may be partially related to ceiling effects given that students may test out of foundational skills domains in Grades 2, 3 and 4 per the Diagnostic’s test flow (see Appendix A).

As Table 8 shows, students in the Grade 2 cohort who placed Early On Grade Level in all six domains in winter of Year 1 were predicted to have an overall score of 569 by the end of Year 3, or spring of Grade 4. A score of 569 is the intercept and can be interpreted as the baseline for the Grade 2 cohort for this model. For Grade 2 students who placed Early On Grade Level in all domains except Vocabulary, in which they placed One Grade Level Below, the predicted overall score in Grade 4 was 558, or 11 scale score points lower than the baseline. For each of the other domains, the predicted scale score was also lower if the Year 1 domain was One Grade Level Below instead of Early On Grade Level, ranging from 2 points lower (i.e., Phonological Awareness) to 8 points lower (i.e., Comprehension: Literature and Comprehension: Informational Text).

The model also predicted scores when multiple domain placements were different from Early On Grade Level. Performing One Grade Level Below in all three language domains in Year 1 predicted a lower scale score in Year 3 (i.e., 542) compared to performing One Grade Level Below in all three foundational skills domains, which predicted a scale score of 558 in Year 3. That is, students were predicted to perform 27 points lower than baseline when performance in the language domains in Year 1 was One Grade Level Below, compared to 11 points lower than baseline when performance in the foundational skills domains in Year 1 was One Grade Level Below. The full model R^2 for this model was .55, indicating that the model explained about 55% of the variance in Grade 2 scores.

Table 8: Predicting Grade 4 Overall Reading Score Based on Grade 2 Reading Domain Placements

Grade 2			
Early On Grade Level in Grade 2	One Grade Level Below in Grade 2	Predicted Overall Reading Score in Grade 4	Difference from Baseline
Baseline—All domains (i.e., Phonics, Phonological Awareness, High-Frequency Words, Vocabulary, Comprehension: Literature, Comprehension: Informational Text)	None	569	N/A
All domains except Phonics	Phonics	562	-7
All domains except Phonological Awareness	Phonological Awareness	567	-2
All domains except High-Frequency Words	High Frequency Words	566	-3
All domains except Vocabulary	Vocabulary	558	-11
All domains except Comprehension: Literature	Comprehension: Literature	561	-8
All domains except Comprehension: Informational Text	Comprehension: Informational Text	561	-8

All language domains (i.e., Vocabulary, Comprehension: Literature, Comprehension: Informational Text)	Phonics, Phonological Awareness, High-Frequency Words	558	-11
All foundational skills domains (i.e., Phonics, Phonological Awareness, High-Frequency Words)	Vocabulary, Comprehension: Literature, Comprehension: Informational Text	542	-27

Note: Difference from baseline when multiple domains are One Grade Level Below may differ slightly from the sum of the differences reported for individual domains due to rounding. More precise point estimates can be found in Appendix C.

Limitations

A multiple regression model was selected as the method of analysis because it provided easily interpretable and actionable results. However, future studies could utilize structural equation modeling to increase the reliability of the incremental validity estimates (Westfall & Yarkoni, 2016) or multilevel modeling to account for any dependence among scores within schools or school districts (Raudenbush & Bryk, 2002).

While the use of a multiple regression model allowed us to make predictions about the performance of students who showed unlikely scoring patterns and to isolate the relationship between specific domains and later overall performance, it should be noted that the most common placement-level pattern in this sample was for a student to score at similar grade levels across domains. When a student placed Early On Grade Level for one domain, it was more common for that student to place Early On Grade Level for all domains. It was less common for a student to place Early On Grade Level in some domains and One Grade Level Below in others. The estimates we presented in this paper are based on the predictions of the model and are not necessarily descriptions of how actual students performed. More studies are needed to replicate our conclusions.

This study used Reading domain placement levels rather than Reading domain scale scores to maintain consistency between grade levels and to provide more interpretable results. However, this reduced the granularity of the Reading score estimates and meant that those estimates were influenced in part by the range of scale score points included within each placement level by grade. Future analyses could select a different methodology that would utilize scale scores to address a similar research question.

Lastly, this study used assessment data collected prior to the COVID-19 pandemic. This choice allowed us to examine longitudinal reading achievement under typical conditions of teaching and learning. The pandemic and the resulting changes to the educational landscape may have lasting impacts on young students' learning in ways we do not yet understand. Further research could examine whether the longitudinal trends we observed here are similar or different, using assessment data collected during or after the pandemic.

Discussion

The results from this analysis provide evidence that domain-level placements in the early elementary grades can be used to predict overall Reading scores two years later. The regression models indicated that foundational skills domain placements were more strongly related to later overall Reading for Grades K and 1, while language domain placements were more strongly related to later overall Reading for Grade 2. Grade K students who performed below grade-level expectations in all three foundational skills domains will, on average, still be one grade level below expectations in overall Reading by the time they finish Grade 2. In other words, students likely to perform below grade level in Grade 2 can be predicted as early as winter of Grade K. Recent research shows that students who are performing below grade level in Grade 2 continue to remain behind their peers and behind grade-level expectations with each passing school year (Dawson, 2022).

These findings are particularly important now more than ever. At the time of this study's publication, we are halfway through the 2022–2023 school year, which means that students who were in Grades K, 1, and 2 when the pandemic hit are now in Grades 3, 4, and 5, respectively. While the analysis is based on data from before the pandemic, we know from research conducted during the pandemic that fewer students were ready for grade-level work in the 2020–2021 school year compared to historical averages—particularly in the lower grades (Curriculum Associates, 2021)—and that students who were furthest behind before COVID-19 (i.e., students who performed two or more grade levels below their chronological grade on the Diagnostic) saw the biggest decrease

in growth during the pandemic when compared to the growth of a pre-COVID-19 cohort of similar students (Dawson, 2021). The most recent results from the Nation's Report Card showed that average reading scores for fourth grade students had the largest score drop in more than 30 years, and the downturn in Reading scores was worst for students in the lowest quartile of performance, that is, for the most vulnerable students (National Assessment of Educational Progress, 2022). Given this context, we can hypothesize that, if anything, more students would be placing below grade level in Reading domains as we prepare to publish this study in winter 2023. This means that, for the second year in a row, we anticipate fewer students will be on grade level this coming spring than prior to the pandemic.

There are three major implications for educators based on this research. The first implication is that understanding students' literacy skills in the early grades is key. We can identify students who may need support with reading based on their domain-level reading skills as early as kindergarten. Early identification of students' strengths as well as any skills deficits in the early elementary grades, paired with an appropriate and evidence-based instructional response, is critical to ensuring students develop the foundational skills necessary to read with comprehension by third grade and beyond (Blachman et al., 2004; Ehri, 2020; Good et al., 2001). Utilizing a computer-administered reading assessment with coverage in critical early literacy domains enables teachers to identify students' domain-level reading skills and immediately deliver high-quality, evidence-based instruction. This is particularly important given the educational interruptions caused by COVID-19.

The second implication is that performance in foundational skills domains in the early grades is closely linked to later overall reading proficiency. Providing educators with domain level data that can be used within an evidence-based curriculum aligned to the Science of Reading can support all students in learning to read on grade level. Students do not learn how to read spontaneously but rather must be taught how to read explicitly. Foundational reading skills should be taught explicitly and systematically to all students in the early elementary grades so students can read at grade-level expectations in later years (Blachman et al., 2004; Ehri, 2020). While evidence-based instructional strategies and curricula that can be tailored to meet students' needs at the domain level are essential, it is equally as important to prevent the need for intervention by providing every student with an equitable, science-backed curriculum.

The third implication is that classroom-administered assessments can be efficient and sufficient. The Science of Reading describes the skills and processes necessary to develop skilled reading. These include, but are not limited to, phonological awareness, phonics, letter-sound correspondence, fluency, background knowledge, sight word recognition, verbal reasoning, print concepts, and vocabulary (Duke & Cartwright, 2021; The Reading League, 2022; Scarborough, 2001). This research supports the relationship between a range of early reading skills and subsequent reading achievement using an assessment that does not require one-on-one administration.

In summary, students' starting points in reading do not have to determine their end points. A student's performance in early elementary is an instructionally important indicator, but it does not have to determine their academic trajectory. Every educator should have access to information on students' domain-level reading skills to provide appropriate, targeted, and effective instruction early in elementary school. If teachers are empowered with knowledge about where students are and what they need instructionally to be successful at grade level, including the use of evidence-based strategies for any instructional tier, then perhaps we can make a positive impact on students' reading trajectories.

References

- Bakken, L., Brown, N., & Downing, B. (2017). Early childhood education: The long-term benefits. *Journal of Research in Childhood Education*, 31(2), 255–269. <https://doi.org/10.1080/02568543.2016.1273285>
- Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 5(3), 25–50. <https://doi.org/10.2307/1602366>
- Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., Shaywitz, B. A., & Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology*, 96(3), 444–461.
- Curriculum Associates. (2021). *i-Ready scale score placement tables (2021–2022 school year)*.
- Curriculum Associates. (2022). *The state of student learning in 2022*. <https://stateoflearning.curriculumassociates.com/>
- Dawson, M. (2021). *The impact of COVID-19 on student academic growth in 2020–2021*. Curriculum Associates.
- Dawson, M. (2022). *Student growth during COVID-19: Grade-level readiness matters*. Curriculum Associates.
- Dickinson, D. K., & Neuman, S. B. (2007). *Handbook of early literacy research: Volume 2*. Guilford Press.
- Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly*, 56(S1), S25–S44.
- Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly*, 55(S1), S45–S60. <https://doi:10.1002/rrq.334>
- Good, R. H. III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, 5(3), 257–288.
- Heckman, J., Pinto, R., & Saveliev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6), 2052–2086. <https://doi.org/10.1257/aer.103.6.2052>
- Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. The Annie E. Casey Foundation.
- Kuhfeld, M., & Lewis, K. (2022). *COVID-19 in the early elementary years: A comparison of achievement in spring 2019 and spring 2022*. NWEA.
- Levin, H. M., Belfield, C. R., Muennig, P., & Rouse, C. (2007). *The costs and benefits of an excellent education for all of America's children*. University of Pennsylvania Center for Benefit-Cost Studies of Education.
- McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., Yang, R., Koepp, A., & Shonkoff, J. P. (2017). Impacts of early childhood education on medium- and long-term educational outcomes. *Educational Researcher*, 46(8), 474–487. <https://doi.org/10.3102/0013189X17737739>
- National Assessment of Educational Progress (2022). *NAEP report card: 2022 NAEP reading assessment*. <https://www.nationsreportcard.gov/highlights/reading/2022/>
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction (3rd ed.)*. Harcourt Brace College Publishers.
- R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Sage Publications.

- Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *JAMA*, *285*(18), 2339–2346.
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research (Vol. 1, pp. 97–110)*. Guilford Press.
- Scheffel, D., Lefly, D., and Houser, J. (2012). The predictive utility of DIBELS reading assessment for reading comprehension among third grade English Language Learners and English-speaking children. *Reading Improvement*, *49*(3), 75–92.
- The Reading League. (2022). *Science of reading: Defining guide..* <https://www.thereadingleague.org/what-is-the-science-of-reading/>
- Utchell, L. A., Schmitt A. J., McCallum, E., McGoey, K. E., & Piselli, K. (2016). Ability of early literacy measures to predict future state assessment performance. *Journal of Psychoeducational Assessment*, *34*(6), 511–523.
- Weiss, B. D., Hart, G., & Pust, R. E. (1991). The relationship between literacy and health. *Journal of Health Care for the Poor and Underserved*, *1*(4), 351–363. <https://doi.org/10.1353/hpu.2010.0294>
- Westfall, J. & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS ONE*, *11*(3), 1–22. <https://doi.org/10.1371/journal.pone.0152719>
- World Literacy Foundation. (2015). *The economic & social cost of illiteracy: A snapshot of illiteracy in a global context.* .

Appendices

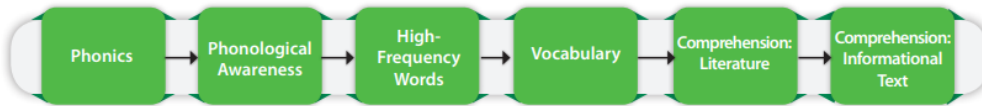
Appendix A: *i-Ready Diagnostic* for Reading Test Flow

The *i-Ready Diagnostic* for Reading’s test flow varies by chronological grade level. All students, regardless of grade level, are presented with items from the Vocabulary, Comprehension: Literature, and Comprehension: Informational Text domains. Students in lower chronological grade levels or performing at lower placement levels may also be assigned Phonics, Phonological Awareness, and High-Frequency Words items. Figure A1 below presents the test flow for reading in Grades K–8.

Figure A1: *i-Ready Diagnostic* Test Flow, Grades K–8

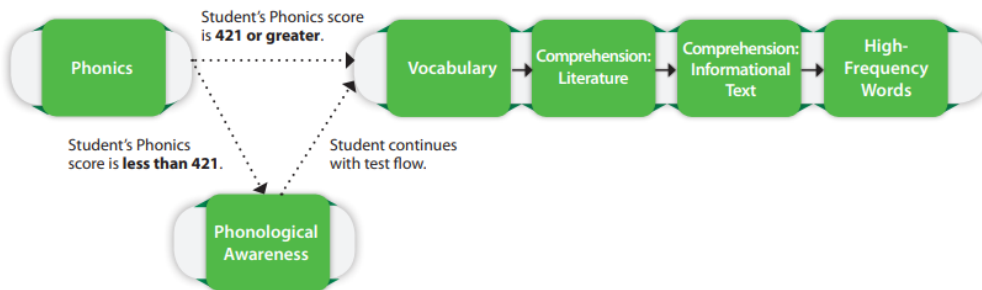
Grades K-1

Students who are in chronological Grades K–1 will receive items in all six Reading domains in the following order:



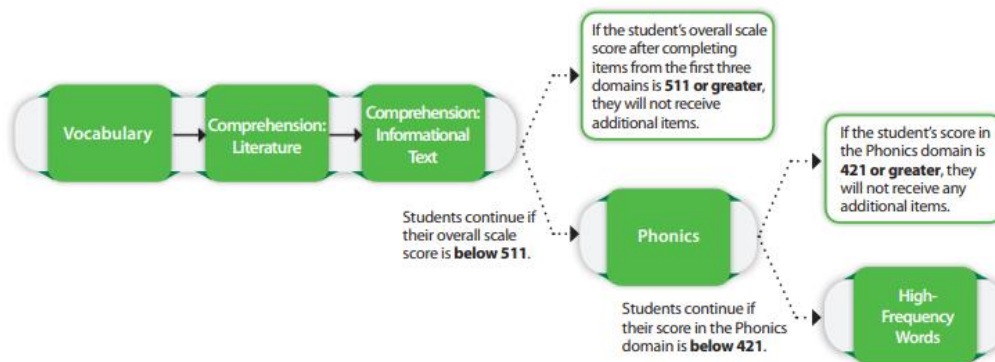
Grade 2

Students who are in chronological Grade 2 are first assessed in the Phonics domain. The student’s **overall** scale score after completing Phonics items determines the rest of their test flow:



Grades 3-8

Students who are in chronological Grades 3–8 are first assessed in the following domains: Vocabulary, Comprehension: Literature, and Comprehension: Informational Text. The student’s **overall** scale score after completing these domains determines if the test ends or continues. If the test continues, the student will receive Phonics, and the score on Phonics determines if the test ends or continues to High-Frequency Words.



Appendix B: *i-Ready Diagnostic* Grade-Level Placements and Achievement-Level Descriptors

The Diagnostic classifies students into criterion-referenced placement levels based on a scale score for both overall Reading achievement and domain-level achievement. Students who place below or above their chronological grade level are classified into a placement grade level (i.e., Levels K–8), while students who place on grade level are assigned a placement of Early, Mid, or Late On Grade Level. Students who place below grade level are assigned a relative placement level of One Grade Level Below, Two Grade Levels Below, or Three or More Grade Levels Below. The Diagnostic’s placement levels are criterion referenced, reflecting what students are expected to know at each grade level and in each content area. Table B1 provides the achievement-level descriptors for the Diagnostic’s Grade-Level Placements.

Table B1: Diagnostic Grade-Level Placement Descriptors

Grade-Level Placement	Achievement-Level Descriptor
One Grade Level Below Two Grade Levels Below Three or More Grade Levels Below	Additional support focused on below-grade level material is recommended to areas for improvement in students’ foundational knowledge. Students in these levels are not close to meeting the expectations of college- and career-ready standards for their grade level.
Early On Grade Level	Students in this level will benefit from on-grade level instruction to help them meet the expectations of college- and career-ready standards for their grade level. Students in Early On Grade Level have only partially met these grade-level expectations.
Mid On Grade Level	Students in this level will benefit from instruction in late on-grade level topics. These students have met the minimum requirements for the expectations of college- and career-ready standards in their grade level.
Late On Grade Level	Students in this level will benefit from late on-grade-level enrichment and will be ready for instruction focused on topics typically covered in the beginning of the subsequent grade level. Students in Late On Grade Level have successfully met or surpassed the grade-level expectations of college- and career-ready standards.
Above Grade Level	Students in this category will benefit from above-grade level instruction. Students in Above Grade Level have successfully met or surpassed all the expectations of college- and career-ready standards for their grade level as well as some expectations from subsequent grade levels.

Appendix C: Results of the Multiple-Regression Model

The full results of the multiple-regression model including intercepts and domain-by-placement-level point estimates are presented in Tables C1 (i.e., Grade K cohort), C2 (i.e., Grade 1 Cohort) and C3 (i.e., Grade 2 cohort). The intercept can be interpreted as the estimated Year 3 reading score for a student who scored Early On Grade Level in all domains assessed in Year 1. The estimated Year 3 reading score for a student with one or more domain-specific placements in Year 1 other than Early On Grade Level can be calculated by adding the point estimate associated with that domain-specific placement to the intercept. Note that the point estimate reported in the B column sometimes differs from the Difference from Baseline column in Tables 6, 7, and 8 due to rounding.

Table C1: Grade K Cohort—Results of the Multiple Regression Model

		<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Grade 2 Intercept (Predicted Score When All Grade K Domain Placements Are Early On Grade Level)		510.41	.45	1144.66	< .01
Phonics	One Grade Level Below	-12.90	.28	-45.30	< .01
	Mid On Grade Level	4.87	.35	14.01	< .01
	Late On Grade Level	6.97	.68	10.19	< .01
	Above Grade Level	8.70	.47	18.69	< .01
Phonological Awareness	One Grade Level Below	-11.43	.28	-40.11	< .01
	Late On Grade Level	6.02	.48	12.46	< .01
	Above Grade Level	12.23	.35	34.88	< .01
High-Frequency Words	One Grade Level Below	-10.49	.37	-28.64	< .01
	Mid On Grade Level	4.21	.40	10.54	< .01
	Late On Grade Level	9.52	.47	20.46	< .01
	Above Grade Level	15.99	.50	32.16	< .01
Vocabulary	One Grade Below	-3.23	.28	-11.45	< .01
	Mid or Late On Grade Level	5.66	.32	17.78	< .01
	Above Grade Level	14.70	1.00	14.66	< .01
Comprehension: Literature	One Grade Level Below	-5.48	.32	-17.29	< .01
	Mid or Late On Grade Level	4.32	.29	14.83	< .01
	Above Grade Level	11.66	.97	12.01	< .01
Comprehension: Informational Text	One Grade Level Below	-5.20	.31	-16.85	< .01
	Mid or Late On Grade Level	5.83	.29	20.00	< .01
	Above Grade Level	13.83	.86	16.07	< .01

Table C2: Grade 1 Cohort—Results of the Multiple Regression Model

		<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Grade 3 Intercept (Predicted Score When All Grade 1 Domain Placements Are Early On Grade Level)		550.77	.69	800.78	< .01
Phonics	Two Grade Levels Below	-28.30	.65	-43.45	< .01
	One Grade Level Below	-7.89	.30	-26.71	< .01
	Mid On Grade Level	.80	.34	2.37	.02
	Late On Grade Level	1.65	.47	3.50	< .01
	Above Grade Level	4.01	.36	11.18	< .01
Phonological Awareness	Two Grade Levels Below	-21.42	.70	-30.69	< .01
	One Grade Level Below	-9.75	.47	-20.69	< .01
	Mid On Grade Level	.62	.50	1.23	.22
	Late On Grade Level	3.59	.53	6.81	< .01
	Above Grade Level	6.69	.52	12.78	< .01
High-Frequency Words	Two Grade Levels Below	-29.33	.60	-49.03	< .01
	One Grade Level Below	-12.90	.35	-36.50	< .01
	Mid or Late On Grade Level	1.89	.37	5.09	< .01
	Above Grade Level	2.21	.39	5.72	< .01
Vocabulary	Two Grade Levels Below	-18.44	.52	-35.29	< .01
	One Grade Level Below	-8.62	.27	-31.93	< .01
	Mid On Grade Level	2.20	.36	6.09	< .01
	Late On Grade Level	9.57	.37	25.56	< .01
	Above Grade Level	21.94	.72	30.67	< .01
Comprehension: Literature	Two Grade Levels Below	-10.36	.61	-16.92	< .01
	One Grade Level Below	-4.60	.29	-16.08	< .01
	Mid On Grade Level	1.00	.37	2.68	.01
	Late On Grade Level	6.65	.37	17.82	< .01
	Above Grade Level	17.05	.56	30.30	< .01
Comprehension: Informational Text	Two Grade Levels Below	-13.50	.56	-24.24	< .01
	One Grade Level Below	-4.69	.29	-16.22	< .01
	Mid On Grade Level	.95	.39	2.47	.01
	Late On Grade Level	7.03	.37	18.76	< .01
	Above Grade Level	18.94	.56	33.77	< .01

Table C3: Grade 2 Cohort—Results of the Multiple Regression Model

		<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Grade 4 Intercept (Predicted Score When All Grade 2 Domain Placements Are Early On Grade Level)		568.83	.41	1371.73	< .01
Phonics	Two Grade Levels Below	-12.99	.43	-30.53	< .01
	One Grade Level Below	-6.57	.22	-29.53	< .01
	Mid On Grade Level	4.74	.33	14.48	< .01
	Late On Grade Level	7.17	.47	15.16	< .01
	Above Grade Level	9.50	.31	30.42	< .01
Phonological Awareness	Two Grade Levels Below	-13.45	.52	-25.97	< .01
	One Grade Level Below	-1.81	.62	-2.93	< .01
High-Frequency Words	Two Grade Levels Below	-16.83	.46	-36.93	< .01
	One Grade Level Below	-2.93	.37	-7.97	< .01
	Above Grade Level	1.59	.29	5.47	< .01
Vocabulary	Two Grade Levels Below	-22.63	.39	-57.75	< .01
	One Grade Level Below	-10.47	.24	-43.16	< .01
	Mid On Grade Level	5.68	.30	19.12	< .01
	Late On Grade Level	12.16	.36	33.72	< .01
	Above Grade Level	19.67	.45	43.47	< .01
Comprehension: Literature	Two Grade Levels Below	-15.75	.41	-38.51	< .01
	One Grade Level Below	-8.03	.26	-31.16	< .01
	Mid On Grade Level	4.75	.29	16.23	< .01
	Late On Grade Level	8.99	.32	28.53	< .01
	Above Grade Level	16.81	.36	46.68	< .01
Comprehension: Informational Text	Two Grade Levels Below	-15.64	.40	-39.36	< .01
	One Grade Level Below	-7.92	.26	-30.87	< .01
	Mid On Grade Level	5.55	.30	18.69	< .01
	Late On Grade Level	11.37	.32	35.18	< .01
	Above Grade Level	20.24	.36	56.73	< .01