OHIO COUNTY SCHOOL DISTRICT

ARP ESSER LEA PLAN

Ohio County Schools will utilize approximately 5% of ARP ESSER funds to implement prevention and mitigation strategies consistent with the Centers for Diseases Control and Prevention guidance on reopening schools in the following ways:

- Provide medical services related training to include methods for prevention of transmission, social distancing, symptom detection, vaccination education in addition to First Aid and CPR.
- Contract with resource officers for COVID prevention and response by utilizing them to help control/limit school visitors, monitor/manage social distancing, assist in distributing materials to students at home, and respond to emergencies connected to COVID at school during the instructional day and during extra-curricular school activities.
- Upgrade HVAC systems at designated schools to improve air quality and prevent COVID.
- Provide air filtration services to improve the air quality and prevent COVID.
- Provide masks, gloves, hand sanitizer, disinfectants, and other cleaning/custodial supplies to prevent the spread of COVID.

Funding provided through the ESSER I and ESSER II grants will be utilized in addition to the ARP ESSER funds to purchase additional necessary demand for medical and hygiene related supplies at adequate levels to prevent, prepare, and respond to COVID.

The Ohio County School District will utilize approximately 90% of ARP ESSER funds to address the academic impact of lost instructional time by providing a wide variety of evidence based interventions to students. These evidence based interventions will be utilized during the regular school year as well as during summer school sessions. Literacy coaches will be assigned to each of the six elementary schools in the district. These literacy coaches will work with students and teachers. While working with students, literacy coaches will utilize Reading Recovery lessons and strategies in a one on one setting as well as small group settings grades k-6. Beyond small group and one on one instruction literacy coaches will work with teachers to ensure best practice literacy instruction is taking place in all classrooms. Literacy coaches will also monitor reading progress data throughout the school year via Reading Running Records and iReady Diagnostic testing results. Literacy coaches and classroom teachers will utilize the following resources to ensure student success:

- Reading Recovery Lessons
- Reading Recovery Literacy Lessons
- Reading Running Records
- Small guided reading groups in all classrooms (Literacy Footprints Materials grades k-6 along with leveled books for school libraries)
- Rise Program (small group interventions grades 1-6)

ARP ESSER funds will be utilized to hire an intervention teacher at each of our six elementary schools as well as 2 intervention teachers at the Ohio County Middle School and 2 intervention teachers at Ohio County High School for three school years. Intervention teachers at the elementary level will focus on learning gaps in the area of math due to lost instructional time. Intervention teachers will review classroom data such as Math Running Records as well as iReady Diagnostic testing results and weekly classwork. The math intervention staff members at the elementary level will collaborate with teachers and share instructional resources with teachers throughout the school as well as working with small groups of students to close instructional gaps. Elementary intervention teachers and elementary classroom teachers will utilize the following resources to ensure student success.

- iReady Diagnostic Assessment
- Math Running Records
- iReady Online Prescriptive Instruction
- Ready Classroom Math Resources
- Various supplemental resources and manipulatives

Intervention teachers at the middle school will focus on learning gaps in the areas of reading and math. These intervention teachers will monitor iReady Diagnostic Data as well as classroom assessments and daily student work. Intervention teachers will collaborate and plan with regular education teachers as well as serving students in small pull-out groups. Middle school intervention teachers and regular education teachers will utilize the following resources to ensure student success.

- iReady Diagnostic Assessment
- iReady Online Prescriptive Instruction for Reading
- IXL Math
- Various supplemental resources and manipulatives

Intervention teachers at the high school level will focus on attainment of credits not earned during the pandemic. These teachers will utilize a credit recovery type system. Students will be assigned to a credit recovery class and will work through assigned courses via OdysseyWare. Intervention teachers will monitor student progress and assist in re-teaching specific content as necessary. The high school intervention teachers will also collaborate with high school teachers when students struggle with specific content areas. The main instructional tool utilized for credit recovery will be OdysseyWare.

ARP ESSER funds will be utilized to fund seven teachers throughout the school district. These teachers will help class size reduction and provide continuity of services during the next three years. These teachers have been placed at specific schools throughout the district based upon need to reduce class size as well as need for specific courses.

ARP ESSER funds will be utilized to fund summer school sessions. Funds will be utilized to cover the cost of salary and benefits for certified teachers and transportation costs (bus drivers and diesel fuel). Summer school sessions will be offered at all schools throughout the Ohio County School District. Teachers will focus on learning loss in the areas of reading and math at the elementary and middle school levels. The high school will focus on attainment of credits in all content areas.

ARP ESSER funds will be utilized to purchase various supplemental instructional resources for students and teachers throughout the school district.

- iReady Resources (Reading & Math)
- Best practice literacy resources (Heinemann...Interactive Read Aloud Collections and Shared Reading Materials)
- Credit Recovery Resources OdysseyWare

Ohio County Schools will spend the remainder of the ARP ESSER funds as follows:

- One-time supplemental pay for full time employees to ensure continuity of services at risk by labor shortage, rising competitive wages, and significantly increased turnover in response to COVID
- Pre-employment drug screening (increased cost due to significantly increased turnover and additional staffing to prepare and respond to COVID, continuity of services)
- Medical services training for staff-COVID prevention, preparation and response

- Insurance required for continuity of services and in response to increased exposure risk due to COVID
- Postage for NTI packets and parent/student information (response to COVID)
- Cell phone service – staff communication both internally and externally with parents and community to prevent, prepare, and respond to COVID, necessary for social distancing
- Purchase school buses
- OneCall license for remote communication with parents (response to COVID) and Ky Virtual Library license.

The Ohio County School District conducted meaningful consultation with a wide variety of stakeholders regarding the use of ARP ESSER funds and designing the district spending plan. Consultation took place via Google Meets, in-person meetings, district staff attending a wide variety of SBDM meetings and parent groups meetings, student interviews, superintendent participating in local radio show to answer questions and share information, teacher meetings, phone calls and in-person meetings with local KEA president. Consultation was conducted with the following groups:

- All school administrators
- SBDM Councils
- Student Groups
- Parent Groups (PTO)
- Teacher Groups (i.e. OCEA/Ohio County Education Association/KEA local affiliate)
- Family Resource Coordinators & related Advisory Councils
- School Board Members
- Elementary Literacy Coaches
- Community Members
- County Judge Executive and Fiscal Court Members
- Ohio County Hospital

EVIDENCE BASED PRACTICES AND RELATED RESEARCH

(ATTACHED AS FOLLOWS)

1. The Rise Framework (The Next Step Forward In Reading Intervention) By: Jan Richardson & Ellen Lewis *Reading intervention framework
2. Reading Recovery (What Works Clearinghouse Information) Reading Recovery strategies and assessments
3. iReady Instructional Resources (iReady Diagnostic Reading and Math, iReady Online Prescriptives Instruction for Reading and Math, Ready Reading & Ready Math)
4. OdysseyWare Credit Recovery System

CHAPTER 2

What Is RISE?

RISE is a tested and proven intensive literacy intervention based on Jan's guided reading lesson framework and implemented 45 to 60 minutes a day for six to eight weeks. It is designed for children reading at Levels C–N who need to improve decoding, spelling, fluency, writing, and retelling. Our research based on field-testing showed students who participated in six to eight weeks of RISE made 6.4 months of progress and showed significant improvement in comprehension, as measured on *Next Step Guided Reading Assessment* (Richardson & Walther, 2013) and *Benchmark Assessment System* (Fountas & Pinnell, 2016). For students who read above Level N and only need to improve their comprehension, use the RISE Up procedures described in Chapters 4 and 5.

There is no reading without comprehension; we want students to understand that reading is a process of constructing meaning from text. While they are reading, they should always ask themselves, "Does this make sense?" If it doesn't, we want them to have a variety of strategies to construct meaning. RISE and RISE Up marshal the full force of language—reading, writing, speaking, and listening. In every intervention session, students are supported by all four language processes.

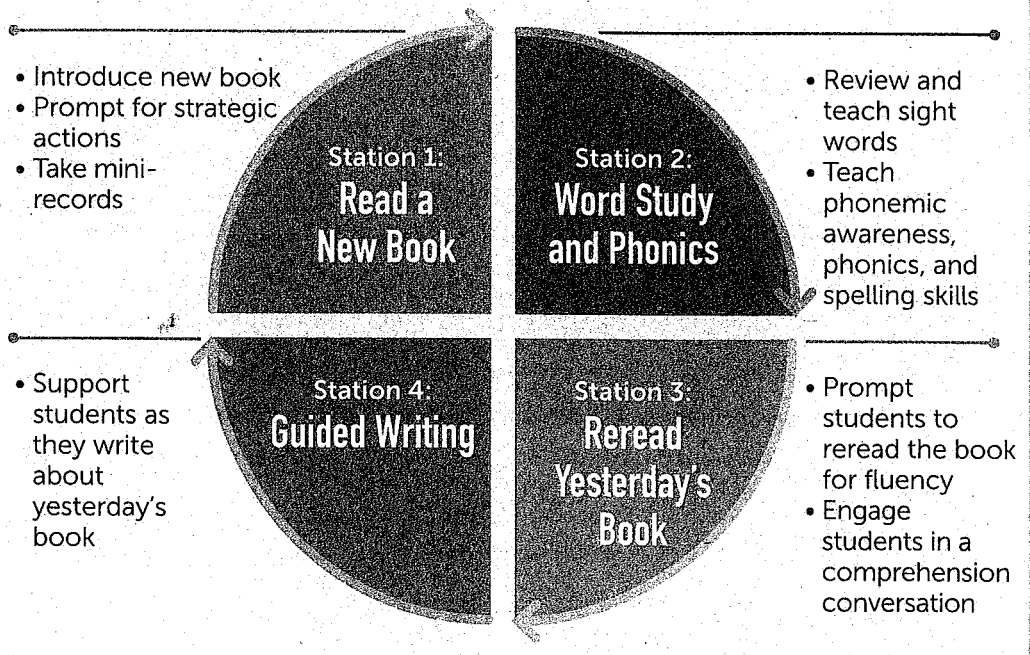
HOW DOES RISE WORK?

The literacy team and the classroom teachers use assessments to select up to 16 students who need intervention and who read at about the same text level. Those students are placed in smaller groups of three or four and rotate through four instructional stations. Each station targets one of Jan's lesson components: reading a new book, word study, rereading and discussing the book, and guided writing.

The RISE leader should be a credentialed teacher—ideally a reading specialist or Reading Recovery® teacher who can train the other members of the team, monitor

student progress, assist with lesson planning, and communicate with classroom teachers. The other RISE instructors can be special-education teachers, Title I teachers, reading interventionists, ELL teachers, literacy coaches, retired teachers, student teachers, teaching assistants, or other adults who routinely work with students in your school.

RISE: Reading Inspires Students to Excel
For children who read at Levels C–N



(45- to 60-minute intervention; 12 to 15 minutes at each station)

Although RISE requires four stations, it can be taught by 1 to 4 instructors. See pages 79–83 for implementation models based on the number of instructors available.

WWC Intervention Report **U.S. DEPARTMENT OF EDUCATION**

What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

Beginning Reading **Updated July 2013***

Reading Recovery®

Report Contents

Program Description[1]

Reading Recovery® is a short-term intervention that provides one-on-one tutoring to first-grade students who are struggling in reading and writing. The supplementary program aims to promote literacy skills and foster the development of reading and writing strategies by tailoring individualized lessons to each student. Tutoring is delivered by trained *Reading Recovery®* teachers in daily 30 minute pull-out sessions over the course of 12–20 weeks.

Research[2]

The What Works Clearinghouse (WWC) identified three studies of *Reading Recovery®* that both fall within the scope of the Beginning Reading topic area and meet WWC evidence standards. All three studies meet standards without reservations. Together, these studies included 227 students in first grade in at least 14 states.

The WWC considers the extent of evidence for *Reading Recovery®* on the reading skills of beginning readers to be small for four outcome domains—alphabetics, reading fluency, comprehension, and general reading achievement. (See the Effectiveness Summary on p. 4 for further description of these domains.)

Effectiveness

Reading Recovery® was found to have positive effects on general reading achievement and potentially positive effects on alphabetics, reading fluency, and comprehension for beginning readers.

Table 1. Summary of findings[3]

		Improvement index *(percentile points)*				
Outcome domain	Rating of effectiveness	Average	Range	Number of studies	Number of students	Extent of evidence
Alphabetics	Potentially positive effects	+21	+9 to +42	2	148	Small
Reading fluency	Potentially positive effects	+46	+32 to +49	1	74	Small
Comprehension	Potentially positive effects	+14	+6 to +26	2	145	Small
General reading achievement	Positive effects	+27	+19 to +38	3	227	Small

Program Information

Background

Reading Recovery® was developed by Dr. Marie M. Clay at the University of Auckland, New Zealand, and is distributed through about 20 university training centers in the United States and supported by the *Reading Recovery®* Council of North America (RRCNA). Address: 500 West Wilson Bridge Road, Suite 250, Worthington, OH 43085-5218. Email: jjohnson@readingrecovery.org. Web: http://www.readingrecovery.org. Telephone: (614) 310-7323.

Program details

In *Reading Recovery®*, teachers tailor one-to-one tutoring lessons to accommodate each student's needs. Depending on these needs, teachers incorporate instruction in topics such as phonemic awareness, phonics, vocabulary, fluency, comprehension, writing, motivation, and oral language. Each lesson consists of reading familiar or novel stories, manipulating letters and words, and writing and assembling stories. Lessons are interactive between the teacher and student, with the teacher carefully monitoring each student's reading behavior. *Reading Recovery®* lessons are discontinued when students demonstrate the ability to consistently read at the average level for their grade—between weeks 12 and 20 of the program. Those who make progress but do not reach average classroom performance after 20 weeks are referred for further evaluation and a plan for future action. Teacher training includes a 1-year university-based training program and ongoing professional development.

Cost

Reading Recovery® is available on a nonprofit, no royalty basis. Because *Reading Recovery®* in the United States is a collaboration between universities and school districts, costs include tuition for initial training and continuing professional development. To establish a *Reading Recovery®* site—comprised of multiple schools in a district or group of districts—a teacher leader must first be trained. This start-up cost includes paying the teacher leader's salary, paying university tuition for the *Reading Recovery®* coursework, and covering the costs of books and materials. Each site must also equip a room with a one-way mirror and sound system to provide subsequent training for teachers.

Teacher leaders work at the site level and provide professional development to *Reading Recovery®* teachers. Ongoing costs include support for the teacher leader and a portion of the *Reading Recovery®* teachers' salaries and benefits. These specially trained *Reading Recovery®* teachers work part of the day in *Reading Recovery®* and the remaining part of the day in other capacities such as teaching small literacy groups or classrooms. According to the program developer, the average US *Reading Recovery®* teacher worked with eight *Reading Recovery®* students and approximately 40 additional students during the 2010–11 school year.

Other related ongoing costs include professional development for both teacher leaders and *Reading Recovery®* teachers, books and materials for lessons, student program materials, and data evaluation fees (which cover the cost of updating a site's roster of teachers and schools, data entry, plus ongoing phone and email support from the Help Desk for teacher leaders). The cost of program materials is approximately $100 per student served (calculated by the RRCNA as an average over the 5-year period from 2007–11). Sites pay an annual data evaluation fee of $350 a site plus $45 per *Reading Recovery®* teacher. Sites implementing the program also pay annual technical support fees, which vary by the university that provides the *Reading Recovery®* training.

Research Summary

The WWC identified 202 studies that investigated the effects of *Reading Recovery®* on the reading skills of beginning readers.

Table 2. Scope of reviewed research

Grade	1
Delivery method	Individual
Program type	Supplement

The WWC reviewed 79 of those studies against group design evidence standards. Three studies (Pinnell, DeFord, & Lyons, 1988; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; and Schwartz, 2005) are randomized controlled trials that meet WWC evidence standards without reservations. Those three studies are summarized in this report. Seventy-six studies do not meet WWC evidence standards.

The remaining 123 studies do not meet WWC eligibility screens for review in this topic area. Citations for all 202 studies are in the References section, which begins on p. 7.

Summary of studies meeting WWC evidence standards without reservations

Pinnell, DeFord, & Lyons (1988) examined the effect of *Reading Recovery®* on the reading skills of first-grade students in urban public schools in Columbus, Ohio who were designated as the lowest 20% of readers in their classroom. In the portion of this study that meets WWC evidence standards without reservations[4], students attending classrooms in which teachers had not previously been implementing the intervention were randomly assigned either to the *Reading Recovery®* intervention or to an alternative compensatory program focused on skills-oriented drill activities. Students in the intervention condition participated in individualized instruction for 30 minutes daily until they reached average levels for the class. Students who reached average levels received, on average, 67 daily lessons. The analysis sample included 74 students (37 in each condition). Outcomes were measured in the spring of first grade.

Pinnell et al. (1994) measured the effect of *Reading Recovery®* on the reading skills of first-grade students enrolled in geographically diverse school districts in Ohio. In the portion of the study that meets WWC evidence standards without reservations, low-achieving students within the same schools were randomly assigned either to the *Reading Recovery®* condition or to a comparison group in which they continued their regular reading program and existing federally-supported educational assistance services. Comparison group teachers were given the opportunity to select the materials to use with comparison group students; options included materials related to basic reading skills and vocabulary development. Students in the intervention condition read an average of five books per lesson and received an average of 33 minutes of daily individualized instruction. The analysis sample included eight schools with 31 students in the intervention condition and 48 students in the comparison condition. Outcomes were measured in February of first grade.

Schwartz (2005) examined the effect of *Reading Recovery®* on the reading skills of first-grade students attending elementary schools in 14 states. Within each participating school, teachers identified two students eligible for *Reading Recovery®*; these students were then randomly assigned to receive the program during the first or the second half of the school year. During the transition period between the first and second half of the school year, students assigned to receive the intervention during the first half of the year (intervention group) had finished the program (by either reaching classroom averages or attending the program for 20 weeks), and students assigned to receive the intervention in the second half of the year (comparison group) had not yet been exposed to *Reading Recovery®*. During this transition period, reading outcomes were measured for 74 students (37 in each condition).

Summary of studies meeting WWC evidence standards with reservations

No studies of *Reading Recovery®* met WWC evidence standards with reservations.

Effectiveness Summary

The WWC review of *Reading Recovery®* for the Beginning Reading topic area includes student outcomes in four domains: alphabetics, reading fluency, comprehension, and general reading achievement. The three studies of *Reading Recovery®* that meet WWC evidence standards reported findings in all four domains. Findings in the alphabetics domain for this review are differentiated by three constructs (as described in the Beginning Reading review protocol): phonemic awareness, letter knowledge, and phonics. Findings in the comprehension domain are differentiated by two constructs: reading comprehension and vocabulary development. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Reading Recovery®* on beginning readers. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 36.

Summary of effectiveness for the alphabetics domain

Two studies that meet WWC standards without reservations reported findings in the alphabetics domain.

One study examined the effect of *Reading Recovery®* on the phonemic awareness construct in the alphabetics domain. Schwartz (2005) reported no statistically significant differences for the phonemic awareness measures—the deletion task and the Yopp-Singer Test of Phoneme Segmentation—but the effects on both measures were positive and considered substantively important based on the WWC criteria (that is, at least 0.25).

Two studies examined the effect of *Reading Recovery®* on the letter knowledge construct in the alphabetics domain. Pinnell, DeFord, & Lyons (1988) did not find a statistically significant effect for *Reading Recovery®* on the Letter Identification subtest of the Observation Survey of Early Literacy Achievement, but the effect was positive and considered substantively important according to WWC criteria. Schwartz (2005) also reported a statistically insignificant effect of *Reading Recovery®* on the Letter Identification subtest of the Observation Survey; this difference was positive but not considered substantively important based on WWC criteria.

Two studies examined the effect of *Reading Recovery®* on the phonics construct in the alphabetics domain. Pinnell, DeFord, & Lyons (1988) found a statistically significant positive effect on the Word Recognition subtest of the Observation Survey. In WWC calculations, there was no statistically significant effect, but the positive effect was large enough to be considered substantively important. Schwartz (2005) found, and the WWC confirmed, a statistically significant positive effect of *Reading Recovery®* on the Word Recognition subtest of the Observation Survey.

The WWC characterizes student findings for Schwartz (2005) as a statistically significant positive effect because the average effect of the four outcomes (across constructs) is positive and statistically significant. Also, the effect on the Word Recognition subtest of the Observation Survey is positive and statistically significant, and no effects are negative and statistically significant for this study. For Pinnell, DeFord, & Lyons (1988), the average effect for the two outcome measures (across constructs) is not statistically significant but is considered to be substantively important based on WWC evidence criteria; therefore, the WWC characterizes these study findings as a substantively important positive effect.

Thus, for the alphabetics domain, among the two studies with a strong design, one showed a statistically significant positive effect and one showed a substantively important positive effect. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the alphabetics domain

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *alphabetics* domain was a statistically significant positive effect in one study and a substantively important positive effect in one study.
Extent of evidence	**Criteria met**
Small	Two studies that included 148 students reported evidence of effectiveness in the *alphabetics* domain.

Summary of effectiveness for the reading fluency domain

One study that meets WWC standards without reservations reported findings in the reading fluency domain.

Schwartz (2005) found, and the WWC confirmed, positive and statistically significant effects of *Reading Recovery®* on the Slosson Oral Reading Test–Revised and the Text Reading Level subtest of the Observation Survey of Early Literacy Achievement.

Thus, for the reading fluency domain, one study with a strong design showed a statistically significant positive effect. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 4. Rating of effectiveness and extent of evidence for the reading fluency domain

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *reading fluency* domain was a statistically significant positive effect.
Extent of evidence	**Criteria met**
Small	One study that included 74 students reported evidence of effectiveness in the *reading fluency* domain.

Summary of effectiveness for the comprehension domain

Two studies that meet WWC standards without reservations reported findings in the comprehension domain.

Two studies examined the effect of *Reading Recovery®* on the reading comprehension construct in the comprehension domain. Pinnell, DeFord, & Lyons (1988) reported, and the WWC confirmed, a substantively important (but statistically insignificant) positive effect on the Reading Comprehension subtest of the Comprehensive Test of Basic Skills (CTBS). Schwartz (2005) reported neither a statistically significant nor a substantively important effect of *Reading Recovery®* on the Degrees of Reading Power Test.

One study examined the effect of *Reading Recovery®* on the vocabulary development construct in the comprehension domain. Pinnell, DeFord, & Lyons (1988) found, and the WWC confirmed, a positive and statistically significant effect of *Reading Recovery®* on the Reading Vocabulary subtest of the CTBS.

Thus, for the comprehension domain, one study with a strong design showed a statistically significant positive effect, and one study with a strong design showed an indeterminate effect. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 5. Rating of effectiveness and extent of evidence for the comprehension domain

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *comprehension* domain was a statistically significant positive effect in one study and an indeterminate effect in one study.
Extent of evidence	**Criteria met**
Small	Two studies that included 145 students reported evidence of effectiveness in the *comprehension* domain.

Summary of effectiveness for the general reading achievement domain

Three studies reported findings in the general reading achievement domain.

Pinnell, DeFord, & Lyons (1988) found, and the WWC confirmed, positive and statistically significant effects of *Reading Recovery®* on three subtests of the Observation Survey of Early Literacy Achievement: Concepts About Print, Hearing and Recording Sounds in Words (Dictation), and Writing Vocabulary.

Pinnell et al. (1994) found, and the WWC confirmed, statistically significant positive effects of *Reading Recovery®* on the Gates-MacGinitie, the Dictation subtest of the Observation Survey, and the Woodcock Reading Mastery Test–Revised.

Schwartz (2005) found, and the WWC confirmed, positive and statistically significant effects of *Reading Recovery®* on three subtests of the Observation Survey: Concepts About Print, Dictation, and Writing Vocabulary.

Thus, for the general reading achievement domain, three studies with strong designs reported statistically significant positive effects. This results in a rating of positive effects, with a small extent of evidence.

Table 6. Rating of effectiveness and extent of evidence for the general reading achievement domain

Rating of effectiveness	Criteria met
Positive effects *Strong evidence of a positive effect with no overriding contrary evidence.*	In the three studies that reported findings, the estimated impact of the intervention on outcomes in the *general reading achievement* domain was a statistically significant positive effect.
Extent of evidence	**Criteria met**
Small	Three studies that included 227 students reported evidence of effectiveness in the *general reading achievement* domain.

Appendix A.1: Research details for Pinnell, DeFord, & Lyons (1988)

Pinnell, G. S., DeFord, D. E., & Lyons, C. A. (1988). *Reading Recovery: Early intervention for at-risk first graders* (Educational Research Service Monograph). Arlington, VA: Educational Research Service.

Table A1. Summary of findings **Meets WWC evidence standards without reservations**

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Alphabetics	74 students	+18	No
Comprehension	71 students	+22	Yes
General reading achievement	74 students	+24	Yes

Setting The study took place in 12 urban public schools in Columbus, Ohio.

Study sample The study authors used several comparison groups to examine the effectiveness of the *Reading Recovery®* program. The study comparison that meets WWC evidence standards includes students attending classrooms of teachers who had not previously been trained in *Reading Recovery®*. Eligible first-grade students were designated as the lowest 20% of readers in their classroom, based on the scores on the Observation Survey of Early Literacy Achievement, teacher judgment, and a standardized test. Thirty-eight students were randomly assigned to participate in the *Reading Recovery®* program, and 37 students were randomly assigned to the comparison group. The analysis sample after sample attrition included 37 students in both the intervention and comparison groups.

Intervention group Students in the *Reading Recovery®* group attended regular education classes. Each student also participated in individualized instruction with a *Reading Recovery®* teacher for 30 minutes daily until the student reached average levels for the class (on average, students who reached average levels received 67 daily lessons).

Comparison group Students in the comparison group attended regular education classes. They also attended an alternative compensatory program focused on a series of skills-oriented drill activities. This program included primarily small group instruction (with minimal individual-level instruction) and was delivered by trained paraprofessionals for approximately 30–45 minutes per day.

Outcomes and measurement Researchers reported outcomes from nine literacy measures, seven of which were included in the WWC review and ratings of effectiveness. Five of the six reported subtests of the Observation Survey[5] were included in the WWC review of this study: two in the alphabetics domain, including Letter Identification and Word Recognition; and three in the general reading achievement domain, including Concepts About Print, Dictation, and Writing Vocabulary. Results from the Observation Survey: Text Reading Level subtest were not reported in this review because the WWC determined that it was not possible to calculate effect sizes that were comparable to other measures. The study authors also reported two outcome measures that fall into the comprehension domain: the Reading Vocabulary subtest and the Reading Comprehension subtest of the Comprehensive Test of Basic Skills (CTBS). Finally, the study included a writing assessment that does not fall within one of the domains specified in the WWC Beginning Reading protocol. For a more detailed description of the included outcome measures, see Appendix B.

Support for implementation

Reading Recovery® teachers received a full year of special training, during which they practiced teaching using *Reading Recovery®* methods and observed other teachers through a one-way mirror. The 20 teachers who provided the *Reading Recovery®* intervention to the analysis sample included in this WWC review received training from a local teacher leader and were in their first year of teaching the intervention during the time of the study.[6]

Appendix A.2: Research details for Pinnell et al. (1994)

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly, 29*(1), 8–39.

Table A2. Summary of findings **Meets WWC evidence standards without reservations**

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
General reading achievement	79 students	+21	Yes

Setting

The study took place in ten school districts (two rural, two suburban, and six urban) in Ohio.

Study sample

The authors studied 403 first-grade students distributed across 43 schools from ten districts. The percentage of students in each district who received public assistance in the form of Aid to Dependent Children ranged from 9% and 42%. Four schools per district implemented one of four reading interventions—*Reading Recovery®*, *Reading Success*, *Direct Instruction Skills Plan*, and *Reading and Writing Group*. Within each school, the ten lowest-scoring students were randomly assigned either to participate in the intervention or to participate in the school's regular reading program. For this report, the WWC looked at results for students in the ten schools (across ten school districts) who were using *Reading Recovery®* as their intervention. These schools all had prior experience implementing *Reading Recovery®*. In the original study design, 100 students were randomly assigned to receive either *Reading Recovery®* or the comparison condition at ten schools. However, random assignment was not successfully implemented at two schools, and there was minor attrition at the remaining schools, resulting in a final analytic sample of 79 students from eight schools (in eight districts). All students were low achieving, which was defined as students who scored below the 37th percentile on a standardized assessment and who were recommended for compensatory help by their teachers.

Intervention group

The intervention group was composed of 31 low-achieving students across eight schools. Intervention students received one-on-one tutoring with a trained *Reading Recovery®* teacher daily for 30 minutes. The activities led by the teacher were aimed at fostering independent reading skills and included: reading both easier and more challenging books, conducting word analysis in context, and participating in activities aimed at improving writing fluency, such as composing sentences and reconstructing cut-up versions of sentences.

Comparison group

The comparison group included 48 students attending the same eight schools as the intervention group. Students assigned to the comparison group received no special instruction, but continued to participate in their regular reading program and existing federally-funded supplemental education services with an instructional focus on developing basic reading and vocabulary skills. Some lessons from the supplemental education program included teachers reading aloud as well as group reading. Comparison group teachers, none of whom had received *Reading Recovery®* training, selected instructional materials based on their own discretion.

Outcomes and measurement

This WWC review focuses on outcomes measured in February of the academic year in which the study took place because, at that point, no comparison group students had been exposed to the intervention. The WWC review does not include assessments that were measured in May of the same academic year because, at that time, a portion of students who had originally been assigned to the comparison condition had participated in the intervention. Three measures were administered to assess student outcomes in the general reading achievement domain: the Dictation subtest of the Observation Survey of Early Literacy Achievement, the Woodcock Reading Mastery Test–Revised, and the Gates-MacGinitie Reading Test. Results from the Observation Survey: Text Reading Level subtest were not reported because effect sizes that were comparable to other measures could not be calculated. For a more detailed description of the included outcome measures, see Appendix B.

Support for implementation

At least two years prior to the study, *Reading Recovery®* teachers received specialized training. During this training period that took place over the course of an academic year, the teachers participated in weekly 2.5 hour long sessions, in which they practiced teaching using *Reading Recovery®* methods and observed other teachers through a one-way mirror. They also received a 1-day orientation at the beginning of the study.

Appendix A.3: Research details for Schwartz (2005)

Schwartz, R. M. (2005). Literacy learning of at-risk first-grade students in the Reading Recovery early intervention. *Journal of Educational Psychology, 97*(2), 257–267.

Table A3. Summary of findings **Meets WWC evidence standards without reservations**

		Study findings	
Outcome domain	**Sample size**	**Average improvement index *(percentile points)***	**Statistically significant**
Alphabetics	74 students	+23	Yes
Reading fluency	74 students	+46	Yes
Comprehension	74 students	+6	No
General reading achievement	74 students	+35	Yes

Setting

The study took place in an unspecified number of elementary schools in 14 states.

Study sample

The study was designed to examine the effect of *Reading Recovery®* on the outcomes of first-grade students. Forty-seven *Reading Recovery®* teachers each identified two students[7] eligible for *Reading Recovery®* based on their low scores on the Observation Survey of Early Literacy Achievement and their own judgment. These 94 students were randomly assigned to enter the *Reading Recovery®* program during either the first or second half of the school year. [Note: The study also included two additional comparison groups of 47 low-average and 47 high-average readers from the same classrooms as the *Reading Recovery®* students who were not expected to participate in the *Reading Recovery®* program. Analysis involving these comparison groups was not eligible for WWC review because the WWC considers only comparisons of students with similar achievement backgrounds in assessing the effectiveness of an intervention.] Because of missing test data, the author's final analytic sample included 74 students distributed across 37 teachers.

Intervention group

Students participated in the one-on-one daily 30-minute tutoring program for up to 20 weeks or until they were judged by their teacher to have met the criteria for termination of the program by reaching average levels of literacy performance. The length of program participation ranged from 12 to 20 weeks. Originally, participants were taught by 47 *Reading Recovery®* teachers who had volunteered to be part of the study, but because of missing test data, data from only 37 teachers and 37 students were included in the author's final analysis. The intervention group was 61% male, 47% Black, 38% White, 12% Hispanic, and 3% Asian. About 60% of the group received free or reduced-price lunch.

Comparison group

The comparison group included students who were randomly assigned to receive *Reading Recovery®* during the second half of the year. Thus, these participants served as a comparison group only during the first part of the year when they received instruction in their regular classroom but no additional supplemental services. The final analysis included data from 37 teachers and 37 students. The comparison group was 41% male, 47% White, 38% Black, and 15% Hispanic. Approximately 57% of the group received free or reduced-price lunch.

Outcomes and measurement

The study author reported outcomes on ten literacy measures, all of which were included in the WWC review and ratings of effectiveness. Six reported subtests of the Observation Survey were included in the WWC review of this study: two in the alphabetics domain, including Letter Identification and Word Recognition; one in the fluency domain (Text Reading Level); and three in the general reading achievement domain, including Concepts About Print, Dictation, and Writing Vocabulary. The study author also reported two additional outcome measures that fall into the alphabetics domain, Phoneme Segmentation and Deletion task, one additional outcome in the fluency domain, Slosson Oral Reading Test–Revised, and one outcome in the comprehension domain, Degrees of Reading Power. For a more detailed description of the included outcome measures, see Appendix B.

Support for implementation

Although the study provided no information about training provided to participating teachers, *Reading Recovery®* teachers typically must complete a year-long certification program.

Appendix B: Outcome measures for each domain

Alphabetics	
Phonemic awareness	
Deletion task	A ten-item version of the Rosner task, this assessment requires students to repeat a word and then say it again after omitting a given syllable or sound. The assessment is not standardized (as cited in Schwartz, 2005).
Yopp-Singer Test of Phoneme Segmentation	Developed by Hallie K. Yopp, the test is an orally administered assessment. A teacher works with each student individually and introduces the test as a word game. The teacher has a list of 22 words that the student is not allowed to see. After the teacher reads each word, the student must repeat all of the sounds in the word separately and slowly (as cited in Schwartz, 2005).
Letter knowledge	
Observation Survey of Early Literacy Achievement: Letter Identification subtest	Students identify upper- and lowercase letters. This assessment, developed by Dr. Marie M. Clay, is not standardized (as cited in Pinnell, DeFord, & Lyons, 1988; Schwartz, 2005).
Phonics	
Observation Survey of Early Literacy Achievement: Word Recognition subtest (also known as the Ready to Read or Ohio Word Test)	Students read 20 common sight words from basic reading texts, and their accuracy is scored. This assessment, developed by Dr. Marie M. Clay, is not standardized (as cited in Pinnell, DeFord, & Lyons, 1988; Schwartz, 2005).
Reading fluency	
Observation Survey of Early Literacy Achievement: Text Reading Level subtest[a]	This subtest measures the percentage of students scoring at the first-grade reading level or higher compared with those scoring lower than first grade. To determine this, students read from passages of increasing difficulty, and each student's error rate and self-correcting behavior are recorded using the running record technique. Students read from leveled texts drawn from a basal reading series until their accuracy rate falls below 90%. Results are translated to a numerical reading level from level one to level 30, which in turn match up to grade-level equivalency. This assessment is not standardized (as cited in Schwartz, 2005).
Slosson Oral Reading Test–Revised (SORT-R3)	Developed by Richard L. Slosson and Charles L. Nicholson, this measure consists of 200 words arranged in order of difficulty, with 20 words per list. Each list represents an approximate reading grade level (for example, list one is equivalent to first grade). Administration ends after all the words on one list are mispronounced. The measure is standardized and norm-referenced (as cited in Schwartz, 2005).
Comprehension	
Reading comprehension	
Comprehension Test of Basic Skills (CTBS): Reading Comprehension subtest	This subtest is a group-administered, standardized assessment of reading comprehension (as cited in Pinnell, DeFord, & Lyons, 1988).
Degrees of Reading Power Test	This test is an untimed standardized assessment requiring students to read a nonfiction passage with a word or set of words missing. Students select an appropriate answer to complete the sentence from a set of four or five alternatives (as cited in Schwartz, 2005).
Vocabulary development	
CTBS: Reading Vocabulary subtest	A group-administered, standardized assessment of vocabulary (as cited in Pinnell, DeFord, & Lyons, 1988).
General reading achievement	
Gates-MacGinitie Reading Test (1978)	A standardized test, this assessment covers vocabulary and comprehension aspects of reading. It evaluates students' abilities to decode initial consonants, consonant clusters, final consonants, and vowels in real English words and also measures their ability to recognize commonly used words without decoding. For reading comprehension, answer choices are given as pictures and words (as cited in Pinnell et al., 1994).
Observation Survey of Early Literacy Achievement: Concepts About Print subtest	Students perform tasks related to printed language concepts (for example, directionality, book handling, and word concepts) while reading a book. This assessment, developed by Dr. Marie M. Clay, is not standardized (as cited in Pinnell, DeFord, & Lyons, 1988; Schwartz, 2005).

Observation Survey of Early Literacy Achievement: Hearing and Recording Sounds in Words (Dictation) subtest	For this subtest, students write the words that are dictated to them in sentence form. This assessment, developed by Dr. Marie M. Clay, is not standardized (as cited in Pinnell, DeFord, & Lyons, 1988; Pinnell et al., 1994; Schwartz, 2005).
Observation Survey of Early Literacy Achievement: Writing Vocabulary subtest	For this subtest, students are given ten minutes to write as many words as they can on a blank sheet of paper. If needed, a standard set of prompts is used to encourage additional attempts to write. The measure is scored by counting the number of correctly spelled words (as cited in Pinnell, DeFord, & Lyons, 1988; Schwartz, 2005).
Woodcock Reading Mastery Test–Revised	A standardized test composed of six subtests, this assessment measures the ability to form associations between visual stimuli and oral responses; ability to recognize upper- and lowercase letters in a variety of fonts; ability to read words aloud; ability to read aloud nonsense words or uncommon words to test phonic and structural analysis skills for pronouncing unfamiliar words; vocabulary ability through the use of antonyms, synonyms, and analogies; and passage comprehension by filling in missing words in a short paragraph (as cited in Pinnell et al., 1994).

[a] For Pinnell et al. (1988) and Pinnell et al. (1994), findings based on the Observation Survey of Early Literacy Achievement: Text Reading Level subtest are not included in the effectiveness ratings because effect sizes and the statistical significance of the findings could not be calculated given the information provided in the studies. The Text Reading Level subtest is reported as reading levels based on ordinal, rather than equal-interval, scales. For example, the increase in fluency measured by scoring at level 3 compared with level 2 on the scale may not be equal to the increase in fluency as measured by scoring at level 24 compared with level 23. The authors no longer had information on the number of students scoring at each level. For more detail, see Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2006). Validity, reliability, and utility of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly, 41*(1) 8–34.

Appendix C.1: Findings included in the rating for the alphabetics domain

Outcome measure	Study sample	Sample size	Mean (standard deviation) Intervention group	Mean (standard deviation) Comparison group	WWC calculations Mean difference	WWC calculations Effect size	WWC calculations Improvement index	*p*-value
Construct: Phonemic awareness								
Schwartz, 2005[a]								
Deletion task	Grade 1	74 students	6.64 (2.56)	5.58 (2.50)	1.06	0.41	+16	> 0.05
Yopp-Singer Test of Phoneme Segmentation	Grade 1	74 students	17.70 (4.93)	15.27 (5.43)	2.43	0.46	+18	> 0.05
Construct: Letter knowledge								
Pinnell, DeFord, & Lyons, 1988[b]								
Observation Survey of Early Literacy Achievement: Letter Identification subtest	Grade 1	74 students	52.27 (1.41)	51.19 (3.17)	1.08	0.44	+17	0.06
Schwartz, 2005[a]								
Observation Survey of Early Literacy Achievement: Letter Identification subtest	Grade 1	74 students	52.18 (1.27)	51.68 (2.78)	0.50	0.23	+9	> 0.05
Construct: Phonics								
Pinnell, DeFord, & Lyons, 1988[b]								
Observation Survey of Early Literacy Achievement: Word Recognition subtest	Grade 1	74 students	13.68 (1.63)	12.51 (2.87)	1.17	0.50	+19	0.04
Schwartz, 2005[a]								
Observation Survey of Early Literacy Achievement: Word Recognition subtest	Grade 1	74 students	14.96 (3.99)	8.87 (4.75)	6.09	1.37	+42	< 0.01
Domain average for alphabetics (Pinnell, DeFord, & Lyons, 1988)						**0.47**	**+18**	**Not statistically significant**
Domain average for alphabetics (Schwartz, 2005)						**0.62**	**+23**	**Statistically significant**
Domain average for alphabetics across all studies						**0.55**	**+21**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. na = not applicable.

[a] For Schwartz (2005), no corrections for clustering or multiple comparisons were needed as the authors adjusted for multiple comparisons. The *p*-values presented here were reported in the original study. For the *Letter Identification* and *Word Recognition* outcomes, the WWC calculated the program group means using a difference-in-differences approach (see WWC Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group post-test means. Mean gains were not available for the two phonemic awareness outcomes, and thus, the WWC reports unadjusted posttest means for the intervention group. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant.

[b] For Pinnell, DeFord, & Lyons (1988), a correction for multiple comparisons was needed and resulted in a WWC-computed critical *p*-value of 0.025 for the *Word Recognition* test; therefore, the WWC does not find the individual results to be statistically significant. However, this study is characterized as having a substantively important positive effect because the mean effect size for the measures of outcomes in the domain is positive and greater than 0.25. For more information, please refer to the WWC Standards and Procedures Handbook, version 2.1, p. 96.

Appendix C.2: Findings included in the rating for the reading fluency domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Schwartz, 2005[a]								
Slosson Oral Reading Test–Revised	Grade 1	74 students	30.58 (14.41)	18.12 (11.87)	12.46	0.93	+32	< 0.01
Observation Survey of Early Literacy Achievement: Text Reading subtest	Grade 1	74 students	0.78	0.05	0.73	2.49	+49	< 0.01
Domain average for reading fluency (Schwartz, 2005)						**1.71**	**+46**	**Statistically significant**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC.

[a] For Schwartz (2005), no corrections for clustering or multiple comparisons were needed as the authors adjusted for multiple comparisons. The *p*-values presented here were reported in the original study. Means presented for the *Text Reading* subtest are the posttest proportions for each group scoring at or above a first-grade reading level (provided by the study author). Effect size is computed as a Cox's index: logged-odds ratio transformation divided by 1.65. See the WWC Handbook, Version 2.1 for the computation of effect sizes for binary outcomes. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant.

Appendix C.3: Findings included in the rating for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Construct: Reading comprehension								
Pinnell, DeFord, & Lyons, 1988[a]								
Comprehension Test of Basic Skills (CTBS) Reading Comprehension subtest	Grade 1	70 students	36.67 (19.27)	28.88 (14.53)	7.79	0.45	+17	0.06
Schwartz, 2005[b]								
Degrees of Reading Power Test	Grade 1	74 students	4.82 (3.88)	4.27 (3.88)	0.55	0.14	+6	> 0.05
Construct: Vocabulary development								
Pinnell, DeFord, & Lyons, 1988[a]								
CTBS Reading Vocabulary subtest	Grade 1	71 students	36.64 (11.93)	26.11 (16.86)	10.53	0.71	+26	< 0.01
Domain average for comprehension (Pinnell, DeFord, & Lyons, 1988)						**0.58**	**+22**	**Statistically significant**
Domain average for comprehension (Schwartz, 2005)						**0.14**	**+6**	**Not statistically significant**
Domain average for comprehension across all studies						**0.36**	**+14**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. na = not applicable.

[a] For Pinnell, DeFord, & Lyons (1988), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were calculated from t-statistics reported in the original study. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant.

[b] For Schwartz (2005), no correction for clustering or multiple comparisons were needed as the authors adjusted for multiple comparisons. The *p*-values presented here were reported in the original study. This study is characterized as having an indeterminate effect because the single effect is neither statistically significant nor substantively important.

Appendix C.4: Findings included in the rating for the general reading achievement domain

			Mean (standard deviation)		WWC calculations			
Outcome measure	Study sample	Sample size	Intervention group	Comparison group	Mean difference	Effect size	Improvement index	*p*-value
Pinnell, DeFord, & Lyons, 1988[a]								
Observation Survey of Early Literacy Achievement: Concepts About Print subtest	Grade 1	74 students	15.81 (2.91)	14.30 (3.08)	1.51	0.50	+19	0.04
Observation Survey of Early Literacy Achievement: Dictation subtest	Grade 1	74 students	30.62 (6.13)	24.38 (6.92)	6.24	0.94	+33	< 0.01
Observation Survey of Early Literacy Achievement: Writing Vocabulary subtest	Grade 1	74 students	32.86 (13.49)	26.05 (14.32)	6.81	0.48	+19	0.04
Pinnell et al., 1994[b]								
Gates-MacGinitie Reading Test (1978)	Grade 1	79 students	36.19 (13.12)	31.00 (nr)	5.19	0.51	+19	< 0.05
Observation Survey of Early Literacy Achievement: Dictation subtest	Grade 1	79 students	31.74 (6.18)	26.75 (nr)	4.99	0.65	+24	0.01
Woodcock Reading Mastery Test–Revised	Grade 1	79 students	39.81 (21.35)	39.49 (nr)	0.32	0.49	+19	< 0.05
Schwartz, 2005[c]								
Observation Survey of Early Literacy Achievement: Concepts About Print subtest	Grade 1	74 students	19.24 (2.55)	16.68 (2.30)	2.56	1.04	+35	< 0.01
Observation Survey of Early Literacy Achievement: Dictation subtest	Grade 1	74 students	35.58 (2.70)	29.08 (7.37)	6.50	1.16	+38	< 0.01
Observation Survey of Early Literacy Achievement: Writing Vocabulary subtest	Grade 1	74 students	42.67 (11.42)	31.00 (12.94)	11.67	0.95	+33	< 0.01
Domain average for general reading achievement (Pinnell, DeFord, & Lyons, 1988)						**0.64**	**+24**	**Statistically significant**
Domain average for general reading achievement (Pinnell et al., 1994)						**0.55**	**+21**	**Statistically significant**

Domain average for general reading achievement (Schwartz, 2005)	**1.05**	**+35**	**Statistically significant**
Domain average for general reading achievement across all studies	**0.75**	**+27**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. nr = not reported. na = not applicable.

[a] For Pinnell, DeFord, & Lyons (1988), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were calculated from *t*-statistics reported in the original study. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant.

[b] For Pinnell et al. (1994), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were reported in the original study. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant.

[c] For Schwartz (1994), no corrections for clustering or multiple comparisons were needed as the authors adjusted for multiple comparisons. The *p*-values presented here were reported in the original study. The WWC calculated the program group means using a difference-in-differences approach (see WWC Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. This study is characterized as having a statistically significant positive effect because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. For more information, please refer to the WWC Standards and Procedures Handbook, version 2.1, p. 96.

Endnotes

* On September 16, 2013, the WWC modified this report in response to an independent review by the quality review team. Based on the review, the WWC changed the disposition for Baenen et al. (1997) from "The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent" to "The study does not meet WWC evidence standards because it is a randomized controlled trial in which the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not shown to be equivalent." Endnote 2 was revised to improve clarity; there were no changes to the substantive information. The WWC has not added studies to the body of evidence or updated the literature search since the July 2013 release of this report.

[1] The descriptive information for this program was obtained from a publicly available source: the program's website (http://www.readingrecovery.org; downloaded December 2011). The WWC requests developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in March 2012, and we incorporated feedback from the developer. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.

[2] The literature search reflects documents publicly available by December 2012. This report has been updated to include reviews of 96 studies that have been reviewed since the previous intervention report was released in December 2008. The additional 96 studies were not within the scope of the review protocol for the Beginning Reading topic area or were within the scope of the review protocol but did not meet evidence standards. In addition, two studies (Iverson and Tunmer, 1993 and Baenen et al., 1997), which met WWC evidence standards with and without reservations, respectively, in the previous report, do not meet WWC evidence standards with or without reservations in this report. These revised dispositions are due to changes in the review standards and the Beginning Reading review protocol. In particular, for Iverson and Tunmer (1993), in the version 1.0 standards, a statistical adjustment for baseline differences was sufficient to demonstrate equivalence in quasi-experimental studies; in the protocol version 2.1 standards, if differences are too great at baseline (greater than 25% of the pooled standard deviation), then the study cannot meet standards (even after a statistical adjustment). For Baenen et al. (1997), the grade retention outcome, initially reported in the December 2008 report, was determined to be ineligible for review due to revisions to the Beginning Reading review protocol (version 2.1). The RCT analysis included a second outcome, the North Carolina End-Of-Grade Reading test, that had a combination of overall and differential attrition rates that exceeded the WWC standards, and the subsequent analytic intervention and comparison groups were not shown to be equivalent. A complete list of all studies reviewed and their dispositions are provided in the references. The studies in this report were reviewed using the Evidence Standards from the WWC Procedures and Standards Handbook, version 2.1, as described in the Beginning Reading protocol (version 2.1). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 36. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[4] In the WWC *Reading Recovery®* intervention report that was published in 2008, the WWC review focused on a slightly different comparison sample of 51 students. During the revised WWC review conducted for this report, it was determined, based on the published documents combined with information obtained through an author query, that the most appropriate random assignment comparison is based on the comparison group of 37 students as reported in Pinnell et al. (1986). Since it is not entirely clear whether the remaining 14 students were randomly assigned in the same manner, the WWC assessed whether the group of 37 *Reading Recovery®* students was equivalent to the larger comparison group of 51 students on pretest scores. These groups were not deemed to be equivalent, and thus, this comparison does not meet WWC evidence standards. Similarly, a second group of students determined to be eligible for *Reading Recovery®* received the standard *Reading Recovery®* pull-out program, with the addition of having regular classroom teachers trained in *Reading Recovery®* (n = 96). The second group was neither randomly assigned to *Reading Recovery®* nor randomly assigned to their classroom teacher, so this portion of the study is considered a quasi-experimental design. It is not included in the intervention rating because the second intervention group with a trained *Reading Recovery®* teacher as a regular classroom teacher goes beyond the standard implementation of the program. Also, this comparison does not meet WWC evidence standards due to lack of statistical adjustment for differences in pretest reading scores as required by the WWC.

[5] The Observation Survey of Early Literacy Achievement was developed by Dr. Marie M. Clay, who also developed *Reading Recovery®*. Although there is no evidence of obvious overalignment between the measure and the intervention (intervention students receiving exposure to the measure during the course of the intervention), it should be noted that the developer of the intervention and the measure were the same.

[6] Twelve teachers received training from a university program and were in their second year of teaching the intervention during the time of the study. These teachers provided the program to students in the non-random assignment portion of the study that did not meet WWC evidence standards.

[7] The teachers initially identified five students. The lowest three students in the class automatically received *Reading Recovery®*, and the remaining two were randomly assigned.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, July). *Beginning Reading intervention report: Reading Recovery®*. Retrieved from http://whatworks.ed.gov

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 36.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 36.

Single-case design A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.

i-Ready

What We've Learned about Unfinished Learning

Insights from Midyear Diagnostic Assessments

Curriculum Associates Research Brief | March 2021

Curriculum Associates

Executive Summary

In marking one year of interrupted schooling due to the COVID-19 pandemic, educators, community members, and policymakers across the country are continuing to ask questions about student learning while reflecting on what teachers and students have achieved under unprecedented circumstances. Questions about mitigating the lost time for teaching and learning and addressing student academic needs highlight the role of language we use and its implications. We recognize that the interruption in schooling happened due to circumstances outside of the school and the classroom, and teaching and learning remains unfinished rather than lost. To reflect that, when we describe where students are not yet prepared for grade-level work, we will use the term *unfinished teaching and learning* or *unfinished learning* instead of learning loss. When we describe where students are on grade level, we will use the terms *ready for grade-level work* or *not ready for grade-level work* instead of performing on grade level or not performing on grade level. For more thoughts on asset-based language, please see *The Language We Use to Describe Learning* on page IV.

Using the *i-Ready Diagnostic*'s criterion-referenced grade-level placement data from more than nine million students, we compared student achievement during winter 2020–2021 to what we would expect during a typical school year. The *i-Ready Diagnostic* assessment asks students to indicate whether they are taking the test in school or out of school, which allows us to look at data trends by testing location. After initially looking at student assessment data from tests taken in school and those tests taken out of school, we came to the conclusion that examining assessment data taken exclusively in school is the closest to a "true" comparison to prior-year achievement. As such, this report discusses findings from in-school testing only.

Our findings shed light on the severity of unfinished learning in 2021. The winter assessment data indicates there are fewer students ready to access grade-level work and more students underprepared to access grade-level work, compared with historical benchmarks—validating educators' concerns about unfinished learning. The unfinished learning is greater for students in schools serving a majority of Black and Latino students, compared to a majority of White students, and unfinished learning is most stark for students underprepared to access grade-level work. Students attending schools in lower-income zip codes are also experiencing greater unfinished learning than students attending schools in higher-income zip codes. The challenge ahead is a great one, but our hope is that data from interim assessments can help educators, district and school leaders, and policymakers understand the landscape of unfinished learning and endeavor to address it together.

Key Findings

- Unfinished learning is greater this winter compared to prior school years.
- Unfinished learning in reading is greater for students in Grades 1–7, particularly in early elementary grades.
- Unfinished learning in mathematics is greater for students in all grades, particularly elementary grades.
- Unfinished learning is greater for students in schools serving majority Black and Latino students.
- Unfinished learning is greater for students in schools located in lower-income zip codes.
- It is too early to tell if students are catching up from starting behind in the fall.

Contents

The Language We Use to Describe Learning

At Curriculum Associates, we are committed to becoming a fully inclusive, anti-racist, multicultural organization. We recognize that systemic bias and racism negatively impact students and educators of color and that common terms and characterizations of student achievement data have been and continue to be problematic. In particular, we are cognizant of how bias is embedded in the language we use to describe what students know and are able to do. For example, deficit-based labels such as *underperforming* unfairly place blame on students who in truth have been underprepared by our society. We know that while teachers and school and district leaders deeply invest in these learners, the cumulative and compounding effects of an array of societal factors have systematically disadvantaged people of color.

We take our role in changing that system very seriously. One of our goals, as a curriculum and assessment provider, is to objectively measure learning to inform instruction, reveal inequities, and contribute to the field of education research. We believe that the deficit-based labels that have long been used to describe student learning have nothing to do with their intellectual capacity, effort, or aptitude. Instead, we choose to honor the potential of students and decouple the words we use to describe student achievement from unfair assumptions and habits. This will take some time, but our work has already begun. As our learning journey continues, we will keep reflecting on the impact of our words and strive to use asset-based language that is empowering for all students, teachers, and educators.

Introduction

In March 2020, schools closed their doors due to the COVID-19 global health crisis. Millions of students faced an interruption to learning unknown to generations before them. Several studies released shortly after the initial shutdowns predicted significant academic consequences due to these school closures. Early prediction models estimated that students would return to school in fall 2020 with only 70% of the learning gains in reading and 50% of the learning gains in mathematics relative to a normal school year (Kuhfeld & Tarasawa, 2020). A similar study estimated that the number of days of instruction "lost" due to the spring closures alone could be up to one year in reading and more than a year in mathematics (Center for Research on Education Outcomes, 2020). This early research largely shaped the national narrative.

Curriculum Associates and others set out to research the impact of school closures on student achievement when the school year began in fall 2020. The resulting body of research on unfinished learning from academia and education testing companies largely agrees that while students are behind this year, they did not experience the anticipated precipitous drop in achievement that was initially predicted, and unfinished learning in mathematics is greater than in reading. Additionally, the research tends to agree that school closures have impacted some students more than others and have exacerbated existing inequalities for students of color and historically underserved communities (Dorn et al., 2020; Kogan & Lavertu, 2021; Catalano, 2020).

More than nine million students who are enrolled in public, private, and charter schools nationwide have taken the *i-Ready Diagnostic* this school year. The findings drawn from the Diagnostic assessment represent approximately 25% of the K–8 public school population and paint a picture of student achievement a year after schools closed their doors. Building on what we learned from fall assessment data, we examined the winter data, first by testing location (in school or out of school), and came to the same conclusion: data from assessments taken exclusively in school, as reported by students during their testing experience, is the closest to a "true" comparison to prior-year performance than data from assessments taken at home. Due to concerns over the comparability of out-of-school testing data, we are reporting out results for in-school assessment data only.

This analysis, shared below, finds that after 12 months of school interruptions due to the pandemic, there are fewer students ready to access grade-level work compared to prior years at this point in the school year, which means there is a greater amount of unfinished learning to address compared to a typical school year. The midyear results suggest some of our youngest students, and those historically underserved, have been impacted the most. After looking at a subset of student assessment data from fall to winter, we have come to the preliminary conclusion that it is too early to tell if students are catching up from starting behind in the fall.

As in any school year, educators face the challenge of supporting students in 2020–2021, and the data from interim assessments helps educators understand individual student needs and plan for resources and instruction. Unlike previous school years, the number of students who need additional supports has increased across grade levels in two critical subject areas: reading and mathematics. There are many research-based recommendations to support teaching and learning. This report suggests how some of those research-based recommendations may help to address unfinished learning.

Why Focus on In-School Assessment Data?

This analysis focused on assessment data from in-school testing locations because it is:

- More consistent with historical testing conditions
- Less variable from student to student, and, therefore:
- A more valid comparison to historical performance

Methodology

Research Questions

The primary research questions addressed in this research paper are as follows:

1. How does unfinished learning during winter 2020–2021 compare to what we have seen historically?
2. How does unfinished learning vary by subject and grade level?
3. How does unfinished learning vary by the racial or ethnic makeup of schools?
4. How does unfinished learning vary by the median household income of schools' locations?
5. How has unfinished learning changed for each grade (1–8) since the fall? Have differences in unfinished learning increased or decreased relative to what we would expect based on a historical average?

Sample Description

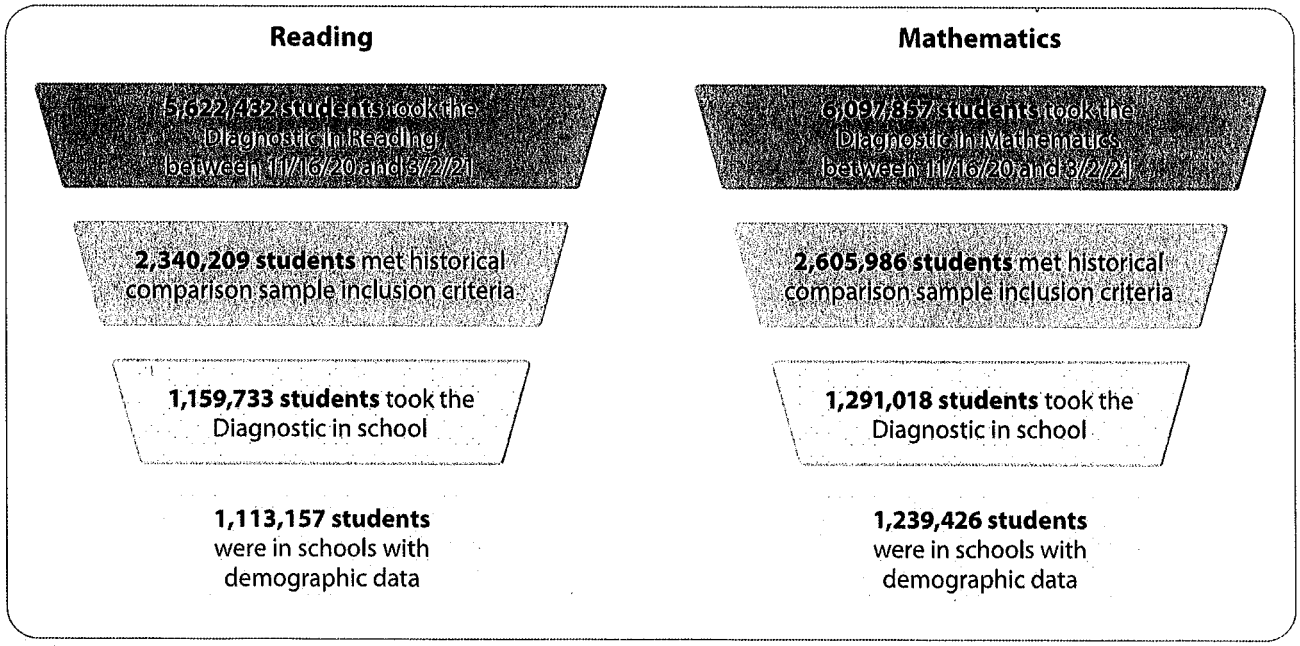
For this study, we examined grade-level placement results from students in winter 2020–2021 compared to prior school years. We constructed a historical average to represent typical performance for students in Grades 1–8 across the three most recent school years: 2017–2018, 2018–2019, and 2019–2020. Student-level data was matched at the school level so the current and historical samples consist of students in the same schools.

In order to have what we considered to be a fair basis of comparison for this analysis, we only included students who tested in school during winter 2020–2021, between November 16, 2020 and March 2, 2021. With these criteria in place, the final analytic sample consisted of 1,159,733 students in Grades 1–8 in the Diagnostic for Reading analysis and 1,291,018 students in Grades 1–8 in the Diagnostic for Mathematics analysis. School-level demographic data was sourced from the National Center for Education Statistics (NCES) Common Core of Data.

This analysis represents students from 49 states, plus the District of Columbia. The number of students per state varied by subject and is not statistically representative of each state.

See Appendix A for more details on the methodology and sample description.

Figure 1: How Was the Winter Assessment Sample Selected?



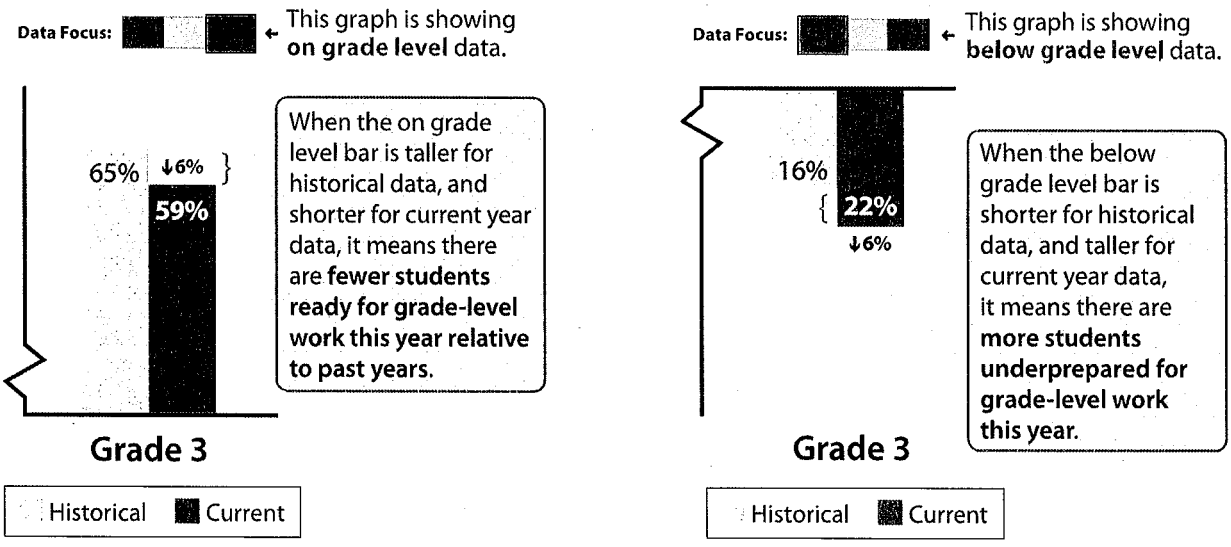
Results

Overview

The following section reports the findings from student assessment data for students who took the Diagnostic in Reading and Mathematics in school. We will begin by sharing the high-level findings across grade levels for each subject this school year relative to the historical average and then discuss the findings for demographic groups by race and ethnicity and income level. We will also look at a subset of students who tested in school during both the fall and winter and look at how their performance levels changed from fall to winter as well as relative to the historical average performance for each testing window.

In this paper, students who placed Early On Grade Level or higher were considered **on grade level** and students who placed Two or More Grade Levels Below were considered **below grade level**. Students who are Early On Grade Level have partially met grade-level college and career readiness standards and students who are Mid or Above Grade Level have met grade-level college and career readiness standards. Students who are Two or More Grade Levels Below are not yet close to meeting grade-level college and career readiness standards and may need additional instruction to fill in gaps in foundational concepts and knowledge. In this paper, we will discuss the national trends we see in the percentage of students who are on grade level and below grade level. We observed somewhat different patterns within each focus area when examining the demographic data in particular that we want the reader to note.

Understanding Grade-Level Placements in This Paper



What Are Grade-Level Placements?

When students take the *i-Ready Diagnostic*, they are given a placement level relative to their chronological grade level that designates the student performance as being on grade, below grade, or above grade. For example, a fifth grader can place below grade at the fourth grade level (One Grade Level Below), at the third grade level (Two Grade Levels Below), and at the second grade, first grade, or kindergarten level (Three or More Grade Levels Below); on grade level (Early On Grade Level, Mid On Grade Level, Late On Grade Level); above grade level as a sixth grader (Above Grade Level), as a seventh grader (Above Grade Level), and at the eighth grade level (Above Grade Level). See Appendix for *i-Ready* placement level descriptors.

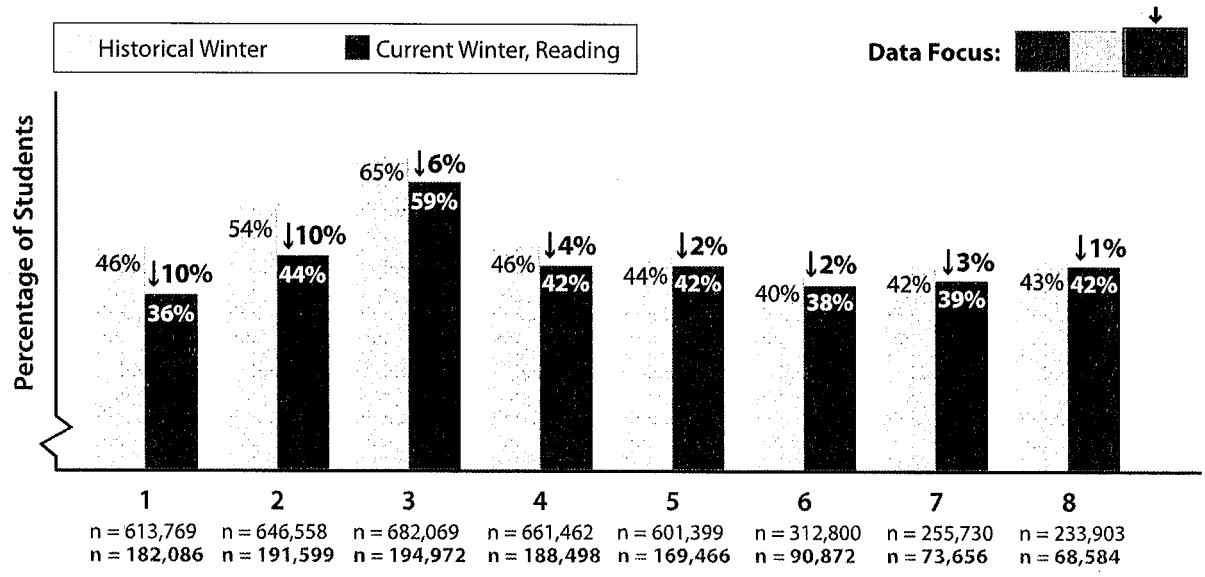
Finding 1

Unfinished Learning in **Reading** Is Greater This Winter Compared to Historical Averages

Reading

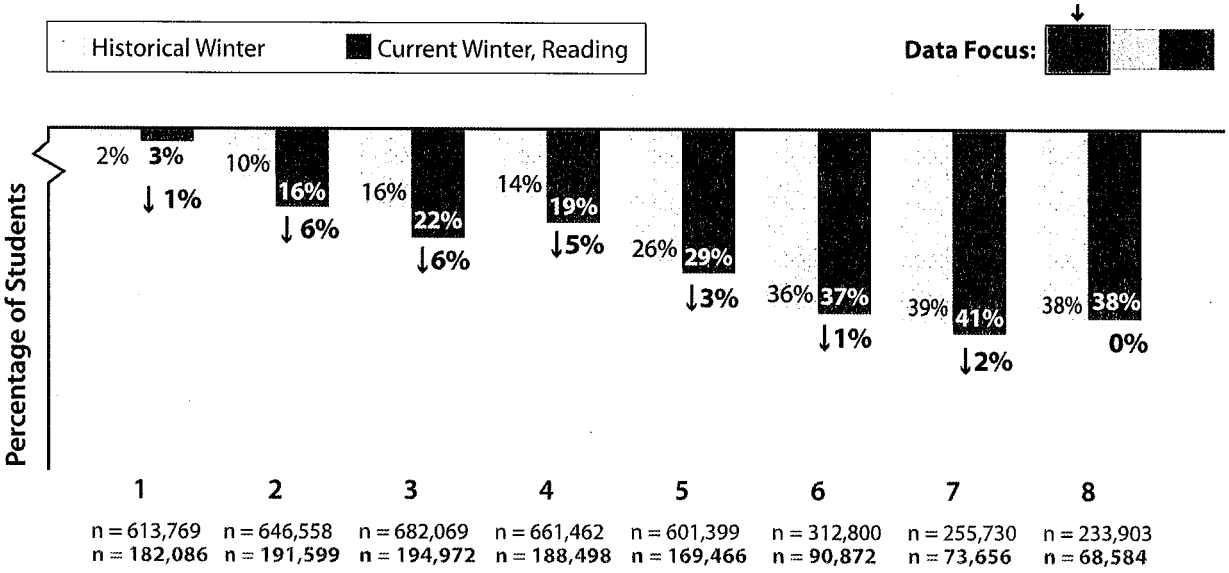
In reading, there is a greater amount of unfinished learning at each grade level, particularly in Grades 1, 2, and 3. The percentage of students who are ready for grade-level work (Early On Grade Level or above) has decreased during the 2020–2021 school year relative to the historical average across all grades.

Graph 1.1: On Grade Level, Reading



Within the same sample, we also looked at the percentage of students who are underprepared for grade-level work (Two or More Grade Levels below). In reading, the percentage of students who are underprepared for grade-level work has increased during the 2020–2021 school year relative to the historical average for students in Grades 1–7, while Grade 8 remains flat.

Graph 1.2: Below Grade Level, Reading



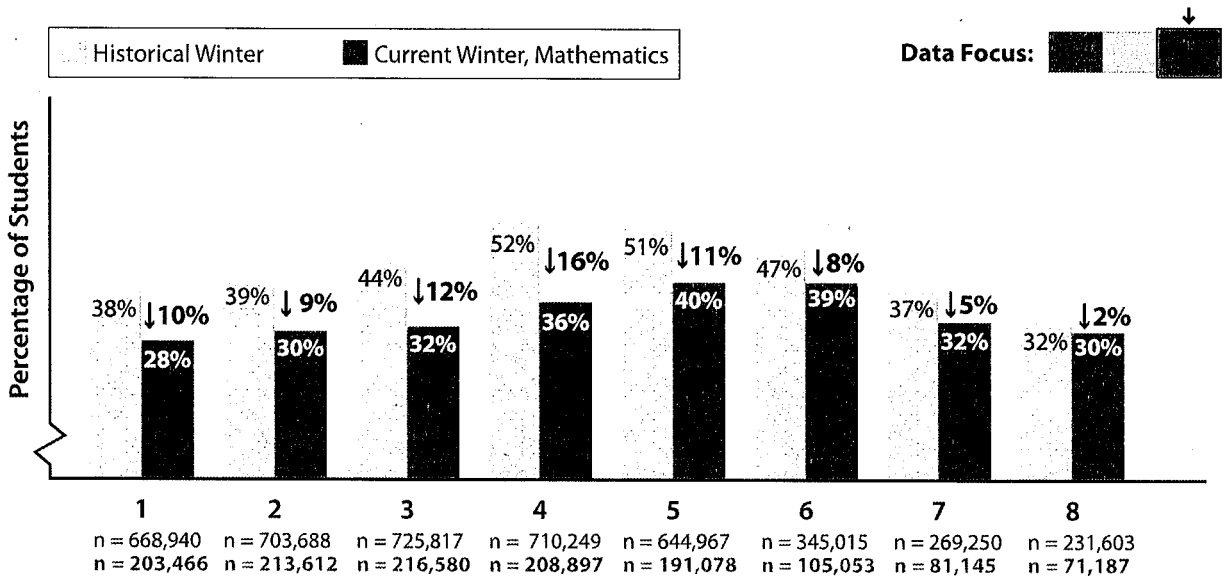
Finding 1

Unfinished Learning in **Mathematics** Is Greater This Winter Compared to Historical Averages

Mathematics

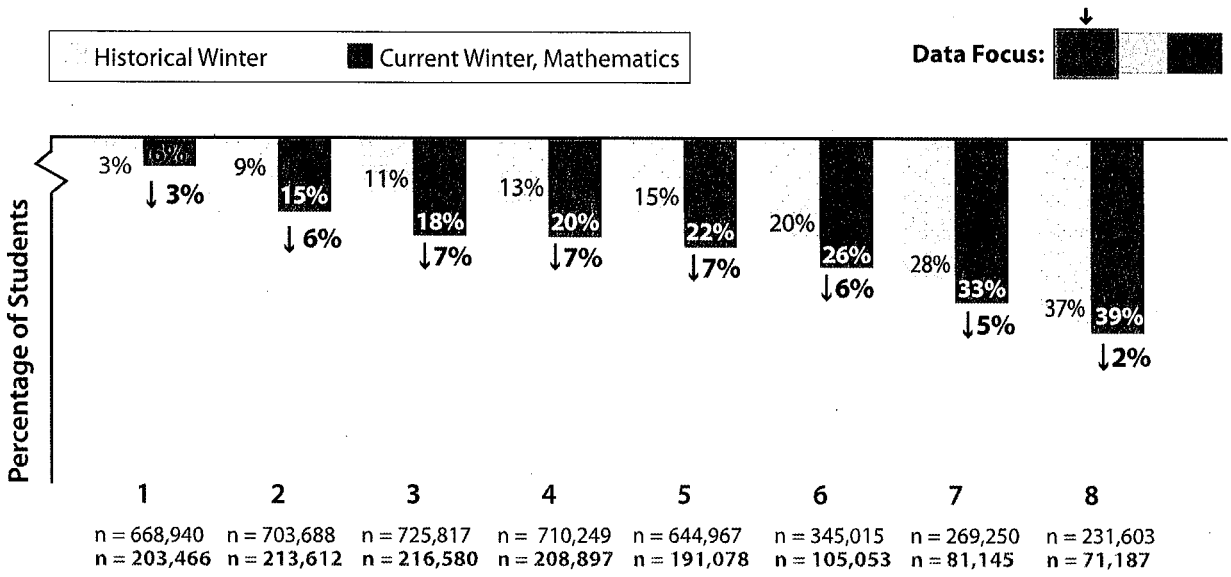
In mathematics, there is a greater amount of unfinished learning across all grades. The percentage of students who are ready for grade-level work (Early On Grade Level or above) has decreased during the 2020–2021 school year relative to the historical average across all grades. Elementary Grades 1–5 and early middle school, Grade 6, show the greatest amount of unfinished learning.

Graph 1.3: On Grade Level, Mathematics



Within the same sample, we also looked at the percentage of students who are underprepared for grade-level work (Two or More Grade Levels Below). In mathematics, the percentage of students who are underprepared for grade-level work has increased during the 2020–20201 school year relative to the historical average for students across all grades. Grades 2–6 show the greatest increases in unfinished learning.

Graph 1.4: Below Grade Level, Mathematics



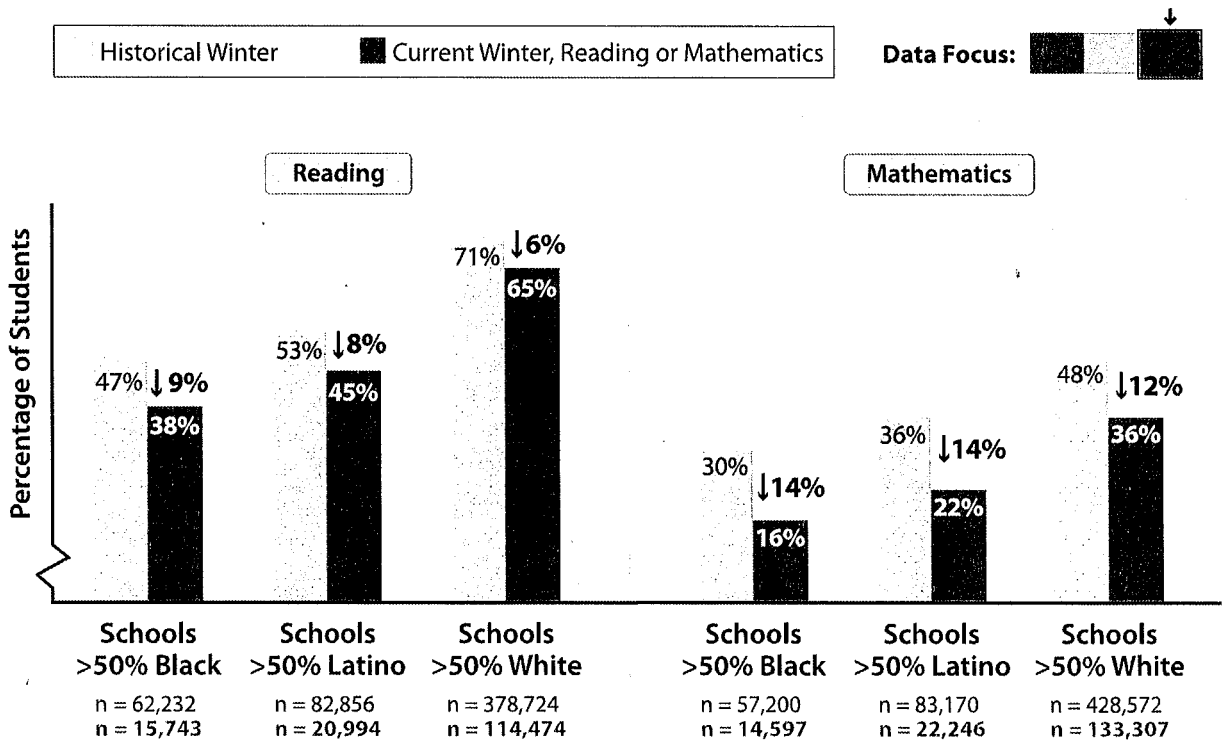
Finding 2

Unfinished Learning Is Greater for Students in Schools Serving a Majority of Black or Latino Students

In this section, we examine the data disaggregated by school-level demographic information in order to look at schools that serve a majority of Black, Latino, and White students. While the majority of Black, Latino, and White schools may contain varying levels of diversity, we chose to group schools this way in order to ensure we had a sufficient sample size for each school-level demographic group.

To illustrate this finding, we are highlighting the results for Grade 3. In reading and mathematics, the percentage of Grade 3 students who are ready for grade-level work has decreased relative to the historical average for students in schools serving a majority of Black, Latino, and White students. The decreases are similar across these three groups within each subject and the historical averages reveal inequities that predate the pandemic.

Graph 2.1: On Grade Level by Demographic Group: Grade 3 Reading and Mathematics

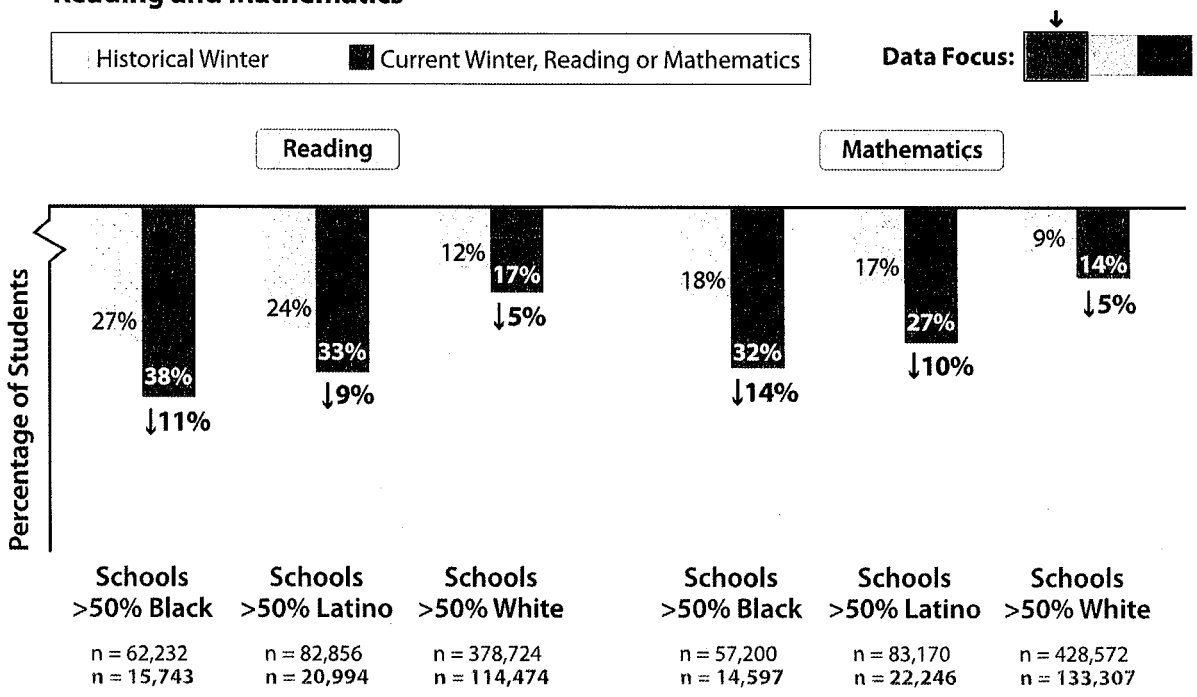


Why Focus on Grade 3?

Throughout this paper, results for Grade 3 students will be illustrated as Grade 3 is a pivotal year for student learning, and research shows performance in Grade 3 is predictive of high school outcomes (Hernandez, 2011).

When looking within the same sample at the percentage of Grade 3 students who are underprepared for grade-level work, however, we can see a larger increase in unfinished learning in Reading for students in schools serving a majority of Black students (11 percentage points) and Latino students (9 percentage points) compared to students in schools serving a majority of White students (5 percentage points). This is also true for schools serving a majority of Black (14 percentage points) and Latino students (10 percentage points) than White students (5 percentage points). The historical averages reveal inequities that predate the pandemic.

Graph 2.2: Below Grade Level by Demographic Group: Grade 3 Reading and Mathematics



Across all grades, the percentage of students who are ready for grade-level work has decreased across schools that serve majority Black, Latino, and White students in reading and mathematics. The following tables present the percentage of students by placement level, subject, and grade for each of the three demographic groups represented on the graphs on the previous page. The results for students in schools serving less than 25% Black, Latino, and White students, as well as students in schools serving between 25% and 50% Black, Latino, and White students, are included in the Appendix.

Table 1: Percentage of Students On Grade Level by Demographic Group, Winter Testing Window; Reading and Mathematics, Grades 1–8

Reading

Percentage On Grade Level						
Grade	>50% Black		>50% Latino		>50% White	
	Historical Winter	Current Winter	Historical Winter	Current Winter	Historical Winter	Current Winter
1	33%	22%	38%	29%	50%	39%
2	36%	26%	44%	33%	59%	49%
3	47%	38%	53%	45%	71%	65%
4	27%	22%	37%	32%	53%	49%
5	24%	22%	34%	31%	49%	47%
6	24%	22%	33%	32%	44%	42%
7	25%	22%	36%	34%	46%	42%
8	27%	24%	39%	40%	47%	45%

Mathematics

Grade	>50% Black		>50% Latino		>50% White	
	Historical Winter	Current Winter	Historical Winter	Current Winter	Historical Winter	Current Winter
1	24%	14%	30%	20%	43%	32%
2	24%	14%	30%	20%	44%	34%
3	30%	16%	36%	22%	48%	36%
4	39%	20%	46%	28%	56%	40%
5	33%	21%	42%	29%	56%	44%
6	32%	24%	42%	35%	50%	42%
7	25%	20%	29%	28%	42%	35%
8	22%	18%	20%	29%	36%	32%

Table 2: Percentage of Students Below Grade Level by Demographic Group, Winter Testing Window; Reading and Mathematics, Grades 1–8

Reading

Percentage Below Grade Level						
Grade	>50% Black		>50% Latino		>50% White	
	Historical Winter	Current Winter	Historical Winter	Current Winter	Historical Winter	Current Winter
1	4%	6%	4%	6%	1%	2%
2	18%	28%	16%	25%	7%	12%
3	27%	38%	24%	33%	12%	17%
4	24%	35%	19%	28%	11%	14%
5	44%	48%	34%	38%	21%	24%
6	52%	55%	43%	44%	31%	33%
7	57%	61%	46%	46%	35%	37%
8	56%	58%	44%	40%	34%	34%

Mathematics

Grade	>50% Black		>50% Latino		>50% White	
	Historical Winter	Current Winter	Historical Winter	Current Winter	Historical Winter	Current Winter
1	6%	11%	5%	10%	2%	5%
2	16%	27%	14%	24%	7%	12%
3	18%	32%	17%	27%	9%	14%
4	21%	35%	17%	29%	10%	16%
5	26%	37%	21%	30%	13%	18%
6	31%	40%	25%	31%	17%	23%
7	40%	48%	36%	40%	24%	30%
8	49%	57%	50%	43%	32%	36%

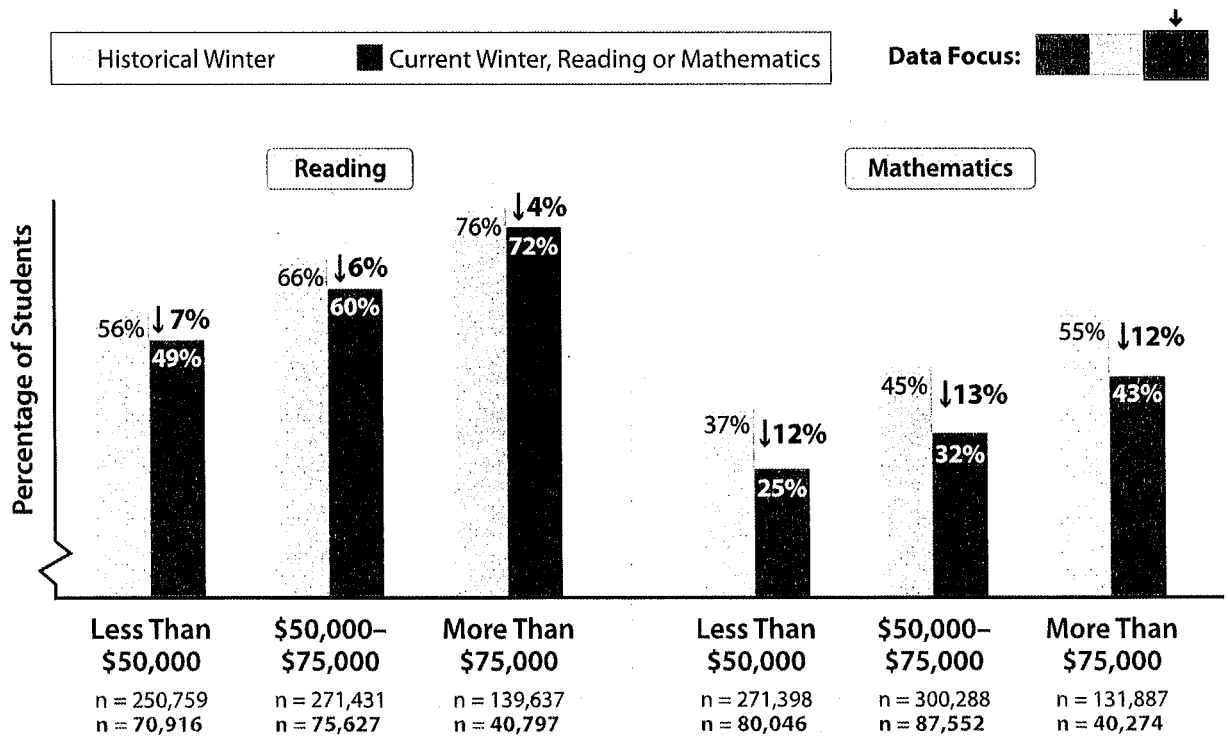
Finding 3

Unfinished Learning Is Greater for Students in Schools Located in Lower-Income Zip Codes

In this section, we examine the data disaggregated by the median annual household income associated with a school's zip code. Across grade levels and subjects, the percentage of students who are ready for grade-level work has decreased this winter relative to the historical average for students, regardless of income bracket.

To illustrate this finding, we are highlighting the results for Grade 3. In reading, the Grade 3 decline relative to the historical average is a little lower for students in schools where the income is greater than $75,000 (4 percentage points) than the decline for students in schools where the income is less than $50,000 (7 percentage points) or $50,000 to $75,000 (6 percentage points). The declines are relatively stable for mathematics across all three income groups.

Graph 3.1: On Grade Level by Income: Grade 3, Reading and Mathematics



As shown below, the percentage of Grade 3 students who are underprepared for grade level work increased for students across schools regardless of income bracket. In reading, the Grade 3 declines relative to the historical average are steeper for students in schools in zip codes where the median household income is below $50,000 annually (8 percentage points) compared with students in schools in zip codes where the median household income is between $50,000 to $75,000 (6 percentage points) and students in schools in zip codes where the median household income is greater than $75,000 (3 percentage points). This is also true for Grade 3 mathematics (8, 6 and 4 percentage points, respectively).

Graph 3.2: Below Grade Level by Income: Grade 3, Reading and Mathematics

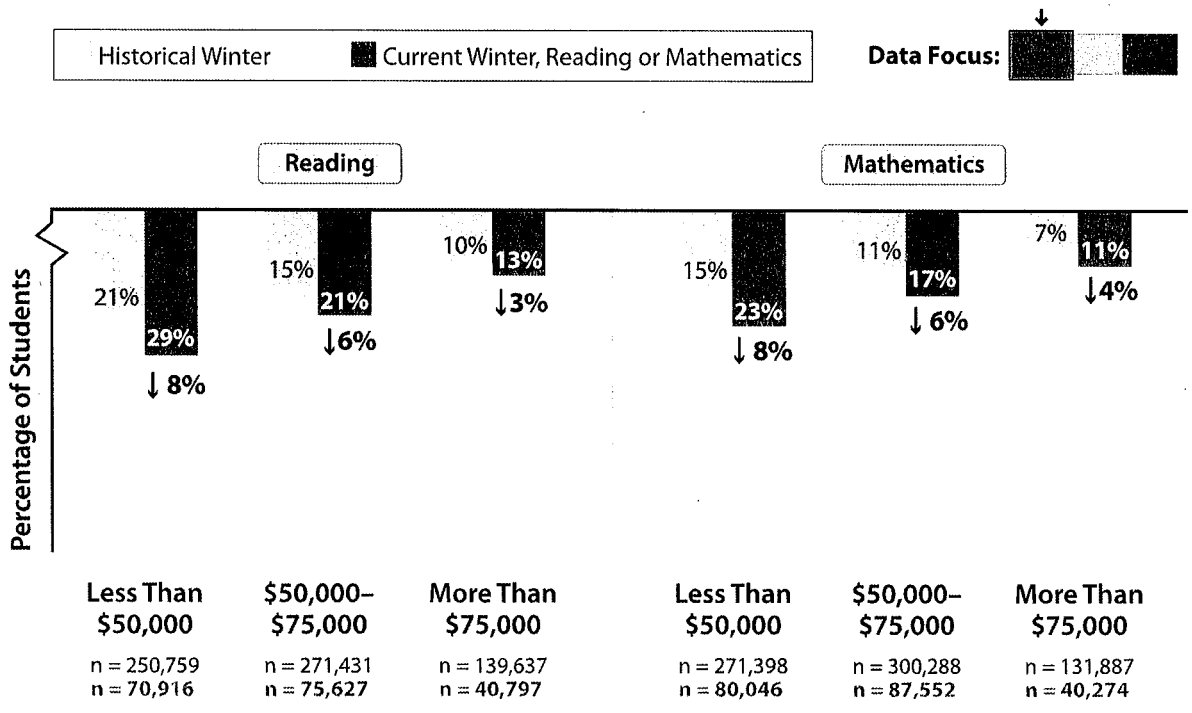


Table 3 presents the percentage of students by placement level, subject, and grade for each of the three income groups.

Table 3: Percentage of Students On Grade and Below Grade Level by Subject and Grade by Income Group, Winter Testing Window; Reading and Mathematics, Grades 1–8

	Percentage On Grade Level					
	Grade	<$50,000		$50,000–$75,000		>$75,000	
		Historical	Current	Historical	Current	Historical	Current
Reading	1	38%	28%	47%	36%	56%	48%
	2	45%	34%	55%	44%	65%	57%
	3	56%	49%	66%	60%	76%	72%
	4	37%	33%	48%	44%	60%	57%
	5	34%	33%	45%	43%	57%	55%
	6	33%	31%	42%	41%	52%	50%
	7	35%	33%	44%	41%	56%	53%
	8	35%	35%	46%	44%	57%	55%
Mathematics	1	31%	21%	39%	29%	50%	40%
	2	32%	22%	40%	31%	51%	42%
	3	37%	25%	45%	32%	55%	43%
	4	46%	29%	53%	37%	63%	48%
	5	43%	32%	53%	41%	63%	52%
	6	40%	32%	49%	41%	61%	51%
	7	31%	26%	40%	35%	51%	44%
	8	27%	24%	34%	31%	45%	42%

	Percentage Below Grade Level					
	Grade	<$50,000		$50,000–$75,000		>$75,000	
		Historical	Current	Historical	Current	Historical	Current
Reading	1	3%	5%	2%	3%	2%	2%
	2	14%	21%	10%	15%	6%	9%
	3	21%	29%	15%	21%	10%	13%
	4	18%	26%	13%	18%	8%	11%
	5	33%	37%	25%	28%	16%	18%
	6	42%	44%	33%	34%	24%	25%
	7	46%	48%	37%	39%	26%	27%
	8	46%	45%	35%	36%	26%	25%
Mathematics	1	4%	8%	3%	6%	2%	4%
	2	12%	20%	9%	14%	5%	9%
	3	15%	23%	11%	17%	7%	11%
	4	17%	26%	12%	19%	8%	12%
	5	20%	28%	14%	20%	9%	14%
	6	24%	31%	18%	24%	11%	17%
	7	33%	39%	26%	31%	18%	22%
	8	42%	45%	35%	37%	24%	26%

Finding 4

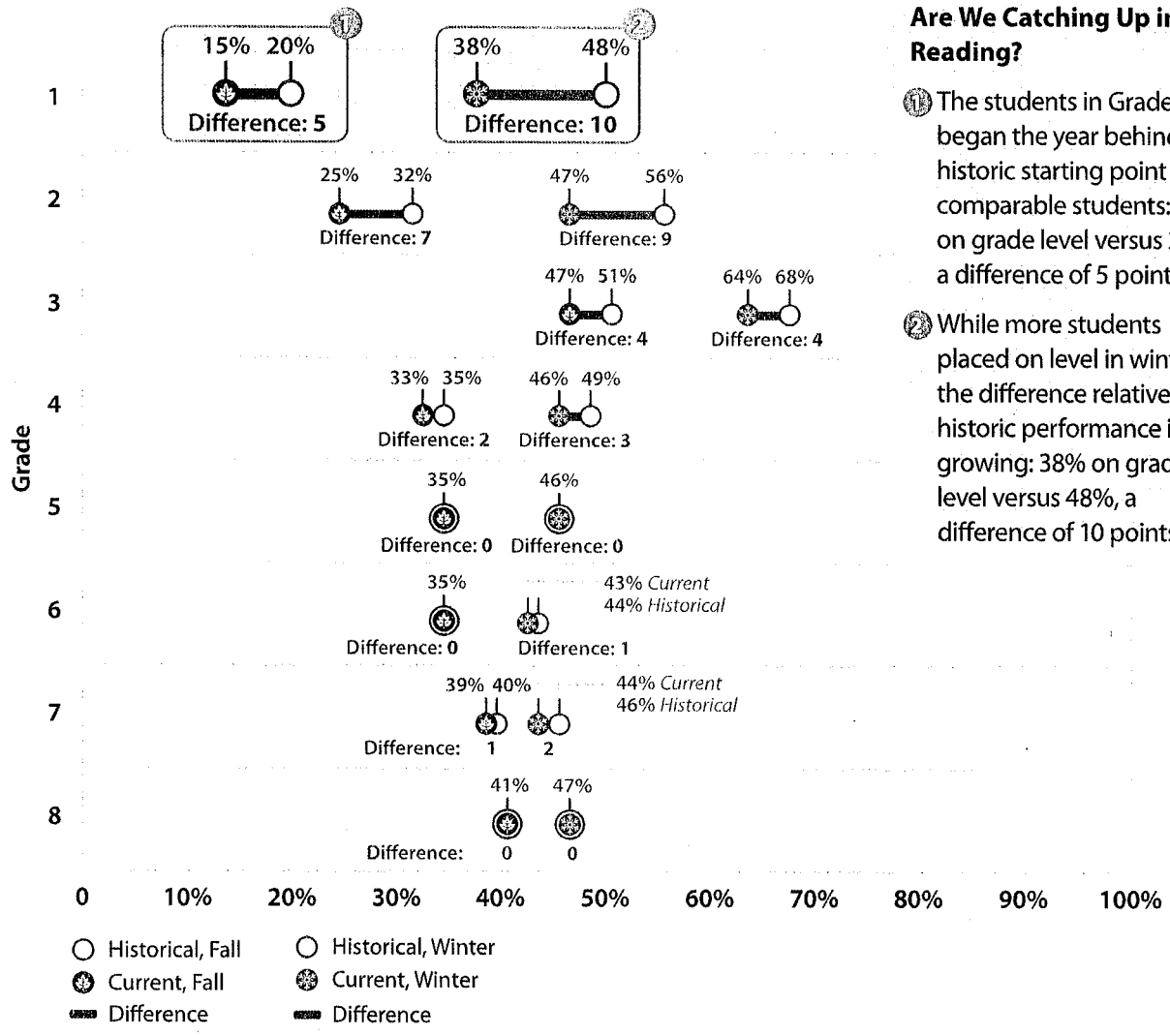
It Is Too Early to Tell If Students Are Catching Up from Starting behind in the Fall

In this section, we examine the change in grade-level placements for a subset of students who indicated they took the Diagnostic in school during both the fall and the winter assessment windows. First, we looked at the percentage of students who were ready for grade-level work and whether, in each season, we saw unfinished learning relative to the historical average. Next, we compared the difference in percentage points between the historical average and the current school year at fall and at winter.

Our visual analysis of data across two time points (fall and winter) shows that there is variability across subjects and grade levels. In some subject and grade levels, the difference between the current school year and the historical average increased from fall to winter and in some subjects and grade levels the difference decreased. When looking at the percentage of students who were ready for grade-level work, a decrease in the differences indicates that students are catching up from where they started behind in the fall, and an increase in the differences indicates that students are not catching up from where they started behind in the fall. Given the variability we saw across subjects and grade levels, the midyear results are inconclusive.

Graphs 4.1 and 4.2 display the differences from fall to winter for the percentage of students who were ready for grade-level work for reading and mathematics, respectively. We recommend interpreting with caution, as the results are limited in generalizability due to the sample constraints. Specifically, the number of students in this subsample who had both a fall and winter Diagnostic taken in school is just under half of the total number of students whose findings are reported in the other sections of this report.

Graph 4.1: Difference between Percentage of Students On Grade Level in Fall and Winter Compared to Historical in Reading



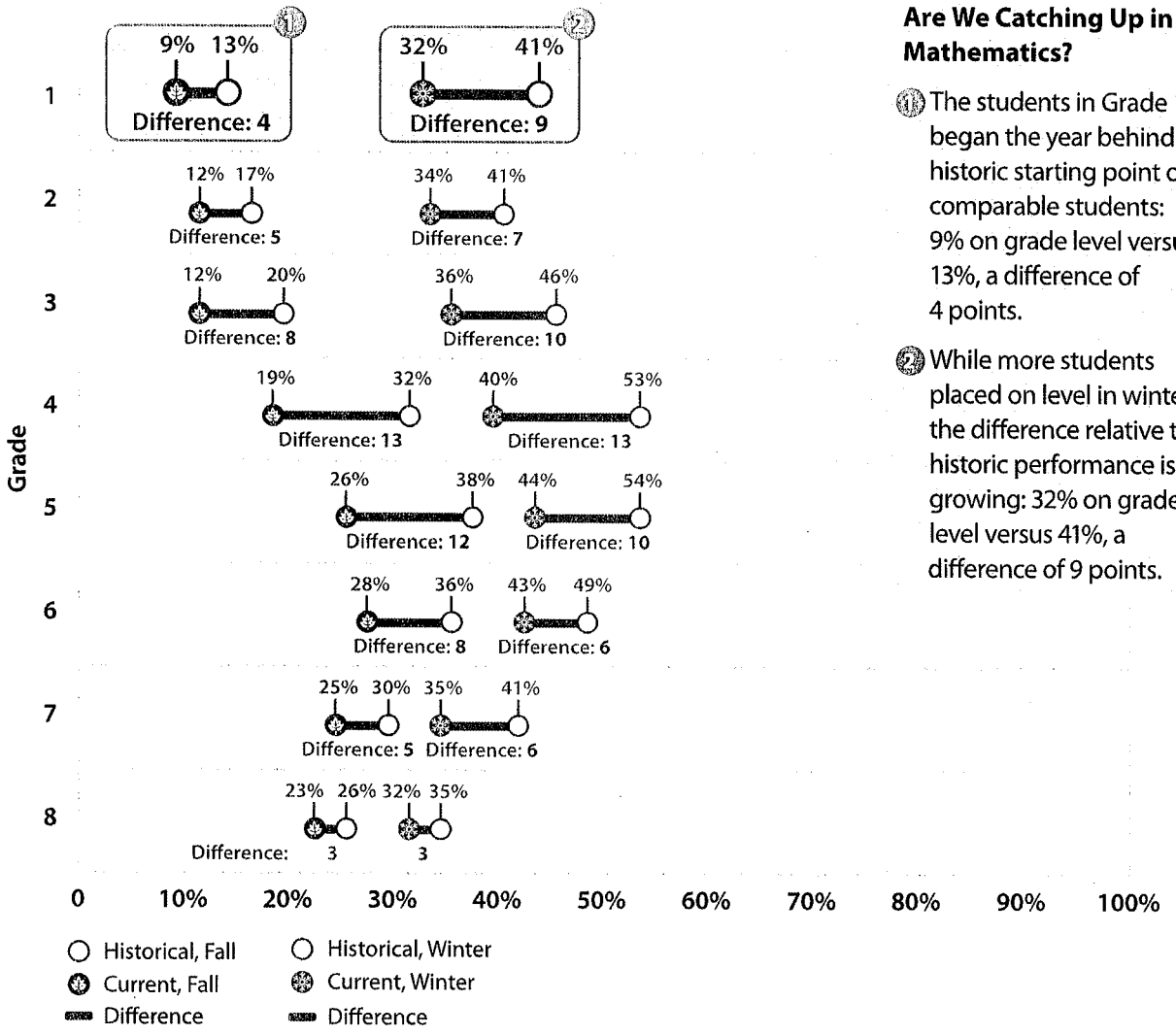
Are We Catching Up in Reading?

① The students in Grade 1 began the year behind the historic starting point of comparable students: 15% on grade level versus 20%, a difference of 5 points.

② While more students placed on level in winter, the difference relative to historic performance is growing: 38% on grade level versus 48%, a difference of 10 points.

See Appendix Table 6 for the differences and difference in differences.

Graph 4.2: Difference between Percentage of Students On Grade Level in Fall and Winter Compared to Historical in Mathematics



Are We Catching Up in Mathematics?

(1) The students in Grade 1 began the year behind the historic starting point of comparable students: 9% on grade level versus 13%, a difference of 4 points.

(2) While more students placed on level in winter, the difference relative to historic performance is growing: 32% on grade level versus 41%, a difference of 9 points.

See Appendix Table 6 for the differences and difference in differences.

Addressing Unfinished Learning

The results from the winter *i-Ready Diagnostic* assessments in reading and mathematics suggest the challenges of addressing unfinished learning will be persistent and significant. Educators play a unique and critical role in helping mitigate the effects of months of interrupted learning, and their understanding of student needs will inform the best tactics to employ. Below are several actions educators can take, drawn from the priorities we hear from our partners, insights from recent research, and the expertise of national organizations, including the Council of Chief State School Officers (CCSSO) and the Council of the Great City Schools (CGCS).

Ensure assessments deliver clear and actionable data. An assessment serves student learning only when it provides a clear view of student needs and related instructional supports. Conflicting reports on the state of unfinished learning underscore how critical it is that assessments accurately account for missing students, report on sociodemographic implications, and capture whether students test in or out of school since this impacts the fidelity of assessment. With instructional time more precious than ever, educators must be able to quickly and deeply understand how best to help each learner find success at grade level (CCSSO and CGCS, 2014). An effective, informative assessment should be criterion-referenced against benchmarks and identify prerequisite skills needed to reach the ultimate goals of grade-level proficiency and college and career readiness. Tools should also make it simple for teachers to regularly monitor student performance to ensure the most vulnerable students who may potentially be falling further behind get the support they need. This may be especially important given the urgency that the US Department of Education has placed on understanding the pandemic's impact on learning and uncertainty around end-of-year summative testing.

Choose high-quality, rigorous curriculum. To address unfinished learning, the CGCS underscores the importance of focusing not on remediation and reteaching, but rather recommends: "School and district curriculum leaders should keep the focus on grade-level work and rigor, addressing learning gaps as needed within the context of grade level" (CGCS, 2020). In a recent report on leadership recommendations for school reentry, Chiefs for Change and the Johns Hopkins University Institute for Education Policy echo the critical importance of "comprehensively adopting high-quality instructional materials with robust teacher supports." To ensure all students, and particularly those whose learning has been most impacted during the pandemic, have equitable opportunities to reach grade-level proficiency, we must provide them with engaging, high-quality, grade-level work (TNTP, 2018). Teachers need curricular supports designed to uphold instructional content priorities and make teaching to college- and career-readiness standards more efficient (Student Achievement Partners, 2020).

Set ambitious yet attainable goals for all students. Research shows students make greater learning gains when their teachers hold high expectations about their ability to meet grade-level standards (TNTP, 2018). This can change learning trajectories, as the impact of learning from a teacher with rigorous, high standards can improve student performance long after moving to a new class (Gershenson, 2020). To support high expectations, teachers must be confident the resources in their toolkit are the best tools to move all learners toward grade-level performance. In supporting appropriately challenging growth goals, educators need assessments that guide grade-level instructional priorities, clear grade-level benchmarks, and scaffolds to address underlying unfinished learning. With high expectations in place, incremental goal setting and monitoring progress are a natural next step as teachers and students work together toward the ultimate shared goal of attaining grade-level proficiency.

Prioritize coherence. As educators implement new and creative instructional opportunities to address unfinished learning, it is critical to ensure that student assessment data provides a clear, composite view of student learning across settings. Ensuring the data from each learning environment is talking to the others helps save time and avoids redundant activities, providing a comprehensive, accessible picture of progress and needs at any given time. Coherence reduces overlapping assessments, in keeping with the guiding principles for assessment set forth by the CCSSO and the CGCS, which states "Assessments should be administered in only the numbers and duration that will give us the information that is needed and nothing more. Multiple assessments of the same students for similar purposes should be minimized or eliminated" (CCSSO and CGCS, 2014). Coherence looks like after-school learning connected to in-school learning, summer learning naturally picking up where the school year ended, unfinished learning addressed alongside grade-level learning, and IEPs directly informing tutoring. When data reflects the sum of learning activities, it empowers teachers to better serve students and helps students get what they need without added burden as they move between settings.

Engage students. Programs designed to be culturally and linguistically responsive (CLR) invite students to see their personal and cultural experiences reflected in the content and engage more deeply in learning. When students are validated and affirmed, they understand the cultural and linguistic experiences they bring to their learning to be assets, which supports connection to the material. CLR teaching can impact student gains and support grade-level attainment goals. In a large study, students with teachers who identified as "high implementers" of a CLR program scored significantly higher on their spring benchmark test in reading than students with teachers who were "low implementers" of the program (Powell, Cantrell, Malo-Juvera & Correll, 2016).

While the suggestions above are intended to guide decisions about learning tools to support educators, the work of addressing and overcoming unfinished learning will be complex and involve a range of non-academic supports. For many, addressing learning needs must begin with addressing social-emotional wellness, as students need help processing and working through trauma experienced in the past year. For some schools, supporting learning may take the form of partnerships with community organizations, innovative summer programs, or family engagement initiatives. Educators who have done heroics in the past year to support learning must balance the exhaustion they feel with resolve to tackle the challenges ahead, and they deserve the best possible supports to do this critical work.

Limitations

The findings in this paper rely on student self-reported data on the location of where they took the *i-Ready Diagnostic* test. We acknowledge this is an imperfect measure. Over half of students who took the *i-Ready Diagnostic* this winter tested remotely and are not reflected in this report. In addition, we know from comparing the in-school and out-of-school data that students who tested in school were more likely to attend schools serving a majority of White students and are more likely to be in towns and rural areas. Ultimately, we chose to focus our findings on the in-school testing results due to higher data consistency with in-school testing data as compared to out-of-school testing data.

The findings in this paper describe the school-level demographics, which is not the same as relying on student-level demographics. Schools consisting of more than 50% of one racial or ethnic group may still be fairly diverse, and we recognize that using school-level demographics does not capture that diversity nor the variability in unfinished learning within each school demographic group. We do not have visibility into where students spent most of their time learning during the 2020–2021 school year. Where a student took an assessment should not be conflated with where a student is learning (e.g., entirely in a traditional school building, entirely remote in their home or another location outside of their school building, or in multiple locations as part of a hybrid model). In this analysis, student use of *i-Ready Personalized Instruction* was not taken into account.

In order to describe the change in grade-level performance from fall to winter, we limited the analysis to only those students who took an *i-Ready Diagnostic* in school during both the fall and winter assessment window. This group represents approximately 40% of the general analysis population described in this paper and less than 10% of the total *i-Ready Diagnostic* testing population. Given the further constrained sample in combination with the variation in grade- and subject-level results, we do not wish to draw a sweeping conclusion about the grade-level fall-to-winter findings. At the same time, we do want to share what we know with educators and plan to continue to monitor student assessment data for the remainder of the school year.

Conclusion

Our analysis of midyear assessment data shows more students have unfinished learning and fewer students are ready for grade-level work this school year than in prior school years in both reading and mathematics. The students who are most affected are students in elementary school, students attending schools that serve a higher proportion of Black and Latino students, and students attending schools in lower-income zip codes. Our grade-level analysis for the subset of students who took their assessments in school during both testing windows (fall and winter) is inconclusive as to whether students are catching up from where they started behind this fall. We know that educators are always focused on addressing unfinished learning. It is our hope that this report provides a clearer picture of where students are this winter to support educators in their work this spring and summer. We will continue to investigate the impact of the pandemic on student learning and release subsequent research publications and issue briefs as the data becomes available.

Appendix

Methodology and Sample Description

Students who took an *i-Ready Diagnostic* test during fall and/or winter of the 2020–2021 school year were eligible for inclusion in this study. To be considered in school, the student had to both self-report that their test was taken in school and belong to a school where the number of students testing in school this year was comparable to last year.

In the historical sample, we kept all students from the selected schools under the assumption that all students tested in school prior to school closures. Because many schools contain a mixture of in-school and remote testers this year, the 2020–2021 student counts will generally appear lower than the average single-year student counts from the historical sample.

All analyses were conducted at the student level. For analyses with school-level demographic variables, the school-level demographic group is treated as a student-level variable. Therefore, the interpretation is, for example, "students in schools located in lower-income zip codes tend to perform lower than students in schools located in higher-income zip codes."

Out-of-school testing data had more variability in terms of both scores and test administration data, such as test duration, number of testing sessions, and number of devices used. For this reason, we focus most of our findings on the in-school testing population as it is the fairest basis of comparison to a typical school year.

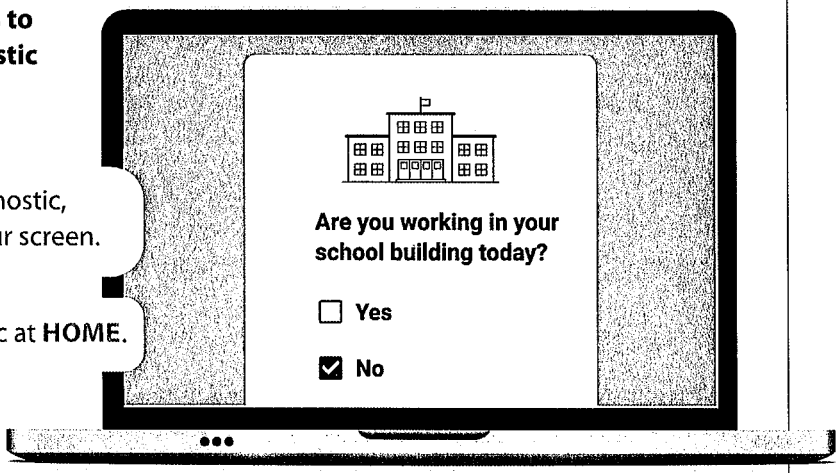
How Was Student Testing Location Determined?

Figure 2. How Was Location Determined?

Platform Popup Question for Students to Select If They Were Taking the Diagnostic in School or in an Out-of-School Test Environment

Once you **START** or **RESUME** your Diagnostic, you will see this question pop up on your screen.

Click **NO** if you are taking the Diagnostic at **HOME**.



Sample Inclusion Criteria

Students who met the following criteria were included in the general analysis:

- Enrollment in Grades 1–8
- Self-report that their test was taken in school
- Belong to a school district that had at least one enrolled student in the three most-current school years (2018–2019, 2019–2020, 2020–2021) for their test subject
- Belong to a school that tested students in their subject and grade during the winter of the current (2020–2021) and prior (2019–2020) school years
- Belong to a school where the percentage of students tested in school in 2020–2021 was between 50% to 200% of same testing window in 2019–2020
- Belong to a school where at least five students tested in school for their test subject and grade
- The winter test was not rushed.
- For demographic analyses, school must be included in NCES CCD in 2018–2019

In order to be included in the fall-to-winter change in the grade-level placement sub-analysis, students had to additionally have taken a Diagnostic in school during the fall testing window (August 1, 2020 through November 15, 2020).

School-Level Demographic Groups

In order to answer the research questions pertaining to race and ethnicity and median household income, we developed the following reporting groups based on available school-level demographics for the population of students who tested in school. Students were grouped based on whether their school served:

- Less than 25% Black students, 25% to 50% Black students, or more than 50% Black students
- Less than 25% Latino students, 25% to 50% Latino students, or more than 50% Latino students
- Less than 25% White students, 25% to 50% White students, or more than 50% White students
- Located in zip codes where the median household income is less than $50,000, ranges from $50,000 to $75,000, or is more than $75,000

While the more than 50% Black, Latino, and White schools may contain varying levels of diversity, we chose to group schools this way in order to ensure that we had a sufficient sample size for each school-level demographic group.

The school-level data on race and ethnicity used in this analysis was sourced from the NCES, which asks students to identify as American Indian or Alaska Native, Asian, Black or African American, Hispanic, Native Hawaiian or Other Pacific Islander, White, or Two or More Races. Throughout this paper, we use the term "Black" to refer to the NCES category of Black or African American and the term "Latino" to refer to the NCES category of Hispanic.

We recognize language changes with time and each demographic group described is not monolithic, nor are all individuals within any designated demographic group in agreement on preferred language. As a company, we will continue to review, reflect on, and evolve the terminology with the goal of using bias-free, inclusive, and sensitive-language labels.

Additional Sample Description Data

Student counts and school-level demographic data are provided for both the in-school testing population (reported) and the out-of-school testing population (not reported).

Appendix Table 1.1: Number of Students by Subject and Grade In-School Testing Population, Winter

In School				
Grade	Reading		Mathematics	
	Historical	Current	Historical	Current
1	613,769	182,086	668,940	203,466
2	646,558	191,599	703,688	213,612
3	682,069	194,972	725,817	216,580
4	661,462	188,498	710,249	208,897
5	601,399	169,466	644,967	191,078
6	312,800	90,872	345,015	105,053
7	255,730	73,656	269,250	81,145
8	233,903	68,584	231,603	71,187

Appendix Table 1.2: Number of Students by Subject and Grade Out-of-School Testing Population, Winter

Out of School				
Grade	Reading		Mathematics	
	Historical	Current	Historical	Current
1	362,355	115,091	403,434	128,032
2	456,555	147,145	500,109	161,032
3	503,083	155,421	544,120	170,650
4	562,089	170,336	614,648	189,753
5	543,857	167,182	607,885	190,966
6	482,282	147,453	509,017	163,005
7	453,033	138,225	498,219	159,900
8	435,363	139,623	449,877	151,630

Note: Diagnostic test results for students who tested out of school are not included in the report findings.

Appendix Table 2.1: School-Level Demographic Characteristics In-School and Out-of-School Testing Population, Winter

In School				
	Reading		Mathematics	
	Average	Range	Average	Range
% American Indian	0.4%	0–89%	0.4%	0–89%
% Asian	2.5%	0–83%	2.3%	0–83%
% Black	15.8%	0–100%	14.4%	0–100%
% Hawaiian or Pacific Islander	0.3%	0–70%	0.3%	0–70%
% Latino	19.6%	0–100%	18.7%	0–100%
% White	57.0%	0–100%	59.6%	0–100%
Median Annual Household Income	$59,592	$10,554 to $235,714	$59,357	$10,554 to $235,714
Student Enrollment	497	15 to 8,761	490	9 to 3,213

Appendix Table 2.2: School-Level Demographic Characteristics In-School and Out-of-School Testing Population, Winter

Out of School				
	Reading		Mathematics	
	Average	Range	Average	Range
% American Indian	0.7%	0–95%	0.8%	0–100%
% Asian	9.1%	0–86%	8.7%	0–86%
% Black	20.5%	0–100%	20.9%	0–100%%
% Hawaiian or Pacific Islander	1.1%	0–82%	1.1%	0–82%
% Latino	37.5%	0–100%	37.0%	0–100%
% White	26.4%	0–100%	27.1%	0–100%
Median Annual Household Income	$66,261	$13,087 to $235,714	$65,590	$13,087 to $223,434
Student Enrollment	530	4 to 11,173	526	4 to 11,173

Note: Diagnostic test results for students who tested out of school are not included in the report findings.

Additional Results

Appendix Table 3: Percentage of Students On and Below Grade Level in Reading by Demographic Group In-School Testing Population, Winter

Percentage On Grade Level												
Grade	Less Than 25% Black		25% to 50% Black		Less Than 25% Latino		25% to 50% Latino		Less Than 25% White		25% to 50% White	
	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current
1	48%	38%	39%	29%	47%	37%	44%	35%	35%	25%	45%	35%
2	57%	46%	45%	35%	56%	46%	51%	41%	39%	29%	52%	41%
3	68%	62%	55%	49%	67%	62%	62%	54%	48%	40%	63%	55%
4	50%	46%	37%	32%	48%	45%	44%	39%	31%	26%	44%	39%
5	47%	45%	34%	33%	46%	44%	41%	38%	29%	26%	40%	38%
6	42%	40%	32%	31%	41%	39%	33%	31%	26%	24%	32%	31%
7	44%	41%	35%	33%	44%	40%	36%	33%	30%	28%	33%	30%
8	45%	44%	34%	33%	44%	43%	36%	35%	29%	30%	33%	31%

Percentage Below Grade Level												
Grade	Less Than 25% Black		25% to 50% Black		Less Than 25% Latino		25% to 50% Latino		Less Than 25% White		25% to 50% White	
	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current
1	2%	3%	4%	5%	2%	3%	3%	4%	5%	7%	3%	4%
2	9%	14%	14%	21%	9%	14%	12%	18%	18%	28%	11%	17%
3	14%	20%	22%	31%	15%	20%	18%	25%	27%	37%	17%	25%
4	12%	17%	19%	28%	13%	17%	15%	23%	22%	34%	15%	21%
5	23%	26%	34%	38%	24%	27%	28%	33%	39%	44%	29%	33%
6	33%	34%	44%	45%	34%	36%	43%	44%	51%	52%	44%	45%
7	37%	39%	47%	49%	38%	40%	46%	48%	54%	54%	49%	52%
8	35%	36%	48%	47%	37%	37%	45%	45%	54%	51%	49%	50%

Appendix Table 4: Percentage of Students On and Below Grade Level in Mathematics by Demographic Group In-School Testing Population, Winter

Percentage On Grade Level												
Grade	Less Than 25% Black		25% to 50% Black		Less Than 25% Latino		25% to 50% Latino		Less Than 25% White		25% to 50% White	
	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current
1	41%	31%	29%	19%	40%	30%	35%	26%	27%	17%	35%	25%
2	42%	32%	31%	21%	42%	32%	36%	27%	27%	18%	36%	26%
3	46%	34%	37%	24%	46%	34%	42%	28%	32%	19%	41%	28%
4	55%	38%	45%	29%	54%	38%	50%	33%	42%	23%	50%	33%
5	54%	42%	43%	32%	53%	42%	48%	35%	38%	25%	46%	34%
6	49%	41%	40%	33%	49%	41%	38%	31%	35%	27%	41%	33%
7	40%	34%	30%	27%	40%	34%	28%	24%	25%	22%	29%	25%
8	34%	31%	25%	23%	35%	31%	23%	21%	17%	22%	26%	22%

Percentage Below Grade Level												
Grade	Less Than 25% Black		25% to 50% Black		Less Than 25% Latino		25% to 50% Latino		Less Than 25% White		25% to 50% White	
	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current	Historical	Current
1	3%	6%	5%	9%	3%	6%	4%	8%	6%	10%	4%	7%
2	8%	13%	13%	21%	8%	13%	11%	18%	15%	26%	11%	18%
3	10%	16%	15%	25%	10%	16%	13%	21%	18%	31%	13%	21%
4	12%	18%	17%	27%	12%	18%	15%	24%	20%	34%	15%	24%
5	14%	20%	20%	28%	14%	20%	17%	25%	24%	34%	18%	26%
6	18%	24%	24%	32%	18%	25%	26%	33%	30%	38%	24%	32%
7	26%	32%	35%	39%	26%	32%	36%	41%	40%	46%	36%	41%
8	34%	37%	47%	47%	34%	38%	47%	49%	54%	52%	45%	49%

Additional Results for Students with Fall and Winter Data

Appendix Table 5: Number of Students by Subject and Grade Level In-School Testing Population, Fall and Winter

Grade	Number In School for Current School Year			
	Reading		Mathematics	
	Historical	Current	Historical	Current
1	219,101	72,956	261,650	88,184
2	236,688	80,033	275,686	93,169
3	220,684	72,539	268,005	87,441
4	224,998	71,932	258,867	84,448
5	203,255	64,717	239,188	77,150
6	126,776	40,623	138,912	45,366
7	100,876	31,678	107,186	34,658
8	93,508	30,415	89,913	29,837
All	1,425,886	464,893	1,639,407	540,253

Appendix Table 6: Percentage of Students Placing On Grade Level In-School Testing Population

Reading

Percentage Early On Grade Level or Higher						
Grade	Fall			Winter		
	Historical	Current	Difference	Historical	Current	Difference
1	20%	15%	5	48%	38%	10
2	32%	25%	7	56%	47%	9
3	51%	47%	4	68%	64%	4
4	35%	33%	2	49%	46%	3
5	35%	35%	0	46%	46%	0
6	35%	35%	0	44%	43%	1
7	40%	39%	1	46%	44%	2
8	41%	41%	0	47%	47%	0

Mathematics

Percentage Early On Grade Level or Higher						
Grade	Fall			Winter		
	Historical	Current	Difference	Historical	Current	Difference
1	13%	9%	4	41%	32%	9
2	17%	12%	5	41%	34%	7
3	20%	12%	8	46%	36%	10
4	32%	19%	13	53%	40%	13
5	38%	26%	12	54%	44%	10
6	36%	28%	8	49%	43%	6
7	30%	25%	5	41%	35%	6
8	26%	23%	3	35%	32%	3

Figure 3: *i-Ready Diagnostic* Placement Level Descriptors

i-Ready Placement Level Descriptors

	[redacted]	Two Grade Levels Below	One Grade Level Below	Early On Grade Level	[redacted]	
Placement relative to grade-Level college and career-readiness standards	Are not close to meeting			Only partially met	Met	
Instructional Recommendations	**Likely need intensive intervention of foundational concepts.** Students who perform below grade level are not likely to be proficient on their state summative test, though it is possible.	**May need intensive intervention of material that is two grade levels below** to help fill in gaps in students' foundational knowledge.	**May benefit from review or remediation of material that is one grade level below.**	**Will benefit from on-grade level instruction** to help them meet the expectations of college- and career-readiness standards for their grade level.	**Mid On Grade Level:**	Will benefit from instruction in late on-grade level topics.
					Late On Grade 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.
					Above Grade Level:	Will benefit from above-grade level instruction.

About the *i-Ready Diagnostic*

The Diagnostic is a computer-adaptive assessment for students in Grades K–12 for Reading and Mathematics that provides valid and reliable criterion-referenced and normative scores. The Diagnostic can be administered, typically, at three time points during the school year: fall, winter, and spring.

In addition to a scale score and a norm-referenced percentile-rank score, the Diagnostic provides five criterion-referenced Grade-Level Placements: Mid or Above Grade Level, Early On Grade Level, One Grade Level Below, Two Grade Levels Below, and Three or More Grade Levels Below. Unlike normative scores, these placement levels articulate the high expectations students must achieve to be considered as having attained grade-level knowledge and skills. These placement levels are designed to help educators understand what level of instruction students are prepared for across the school year.

References

Catalano, F. (2020). Learning loss is everywhere but how do the results compare? https://www.edsurge.com/news/2020-12-14-learning-loss-is-everywhere-but-how-do-the-reports-compare

Council of Chief State School Officers and Council of the Great City Schools (2014). Commitments from CCSSO and CGCS on high-quality assessments. https://ccsso.org/sites/default/files/2017-11/CSSOCGCSAssessmentCommitments10152014%20%281%29.pdf

Council of the Great City Schools (2020). Addressing unfinished learning after COVID-19 school closures. https://www.cgcs.org/cms/lib/dc00001581/centricity/domain/313/cgcs_unfinished%20learning.pdf

Curriculum Associates (2020). A window into the digital world. https://www.curriculumassociates.com/products/i-ready/usage-trends-digital-divide

Curriculum Associates (2020). Overcoming the digital divide: Distance learning successes during the pandemic. https://www.curriculumassociates.com/products/i-ready/overcoming-the-digital-divide

Curriculum Associates (2020). Understanding student needs: Early assessment results from fall 2020. https://www.curriculumassociates.com/fall2020data

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (December 2020). COVID-19 and learning loss—disparities grow and students need help. McKinsey & Company. https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help

Gershenson, Seth. (February 2020). Great expectations: The impact of rigorous grading practices on student achievement. Washington, DC: Thomas B. Fordham Institute. https://fordhaminstitute.org/national/research/great-expectations-impact-rigorous-grading-practices-student-achievement

Kogan, V., & Lavertu, S. (January 2021). The COVID-19 pandemic and student achievement on Ohio's third-grade English language arts assessment. http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf

Hernandez, D. (2011). Double jeopardy: How third-grade reading skills and poverty influence high school graduation. https://files.eric.ed.gov/fulltext/ED518818.pdf

Kuhfeld & Tarasawa (2020). The COVID-19 slide: What summer learning loss can tell us about the potential impact of school closures on student academic achievement. NWEA Research. https://www.nwea.org/content/uploads/2020/05/Collaborative-Brief_Covid19-Slide-APR20.pdf

Powell, R., Cantrell, S. C., Malo-Juvera, V., & Correll, P. (2016). Operationalizing culturally responsive instruction: Preliminary findings of CRIOP research. *Teachers College Record, 118*(1), 1–46.

Student Achievement Partners (2020). 2020–2021: Priority instructional content in English language arts/literacy and mathematics. https://achievethecore.org/page/3267/2020-21-priority-instructional-content-in-english-language-arts-literacy-and-mathematics

The Center for Research on Education Outcomes at Stanford University (2020). Estimates of learning loss in the 2019–2020 school year. https://credo.stanford.edu/sites/g/files/sbiybj6481/f/credo_days_of_learning_description_r5.pdf

TNTP (2018). The opportunity myth: What students can show us about how school is letting them down—and how to fix it. https://tntp.org/assets/documents/TNTP_The-Opportunity-Myth_Web.pdf

Will, M. (2020). 6 lessons learned about better teaching during the pandemic. https://www.edweek.org/technology/6-lessons-learned-about-better-teaching-during-the-pandemic/2020/11

The mission of Curriculum Associates is to make classrooms better places for teachers and students.

i-Ready

Built to address the rigor of the new standards, *i-Ready* helps students make real gains. *i-Ready* collects a broad spectrum of rich data on student abilities that identifies areas where a student is struggling, measures growth across a student's career, supports teacher-led differentiated instruction, and provides a personalized instructional path within a single online solution.

To learn more about evidence on the impact of *i-Ready*, please visit CurriculumAssociates.com/Research.

@myiready Curriculum Associates @CurriculumAssoc iReady

 What We've Learned about Unfinished Learning (Research Report No. 2021-03). North Billerica, MA: Author. | 03/21 0K

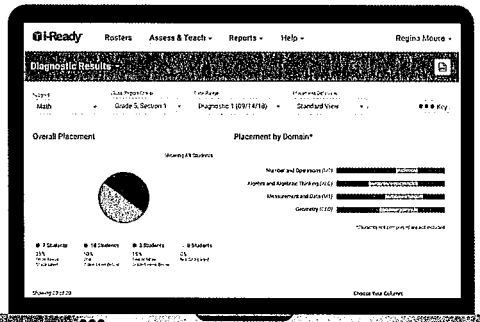
Curriculum Associates

Backed by Practical and Applicable Research

Our assessment and instruction programs are backed by the industry's most practical and applicable research. We conduct timely research in diverse educational settings to provide educators with a meaningful understanding of how our programs can help your students succeed.

i-Ready Diagnostic

Powerful assessment data and intuitive reporting

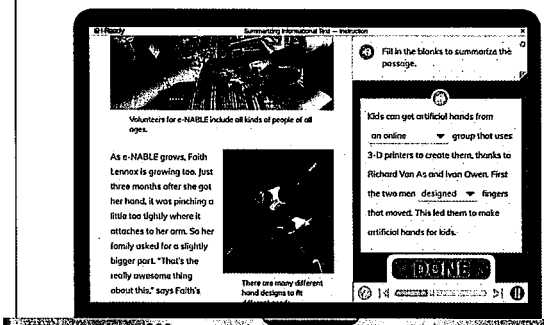


- Pinpoint student needs down to the sub-skill level with a powerful, valid, and reliable adaptive Diagnostic.
- Inform and adjust instruction with timely and actionable data.
- See a complete picture of student performance and growth and eliminate the need for multiple, redundant tests.

i-Ready Diagnostic is strongly correlated to leading state and national assessments. Review our linking studies 4

i-Ready Personalized Instruction

Engaging, individualized digital instruction



- Reach students of all skill levels with digital instruction designed to fill each student's knowledge gaps and help every learner access grade-level content.
- Engage students with interactive instruction designed for modern standards.

Students achieve greater growth with *i-Ready Personalized Instruction*. Review our ESSA evidence 6

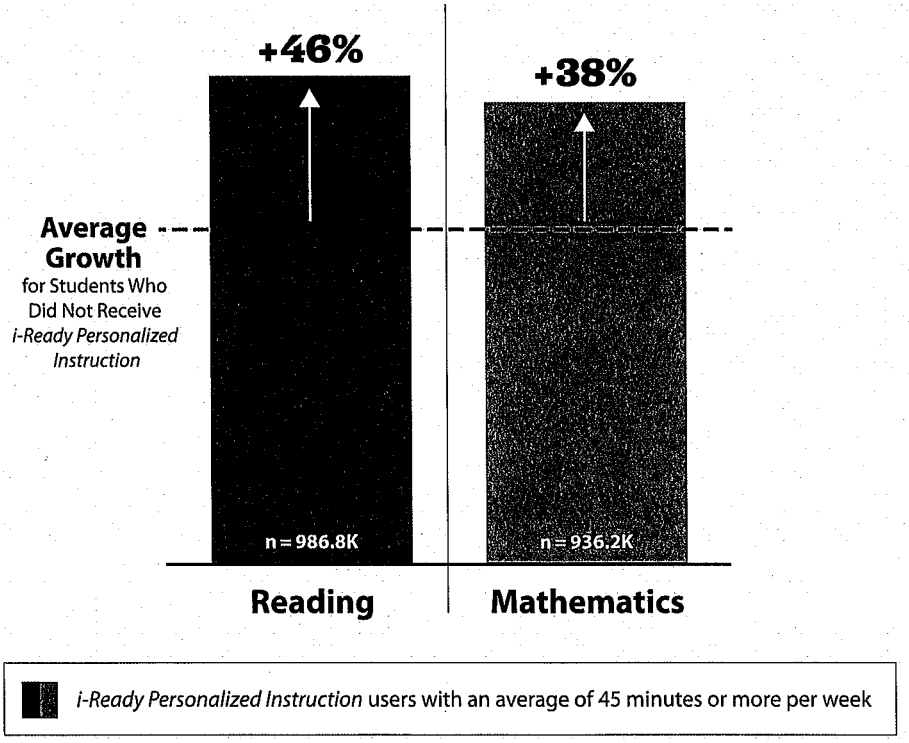
Students receiving *i-Ready Personalized Instruction* experience remarkable gains.

The Curriculum Associates Research team analyzed data from more than one million students who took the *i-Ready Diagnostic* during the 2017–2018 school year. In both Reading and Mathematics, students who used *i-Ready Personalized Instruction* demonstrated substantial learning gains compared to students who did not, and these gains were positive and statistically significant. This large-scale study provides additional support that *i-Ready* is a well-researched program that meets the criteria for "evidence-based" as outlined by the Every Student Succeeds Act (ESSA).

Students achieve greater growth with *i-Ready*.

Students using *i-Ready Personalized Instruction* for an average of 45 minutes or more per subject per week for at least 18 weeks showed statistically significantly greater growth than the average student who did not receive *i-Ready Personalized Instruction* during the 2017–2018 school year.

Score Gains for Students Receiving *i-Ready Personalized Instruction* Relative to Students Not Receiving *i-Ready Personalized Instruction*

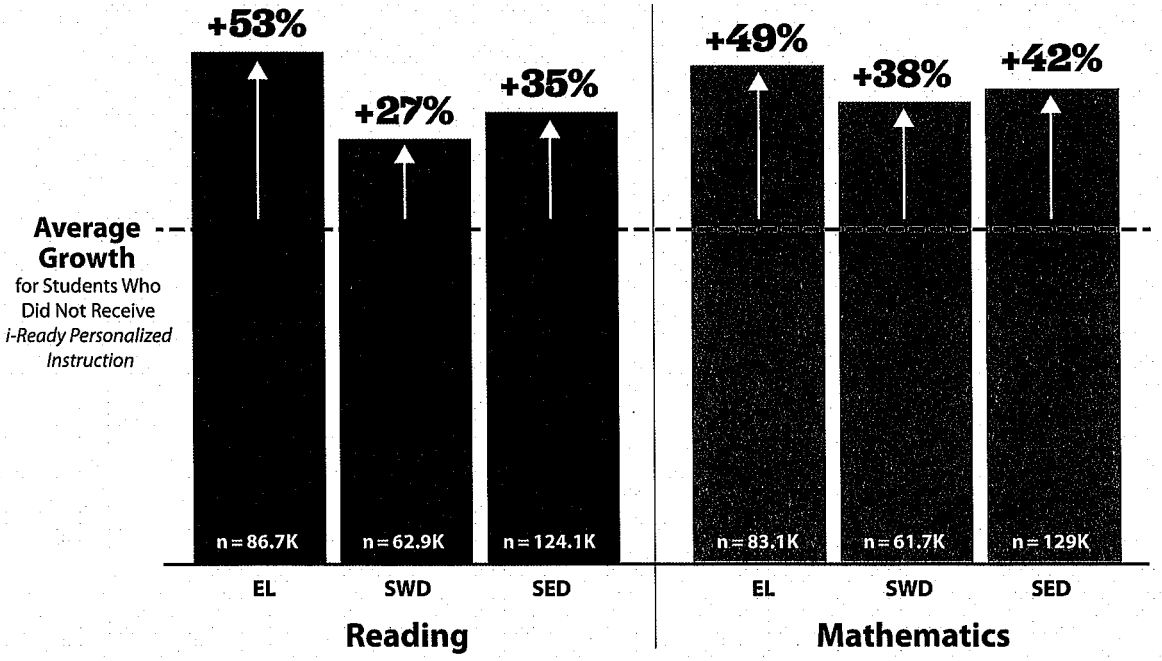


> ***When you start looking at [i-Ready] data and following the trends . . . it's revealing. It's about knowing the truth, and that helps you make better decisions that move schools forward.***
>
> —Melinda Chemin, Reading Coach, FL

i-Ready accelerates growth for key student groups.

Students who are English Learners (EL), students with disabilities (SWD), and students with socioeconomic disadvantages (SED) using *i-Ready Personalized Instruction* all saw statistically significantly greater growth than students from the same subgroups who did not have access to the program during the 2017–2018 school year.

Score Gains for Key Student Groups Receiving *i-Ready Personalized Instruction* Relative to Key Student Groups Not Receiving *i-Ready Personalized Instruction*



Ready Mathematics Blended Core Curriculum
ESSA EVIDENCE

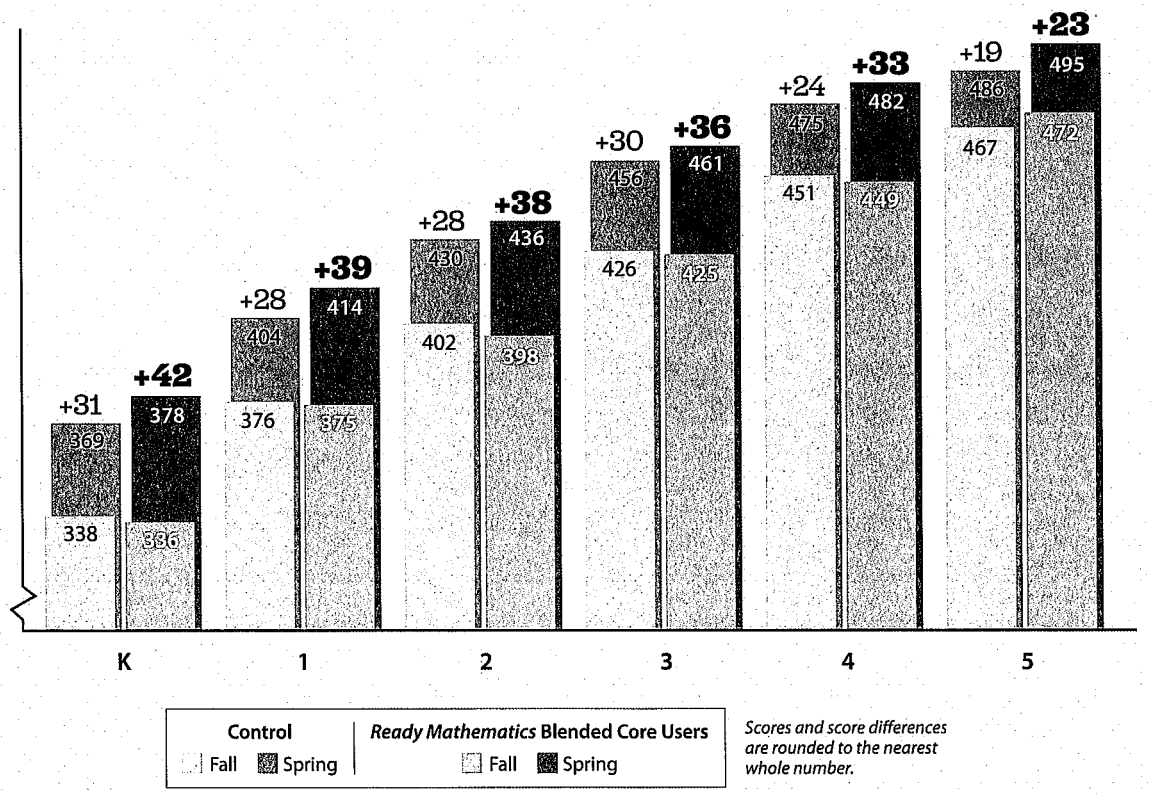
Students using *Ready Mathematics* grow more.

HumRRO, a leading evaluator of educational programs, conducted a study using data from the 2017–2018 school year of more than 21,000 students to understand the impact of the *Ready Mathematics* Blended Core Curriculum* on mathematics achievement for students in Grades K–5. The study found that across grades and diverse student populations, *Ready Mathematics* helped students grow more.

Ready Mathematics, the foundation of *Ready Classroom Mathematics*, drives mathematics achievement.

HumRRO's research found that students using *Ready Mathematics* as their core curriculum along with *i-Ready Personalized Instruction* outperformed comparable student populations using other programs. Because *Ready Classroom Mathematics* curriculum is the next evolution of *Ready Mathematics*, the findings from this study support the efficacy of both programs.

i-Ready Diagnostic (Mathematics) Score Differences



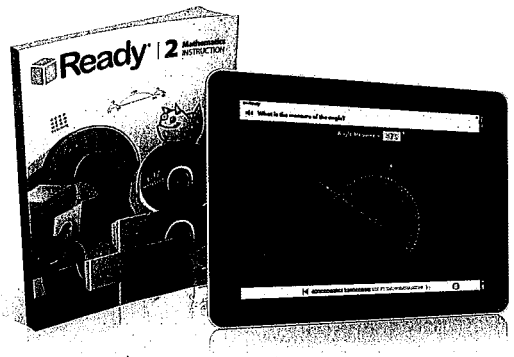
Scores and score differences are rounded to the nearest whole number.

About the Research: For students with comparable starting points, the mean mathematics achievement for the *Ready Mathematics* Blended Core Curriculum group was statistically significantly higher in all Grades K–5 (Hedges' *g* effect sizes range from .17 to .36). The research was designed to meet the What Works Clearinghouse (WWC) evidence standards for quasi-experimental studies as well as the ESSA Level 2 (Moderate) criteria.

**Ready Mathematics* Blended Core Curriculum includes *Ready Mathematics*, *i-Ready Personalized Instruction*, and the *i-Ready Diagnostic*.

Ready Mathematics Blended Core Curriculum

Rigorous, discourse-based mathematics instruction

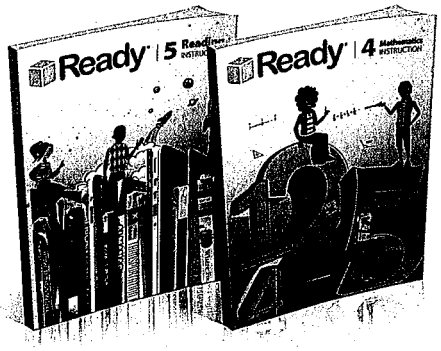


- Develop students' procedural fluency and conceptual understanding through reasoning, modeling, and discussion.
- Reach all skill levels with customized instruction powered by assessment data to address whole class and individual student needs.

***Ready Mathematics* drives achievement. Review our ESSA evidence 8**

Ready Reading and Mathematics

Proven whole class and small group instruction



- Provide rigorous whole class instruction with materials that support teachers and engage students in mastering content standards.
- Access a digital collection of both on- and off-grade level instructional resources to facilitate differentiated instruction for students of all performance levels.

Schools using *Ready* perform better on state assessments. Review our ESSA evidence 9

Grounded in Research
Meet the experts behind *i-Ready* and *Ready* **10**

Ohio County Schools - ARP ESSER Plan
4. OdysseyWare Credit Recovery System

CREDIT RECOVERY

Increase graduation rates and improve student outcomes leveraging flexible, online courses from Odysseyware®.

CREDIT RECOVERY FEATURES

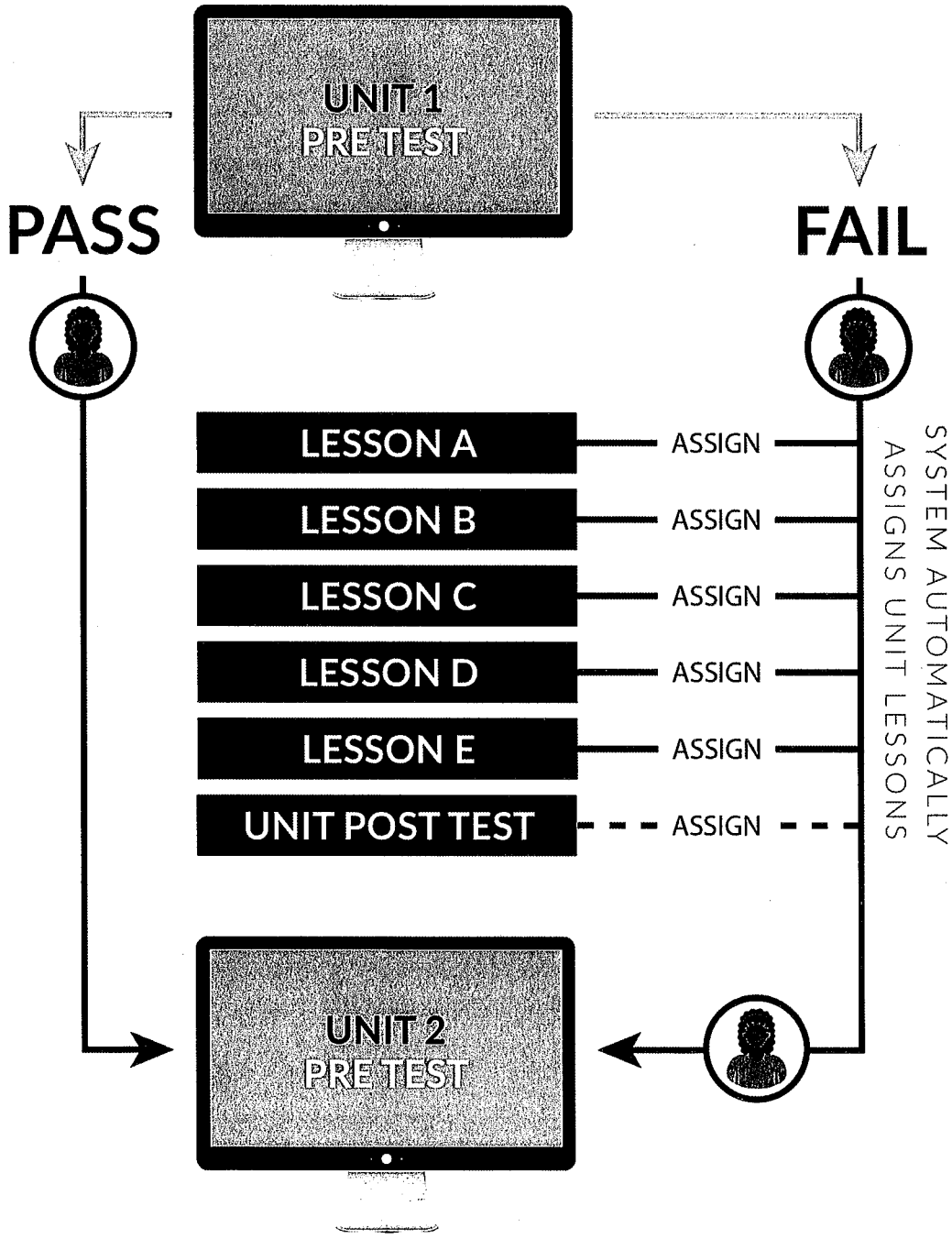
Educators can give students the ability to earn credits for failed or incomplete courses without having to repeat the entire course, enabling students to complete the work at their own pace and focus on just the concepts they have not yet mastered.

Credit Recovery Overview (https://glndocs.s3.amazonaws.com/odw/marketing-resources/pdfs/SSFL_1005_072017_Credit%20Recovery_F-HR.pdf)

CREDIT RECOVERY (CRX) MODE

When Odysseyware curriculum is used in Credit Recovery (CRx) Mode, students are assigned a pre-test at the beginning of each unit to determine if they are already competent in concepts addressed in the unit. If the student achieves a score at or above the predetermined CRx Pass Threshold, the program automatically skips the rest of the unit and the student takes the pretest for the following unit. If not, the system assigns the lessons within the unit to the student.

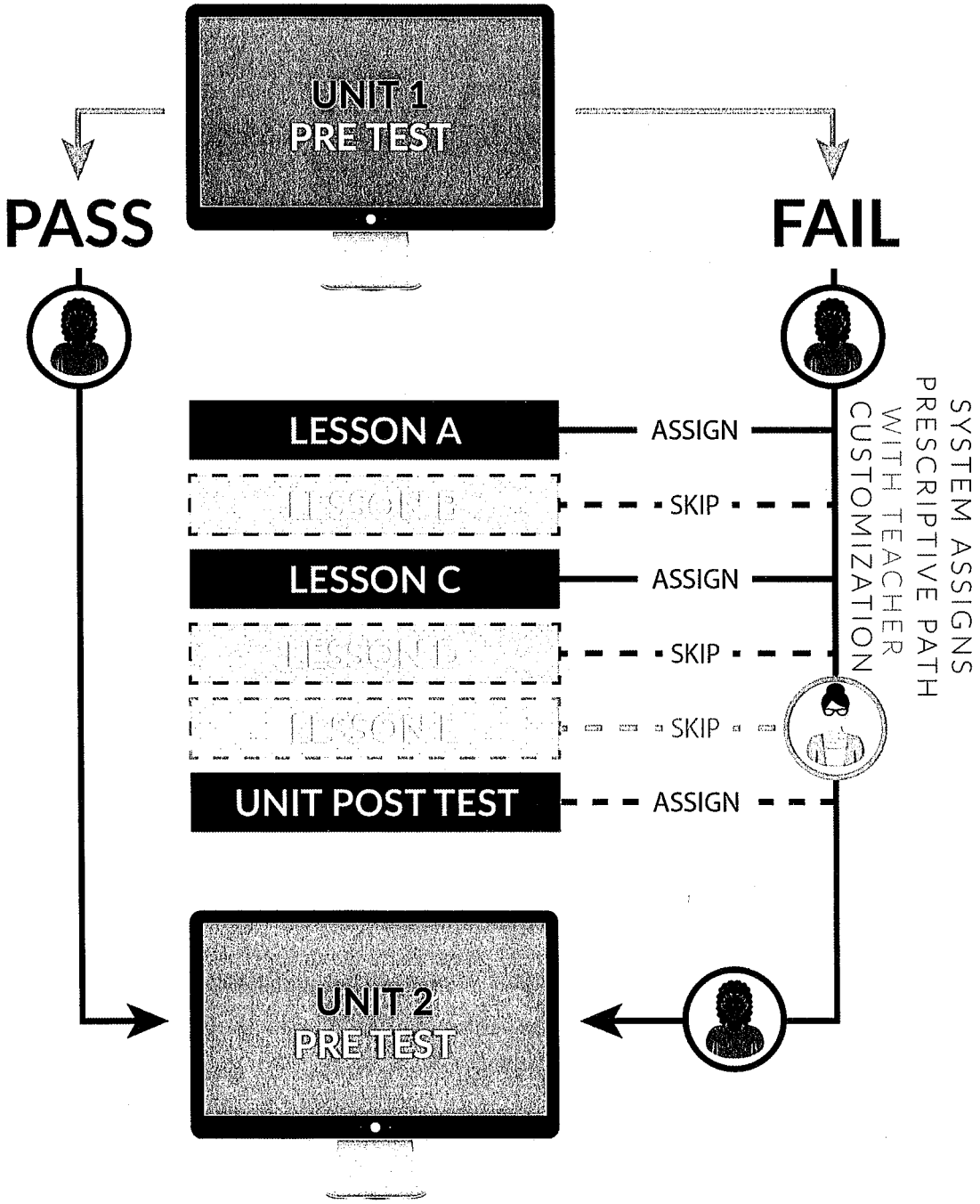
CRx Mode



FLEX CRX MODE

Using Flex CRx Mode, a student who fails the unit pre-test but shows mastery of certain concepts will be able to skip those lessons in the unit. The software automatically creates a custom learning path that assigns only the specific lessons within that unit that contains content the pre-test assessed as needing remediation. The teacher then has the ability to further customize each student's learning path.

Flex CRx Mode



FLEX ASSESSMENTS

The Flex Assessment (../curriculum/customization) feature automatically removes the assessment questions for the lessons skipped in Flex CRx mode, as well as adjusts to match any changes made by the teacher to the lesson sequence.

SCAFFOLDED INSTRUCTIONAL SUPPORTS

Students have a variety of embedded instructional supports (../curriculum/instructional-supports) at their fingertips in every lesson, including literacy, fluency, and audio scaffolds. They also receive immediate feedback throughout lessons and assessments to guide their learning.

CUSTOMIZATION

All courses are fully customizable (../curriculum/customization), so instructors can rearrange/remove content, add lessons from another Odysseyware course, and/or add their own unique web-based content.

Odysseyware Spotlight: Racine Unified School District

Racine Unified School District

Video Testimonial

ADDITIONAL HIGHLIGHTS

Additional benefits of using Odysseyware to fuel your dropout prevention strategy and reach your at-risk/struggling students include:

Flexible Delivery

Odysseyware curriculum can be delivered in a variety of settings, including fully online, as part of a blended learning model, after school, lab setting, alternative program, and/or over the summer.

Anytime, Anywhere Access

Real-Time Dashboard And Reports

Certified Teachers

Moreno Valley, CA

Thomasville, GA

WHAT OTHERS ARE *Saying*

The Odysseyware implementation is fostering a level of learning where students are learning to work

and retain the materials,
creating educational growth for
the student and our school.

Earl Post

Credit Recovery
Facilitator, Braden River
High School, FL


(../)

877.7CLICKS


(https://www.facebook.com/Edgenuity)


(https://twitter.com/Edgenuityinc)

(https://www.linkedin.com/company/edgenuity-inc-/?trk=company_name)

(https://www.youtube.com/channel/UCnlauGncYaQxbMnWxeIabUA)