U.S. Department of Education
November 2015

Evaluation of Response to Intervention Practices for Elementary School Reading

Executive Summary

Rekha Balu
Pei Zhu
Fred Doolittle
Ellen Schiller
Joseph Jenkins
Russell Gersten

ies INSTITUTE OF EDUCATION SCIENCES

NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

This page has been left blank for double-sided copying.

Evaluation of Response to Intervention Practices for Elementary School Reading

Executive Summary

November 2015

Rekha Balu
Pei Zhu
Fred Doolittle
MDRC

Ellen Schiller
SRI International

Joseph Jenkins
University of Washington

Russell Gersten
Instructional Research Group

Jonathan Jacobson
Project Officer
Institute of Education Sciences

NCEE 2016-4000
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education

Arne Duncan

Secretary

Institute of Education Sciences

Ruth Curran Neild

Deputy Director for Policy and Research Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance

Joy Lesnick

Acting Commissioner

November 2015

This report was prepared for the Institute of Education Sciences under Contract ED-04-CO-0111/0003 with MDRC. The project officer is Jonathan Jacobson in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

This report is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should read: Balu, Rekha, Pei Zhu, Fred Doolittle, Ellen Schiller, Joseph Jenkins, and Russell Gersten (2015). *Evaluation of Response to Intervention Practices for Elementary School Reading* (NCEE 2016-4000). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

This report is available on the Institute of Education Sciences website at http://ies.ed.gov/ncee.

Alternate Formats: Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

Acknowledgments

This study represents a collaborative effort among the authors, our colleagues at MDRC and SRI International, and Institute of Education Sciences (IES) staff. This study also benefits from collaboration between the study team and more than 1,200 schools included in the evaluation, plus many other schools that participated in our initial screening of sites. We appreciate the willingness of school staff to provide information on how they deliver reading instruction to students, especially those who are not yet meeting reading standards. Staff in participating schools and districts completed surveys, provided student records on reading assessments and instruction, and supported the study team as it fielded reading assessments. Without their help, the study would not be possible.

We also benefited from the advice of the members of our Technical Working Group who provided expert insights into reading assessments, instruction, the response to intervention (RtI) approach, and research methodology. Members of this group over the course of the project have included Donald Compton of Vanderbilt University, Carol Connor of Arizona State University, Judy Elliot formerly of Los Angeles Unified School District and now an independent consultant, David Francis of the University of Houston, Paul McDermott of the University of Pennsylvania, Rollanda O'Connor of the University of California at Riverside, Sean Reardon of Stanford University, Amy Sichel of the Abington Pennsylvania School District, Jeffrey Smith of the University of Michigan, Deborah Speece of the Virginia Commonwealth University, and Sharon Vaughn of the University of Texas. We also wish to extend our thanks to reviewers of the report, including Jose Blackorby, Gordon Berlin, Howard Bloom, William Corrin, John Hutchins, Michelle Maier, Shira Mattera, and Marie-Andrée Somers.

Many people at four organizations played a crucial role in conducting this study. Many of the MDRC staff members involved in processing the data, conducting the analysis, and preparing this report are named below in the listing of project staff. We especially want to acknowledge the contributions of Ivonne Garcia throughout the project for her role in skillfully organizing, checking quality control, and merging complex data from multiple sources; and of Amanda Ferrandino for serving as our report coordinator and producing multiple versions of numerous exhibits and chapters for the report as we sought the best way to convey the findings. In addition, Margaret Bald, Alison Black, Stephanie Cowell, Sonia Drohojowska, Anne Fenton, Kate Gualtieri, Kathy Jenkins, Ellie Leahy, Nathan Meyers, Seth Muzzy, Ushapriya Narasimhan, Elizabeth Nelson, Gina Price, Shelley Rappaport, Carolyn Thomas, and Robert Weber played important roles in site recruitment, survey design, project management, and report production. At SRI, project staff including Jake Cornett, Sara Kalb, Camille Marder, Christina Saldarriaga, and Katherine Nagle played important roles in early stages of the project. At the Instructional Research Group, Joseph Dimino, Mary Jo Taylor, and Becky Newman-Gonchar were

important vital contributors to the project. At SRM, Linda Kuhn, Tony Lavender, Julie Berry, Maegan McKee, Daniel Mackin, and Betsy Quicksall played crucial roles in data collection.

Following are the Data Processing and Analysis Team for the Report and Restricted Use File (in alphabetical order):

- Catherine Armstrong
- Colin Bottles
- Nick Commins
- Catherine Corbin
- Albert Fang
- Ivonne Garcia
- Zeest Haider
- Kateryna Lashko
- Sophia Litschwartz
- Heilyn Paulino
- Nick Sherefkin
- Sarah Spell
- Mark van Dok
- Edmond Wong

The Authors

Disclosure of Potential Conflicts of Interest

The research team for this evaluation consists of a prime contractor, MDRC, Inc., of New York City, NY, and three subcontractors: SRI International, of Menlo Park, CA; Instructional Research Group, of Los Alamitos, CA; and Survey Research Management (SRM) Corporation, of Boulder, CO. None of these organizations or their key staff has financial interests that could be affected by findings from the evaluation of the two supplemental literacy interventions considered in this report. No one on the eight-member Technical Working Group, convened by the research team to provide advice and guidance, has financial interests that could be affected by findings from the evaluation.

This page has been left blank for double-sided copying.

Executive Summary

The 2004 reauthorization of the Individuals with Disabilities in Education Act (IDEA) allows states and school districts to use a portion of federal special education funds to provide coordinated early intervening services to students at risk of reading failure or other academic or behavioral problems. One of the primary approaches that has emerged is called "Response to Intervention" (RtI). In the context of this report, RtI incorporates a range of assessment, instruction, and intervention principles, including (1) offering multiple tiers of support for students, depending on the level of reading difficulty they may be experiencing; (2) allocating staff to provide that tiered support to students; and (3) collecting and using data to make instructional and intervention decisions for students throughout the school year.

This study describes these RtI practices and compares their prevalence between two different samples: a *reference sample* of schools representative of elementary schools in the 13 states included in the evaluation and an *impact sample* of 146 elementary schools with three or more years of implementing RtI approaches in reading. In the impact sample, the study research team compared the intensity of services provided to reading groups at different reading levels to measure the extent to which support is more intense for students reading below grade level. For the impact analysis, the study research team estimated effects of assignment to reading interventions for students at the margins of eligibility for those services who read below grade level.

This report provides new information on the prevalence of RtI practices in elementary schools, illustrates the implementation of RtI practices for groups of students at different reading levels, and provides evidence on effects of one key element of RtI: assigning students to receive reading intervention services. The findings show, for the 2011-12 school year, that:

- A majority of schools in the 13-state reference sample (56 percent) reported full implementation of the RtI framework, while a higher proportion of impact sample schools (86 percent) in those states reported full implementation.
- Schools in the impact sample adjusted reading services to provide more support to students reading below grade-level standards than to those at or above the standards.
- For those students just below the school-determined eligibility cut point in Grade 1, assignment to receive reading interventions did not improve reading outcomes; it produced negative impacts.

The rest of the Executive Summary describes the evaluation's policy context and specific research questions, defines key terms and analytic approaches, and explains the findings.

Policy Context and Relation to Previous Research

For school year 2008-09, when this study began its planning and design, 70 percent of districts with elementary schools reported using RtI in reading/language arts.[1] The use of the RtI framework is an outgrowth of a change in approach related to special education policy and the process for identifying children with a Specific Learning Disability (SLD) — the disability category most associated with reading difficulties. The previous eligibility standard required educators to document an "educationally significant discrepancy" between achievement of specific skills (for example, reading performance) and general ability (that is, overall intellectual functioning as measured by an IQ test) that could not be explained by visual, hearing, or motor disabilities; emotional disturbances; mental retardation; or environmental, cultural, or economic disadvantage. The 2004 reauthorization of IDEA forbids states from requiring districts to identify SLD students using a discrepancy approach, and it permits districts to use an SLD identification process based on the child's response to scientific, research-based interventions. The law also allows districts to use up to 15 percent of their IDEA Part B special education funds to develop and implement coordinated early intervening services for students not yet identified as needing special education and related services but who need additional academic or behavioral support to be successful in general education classrooms. This funding change allows federal dollars to be used for RtI services.

Over the past 15 years, numerous studies have addressed the effect of interventions delivered to early readers in need of help within an RtI framework. A survey of the recent literature (since 1999) yields 27 studies that report the impact of providing certain types of interventions to students with reading difficulties on a range of reading skill measures. These recent studies support the conclusion of Gersten et al. that well-designed and closely monitored small-group reading interventions could be beneficial to early-grade readers in terms of improving their specific reading skills.[2] The evidence is stronger for first grade than for second or third grades. The effect of such intervention on students' more comprehensive reading skills is less clear. Also not clear is the impact of such interventions if they were to be implemented at a larger scale.

---

[1] M. C. Bradley, Tamara Daley, Marjorie Levin, Fran O'Reilly, Amanda Parsad, Anne Robertson, and Alan Werner, *IDEA National Assessment Implementation Study,* NCEE 2011-4027 (Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2011).

[2] Russell M. Gersten, Donald L. Compton, Carol M. Connor, Joseph Dimino, Lana Santoro, Sylvia Linan-Thompson, and W. David Tilly, "Assisting Students Struggling with Reading: Response to Intervention and Multi-Tier Intervention in the Primary Grades" (Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2009). Website: http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=3.

This evaluation's analysis of RtI implementation and the impact of interventions on reading achievement expands the field's knowledge about RtI in three ways. First, this study describes implementation of RtI practices in multiple states at the school level, unlike previous studies that address RtI adoption at the district or state level. Second, this study describes practices in schools that had adopted RtI on their own and had implemented it for three or more years, rather than for a sample of schools that were monitored by researchers or that received special supports for first-year implementation. Third, while this study's school sample is broader than in earlier studies, the student sample is narrower. Unlike earlier studies, which address the overall effectiveness of RtI, this study's research design answers a question about effective targeting, by comparing the outcomes for students just below and just above the cut point of eligibility for intervention. This approach provides an estimate of the impact of interventions on the students *slightly below* grade-level reading standards, rather than for the full range of students served by interventions. This impact on the marginally eligible student served is important for assessing the effective targeting of intervention resources, but it does *not* assess whether the RtI framework as a whole is effective in improving student outcomes or whether reading interventions are effective for students *well below* grade-level standards.

Research Questions and Study Overview

This study answers three sets of major research questions:

1. **Comparison of practices between school samples.** How did the prevalence of RtI practices differ between a representative "reference" sample of schools and schools selected for the impact evaluation? To what extent were impact sample schools implementing more RtI practices than the reference sample schools? How do special education identification rates in the impact sample compare with rates for the states as a whole?

2. **Comparison of reading services between reading groups at different skill levels.** In impact sample schools (those with three or more years of implementing RtI): To what extent did schools place students in tiers as suggested by earlier RtI models? To what extent did schools adjust tier placement during the school year? To what extent is there variation in how schools organize reading services for specific reading levels? To what extent were services for students reading below grade level more intense than for students reading at or above grade level?

3. **Impacts on reading outcomes of students.** For students who fell just below school-determined standards for each grade on screening tests: What were the effects on reading achievement of actual assignment to receive reading intervention services (in addition to core instruction)? What is the extent of variation in estimat-

ed impacts across RtI schools? How is the estimated impact associated with certain school features or student characteristics?

Key Terms, Sample Selection, and Research Design

"Intervention" in this report generally refers to additional support for students who have difficulty reading. RtI schools may place students in reading groups and deliver services based, in part, on students' scores on *screening tests,* which are brief assessments of skills considered necessary for reading, such as word identification and letter sounds. In this way, how students score on screening assessments is related to the services they receive. Screening tests differ from the end-of-year comprehensive reading tests, which evaluate a wider variety of reading skills.

- **Tier 1.** "Tier 1" refers to the core instruction that all students receive. The National Reading Panel has recommended that reading instruction in the early grades focus on five reading components: phonemic awareness, phonics, fluency, reading comprehension, and vocabulary.[3] Tier 1 is intended to prevent the risk of reading failure for as many students as possible and to avoid inappropriate referrals to special education. Core instruction usually occurs during a period called the "core reading block." Students who receive only core instruction generally read at or above grade level.

- **Tiers 2 and 3.** Students placed in Tier 2 or Tier 3 receive intervention services in addition to Tier 1 core instruction services. Students in Tier 2 generally read at least *somewhat below* grade level based on screening tests. The typical mechanism that schools use to deliver services to students in Tier 2 is an adult-led small reading group — an approach that could be used to provide small-group instruction during the core reading block as well as additional intervention services. Students in Tier 3 generally read *far below* grade level or have not responded to Tier 2 interventions, and they may be assigned to more intensive interventions (characterized by smaller group size, additional intervention time, or both). To address the second research question of how services differ depending on students' reading skills, the descriptive analysis compares services received by reading groups at different skill levels: at or above grade level, somewhat below

[3]National Reading Panel, "Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction" (Washington, DC: U.S. Department of Health and Human Services, National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development, 2000).
Website: http://www.nichd.nih.gov/publications/pubs/nrp/pages/smallbook.aspx?renderforprint=1.

grade level, or far below grade level. This analysis compares features of groups receiving small-group reading instruction during the core reading block as well as features of groups receiving reading intervention services.

Schools purposively selected for inclusion in the impact study reported at least three years' experience with RtI implementation and are referred to as the "impact sample." The impact sample was selected to include schools implementing all of the following practices no later than 2009-10:

- Use of three or more tiers of increasing instructional intensity to deliver reading services to students
- Fielding of screening assessments of all students (universal screening) at least twice a year
- Use of data for placing students in Tier 2 or Tier 3
- Use of progress monitoring (beyond universal screening) for students reading below grade level to determine whether intervention is working for students placed in Tier 2 or Tier 3

Schools in the impact sample provided information about the score on a screening test that they used to determine a student's placement in Tier 2 or Tier 3. This score, referred to as a "cut point" (or "cut score"), allowed the study research team to determine whether schools followed a consistent quantitative decision rule for tier placement.

To address the third research question, which assesses the relationship between assignment of students to Tier 2 or 3 to receive intervention services and their reading outcomes, the study uses a *Regression Discontinuity (RD) design.* This quasi-experimental research design provides a causal impact estimate when random assignment is not possible. Schools participating in the impact evaluation used students' fall screening test scores to determine their assignment to intervention. Students whose scores are below the predefined cut point typically receive treatment (Tier 2 or 3 intervention services) in addition to core instruction, and those whose scores are at or above the cut point typically receive only core instruction (Tier 1). Students at or near either side of the cut score are expected to be comparable to each other, and they form the treatment and comparison groups for the impact analysis. Most but not all of the students with scores just below the cut point were placed in Tier 2, while most of the students with scores just above the cut score were placed in Tier 1.

Samples and Data

Different samples and data were used to answer each of the three main research questions. To study different RtI practices across schools, the *impact sample* of 146 unique schools across 13 states was compared with a random sample of 100 elementary schools in each of the same 13 states (referred to as the *reference sample*),[4] based on data collected through a school administrator survey.

To compare reading services provided to reading groups at different skill levels, in particular for students reading below grade level (or students receiving Tier 2 or 3 intervention services) and students reading at or above grade level (or students in Tier 1 only), the study research team collected survey data in spring 2012 from reading teachers and staff who provided reading intervention services. The survey data report information about reading services provided to reading groups of all reading levels by group, not by individual student.

Finally, to analyze the impacts of assignment to intervention on students' reading achievement, the study research team compared the difference in reading outcomes between students whose fall screening test scores were just above the cut point for Tier 2 intervention set by the schools and those whose scores were just below, based on the RD design described above. This design determines that the impact findings are applicable not to everyone receiving either Tier 2 or Tier 3 intervention, but only to students whose fall screening scores were close to the cut point. Students close to the cut point are largely Tier 2 students but also include a small portion of Tier 3 students.

To carry out this design, the study research team collected individual-level fall screening test scores and resulting tier placements for fall and winter of the 2011-12 school year for all students in grades 1-3 in the 146 impact sample schools.[5] The reading achievement outcomes used in the impact analysis vary by grade. The study research team administered the Early Childhood Longitudinal Study, Kindergarten Cohort, of 2011 (ECLS-K: 2011) Reading Assessment to first graders in the sample to measure their comprehensive reading skills; it also administered a Sight Word Efficiency test (the Test of Sight Word Reading Efficiency, 2nd edition, or TOWRE2) to measure students' decoding fluency skill in Grades 1 and 2. For third-graders, individual-level scores from the spring state reading achievement tests were used to measure students' comprehensive reading skills.

[4]Of the 1,300 schools randomly sampled for the reference sample, 1,105 (or 85 percent) completed the school administrator survey that principals of impact sample schools also received.

[5]Note that number of schools eligible for the impact analysis varies by grade, with 119 eligible schools for Grade 1 analysis, 127 eligible schools for Grade 2 analysis, and 112 eligible schools for Grade 3 analysis.

Summary of Findings

This study reports on services and impacts in the 2011-12 school year — the only year for which data were collected and analyzed. This section reports key findings related to the three types of analysis presented in the report.

Comparison of Practices Between Schools

- **More than half of the reference sample schools in the 13 study states adopted an RtI framework in Grade 1-3 reading for the 2011-12 school year. A higher proportion of impact sample schools than reference sample schools reported full implementation of an RtI framework for Grade 1-3 reading.**

Figure ES.1 shows that a majority of schools in both samples reported full implementation of an RtI framework for reading: 86 percent of impact sample schools, compared with 56 percent of reference sample schools. Because the impact schools were screened for experience with RtI implementation, this difference is to be expected. The study research team also examined the frequency of specific practices that correspond to three key aspects of an RtI framework, described below.

Multiple Tiers of Reading Instruction and Intervention

Although about two-thirds (68 percent to 70 percent) of both school samples reported offering more than 90 minutes per day of core reading instruction, the frequency of offering intervention differed between the two samples. Impact sample schools were more likely to report providing time for Tier 2 intervention at least three times a week than were reference sample schools (97 percent and 80 percent, respectively). Impact sample schools were also more likely to report providing time for Tier 3 intervention at least five times a week than were reference sample schools (68 percent and 47 percent).

Allocation of Staff

Impact sample schools were more likely than reference sample schools to allocate staff to assist teachers with using data (88 percent and 72 percent, respectively) and with reading instruction (69 percent and 56 percent).

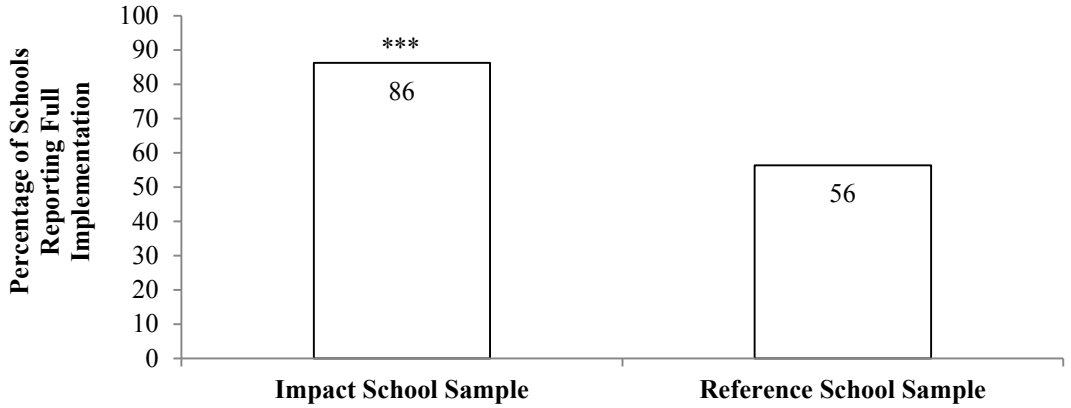
Use of Data to Inform Decisions

Among impact sample schools, 83 percent conducted universal screening assessments of students at least twice a year, compared with 59 percent of reference sample schools. Impact

The Response to Intervention (RtI) Evaluation

Figure ES.1

Full Implementation of RtI in Reading in Grades 1-3



SOURCE: School survey.

NOTES: The survey defined RtI as a "multistep approach to providing early and progressively intensive intervention and monitoring within the general education setting." Respondents could answer that RtI was "fully implemented," "partially implemented," or "not implemented" in reading for each grade. This exhibit reports the percentage of respondents reporting that RtI was "fully implemented" for each of Grades 1, 2, and 3 for which the school responded. Percentages reflect rounding. The statistical significance is indicated as follows: *** at the $p \leq 0.001$ level, ** at the $p \leq 0.01$ level, and * at the $p \leq 0.05$ level.

sample schools were also more likely to follow a prescribed sequence of steps to respond to students who read below grade-level benchmarks (95 percent, compared with 88 percent for reference sample schools). Impact sample and reference school samples were not significantly different in their use of data to monitor student progress following implementation of reading interventions for students suspected of having a Specific Learning Disability.

Comparison of Reading Services Between Reading Groups at Different Skill Levels

- **Impact sample schools followed RtI practices of adjusting student tier placement during the 2011-12 school year. In Grade 1, about three-fourths of students remained in the same reading tier, and one-fourth of students moved between tiers, from fall to winter.**

As shown in Figure ES.2, 59 percent of students in Grade 1 in impact sample schools were placed in Tier 1 as their highest tier in fall 2011. (Results are similar for other grades.) Fewer students were placed in Tier 2 or 3 as the highest tier in which they received services — 25 percent and 16 percent, respectively. This arrangement reflects that Tier 3 was typically reserved for students who had not responded to Tier 2 interventions, although some students were placed directly in Tier 3 in the fall. The majority of students placed in Tier 1 or Tier 3 remained there in the winter: 86 percent of students who began in Tier 1 remained in Tier 1, and 65 percent of students in Tier 3 in the fall remained in Tier 3 in the winter. In contrast, about half the students initially assigned to Tier 2 in the fall remained in Tier 2 in the winter, while the other half moved either to Tier 1 or Tier 3. Across all tiers in Grade 1, 74 percent of students remained in the same reading tier.

The stability of tier placement for the majority of students was coupled with movement to different tiers for other students. These patterns, as well as school reports of the types of data they used to make placement decisions, indicate that schools used screening data to adjust students' tier placement.

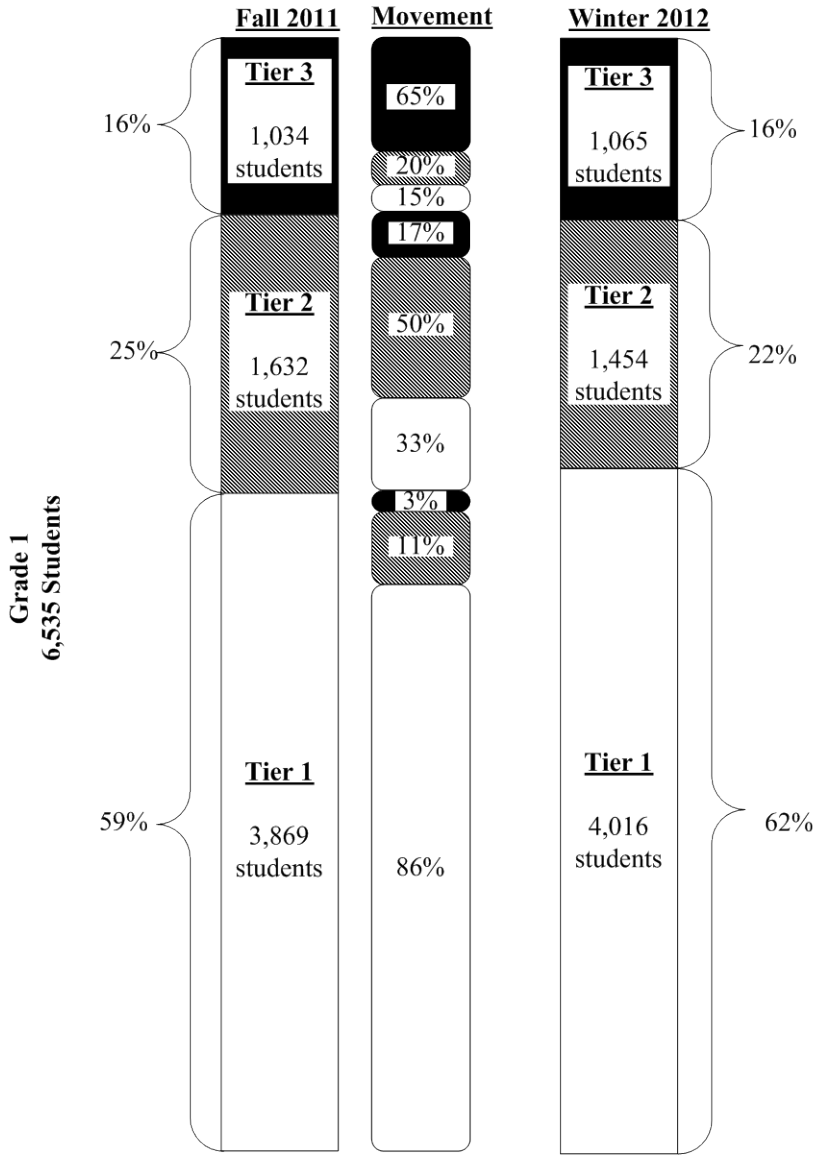
- **Impact sample schools varied in how they organized and delivered reading group services, in some ways differing from descriptions of RtI in prior literature.**
    - **In Grade 1, 45 percent of schools provided intervention services to some groups of students at *all* reading levels, rather than only for reading groups *below* grade level.**
    - **In Grade 1, 67 percent of schools provided at least some reading intervention *during* the core reading block, rather than only *in addition to* the core.**

Although all impact sample schools complied with RtI implementation criteria, some schools showed variations on three aspects of RtI implementation described in prior literature. First, prior studies that designed or monitored the delivery of Tier 2 or 3 intervention services generally served only students reading *below* grade level. In contrast, 45 percent of schools in the impact sample offered reading intervention services to at least some students reading at or above grade level, as well as to those reading below grade level. However, these schools did not necessarily provide intervention services for *all* students at or above grade level. (Results are similar across grades; discussion here focuses on Grade 1.)

Second, previous studies of small-group intervention services often designed intervention as supplemental services that occurred *in addition to* the core reading block time. This study,

Figure ES.2

Student Distribution, by Tier, and Highest Tier Movement



SOURCES: Fall 2011 and winter 2012 tier placement data.

NOTES: Students placed in Tier 1 typically receive only core reading instruction; those placed in Tiers 2 and 3 typically receive core reading instruction plus intervention services. Tier assignment occurs based on results from screening assessments conducted in the fall and winter. Each segment is shaded to represent the proportion of students who remain in that same tier between fall and winter or who move to a different tier (shown in different shading). The Grade 1 school sample size was restricted to 89 schools that had at least one student in each of Tier 1, Tier 2, and Tier 3 in both fall and winter.

in contrast, found that 69 percent of schools in the impact sample offered at least some intervention services *during* the core. In such schools, intervention may have displaced instruction time and replaced some small-group or other instruction services with intervention services. As a result, reading intervention services may have been different from, but not necessarily supplemental to, core reading instruction.

Third, in contrast to more controlled studies of RtI that have relied on non-classroom teaching staff to provide intervention services, the current study included intervention services provided by whoever was designated by schools to provide these services. This study found that, even in schools using the more traditional model of providing intervention services only to readers below grade level, classroom teachers played an additional role and provided intervention services to 37 percent of those groups in Grade 1. These results suggest that impact sample schools adapted time and staff resources to address student needs within an RtI framework.

- **Schools increased the intensity of both small-group instruction during the core and intervention services offered to reading groups below grade level relative to groups reading at or above grade level: group size was smaller, and instruction time was longer. A larger percentage of intervention groups that were below grade level than above it addressed phonics and phonemic awareness.**

The study research team examined whether schools provided more intense services to groups of students reading below grade level than to groups reading at or above grade level, by looking at differences in small-group instruction services during the core reading block (provided by teachers to all students in the class), as well as at differences in reading intervention services delivered either during or outside the core (provided by either teachers or interventionists for students in need of targeted reading support). Results are similar across grades; discussion here focuses on Grade 1.

One way that schools provided more intense services was by reducing the size of groups receiving either instruction or intervention services. For small-group instruction during the core reading block in Grade 1, groups for readers below grade level served about one fewer student than groups reading at or above grade level. For reading intervention services in schools that intervened for groups at all reading levels in Grade 1, there were 1.5 fewer students in intervention groups below grade level than in intervention groups at or above grade level.

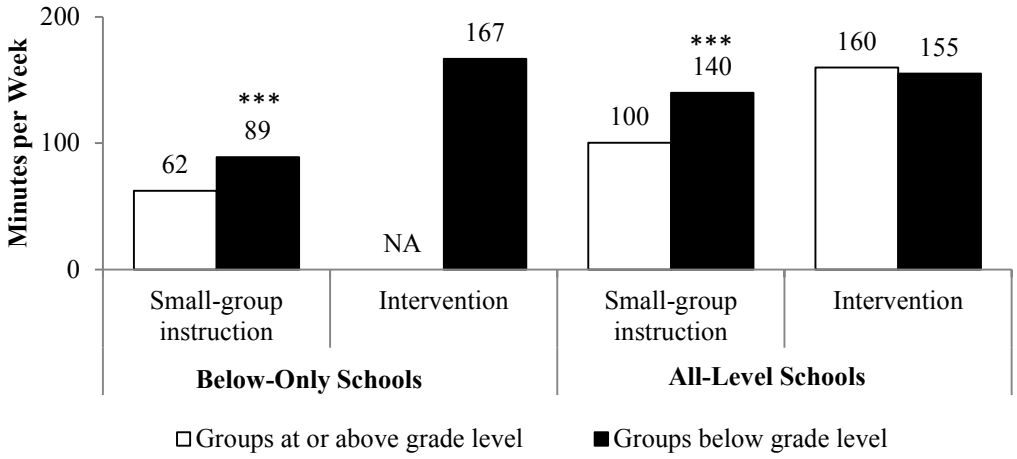
Weekly small-group instruction time during the core in Grade 1 was about 43 percent longer (27 minutes) for groups below grade level than for those at or above grade level, as shown in Figure ES.3. In schools that provided intervention only to readers below grade level in

The Response to Intervention (RtI) Evaluation

Figure ES.3

Service Contrast for Minutes per Week: Difference Between Groups At or Above Grade Level and Below Grade Level in Below-Only and All-Level Schools, for Grade 1

Grade 1



SOURCES: Teacher survey and interventionist survey.

NOTES: "Small-group instruction" refers to services provided by teachers during the core reading block to all students. Intervention services are provided by either teachers or interventionists to students needing targeted reading support, either during or outside the core reading block. The Below-Only school sample represents schools that have at least one of either a Somewhat Below or a Far Below grade-level group receiving intervention services. The All-Level school sample represents schools that have at least one At or Above grade-level group receiving intervention services and at least one of either a Somewhat Below or a Far Below grade-level group (a below-grade-level group) receiving intervention services. No tests were performed between intervention groups in Below-Only schools, which do not provide intervention to At or Above grade-level groups. Means reflect rounding.

Statistical significance is indicated as follows: *** at the $p \leq 0.001$ level, ** at the $p \leq 0.01$ level, and * at the $p \leq 0.05$ level.

the corresponding grade, those groups received 89 minutes per week of small-group instruction time, compared with 62 minutes for groups at or above grade level. In schools that provided intervention services to all reading levels in the corresponding grade, weekly small-group instruction time during the core was 140 minutes per week for groups below grade level, compared with 100 minutes for groups at or above grade level. Unlike the differences in weekly small-group instruction time during the core, the difference in time provided to *intervention*

groups serving students reading below grade level, compared with those reading at or above grade level, is not statistically significant in schools that provided services to all reading levels.

The reading skills that were addressed differed by the reading level of the group. (Results are similar across grades; discussion here focuses on Grade 1.) While 90 percent to 92 percent of groups below grade level for small-group instruction during the core reading block included content on phonics, about half (46 percent to 52 percent) of groups at or above grade level included that content. Among both small groups meeting during the core and reading intervention groups, 70 percent or more of groups both at or above and below grade level included content on fluency, reading comprehension, and vocabulary, regardless of whether the group served students reading below grade level or those reading at or above grade level. These findings suggest that small reading groups and intervention groups focused on multiple skills but that the more elemental skills of phonics were more likely to be addressed by small groups reading below grade level than by small groups at or above grade level.

Impacts on Reading Outcomes of Students

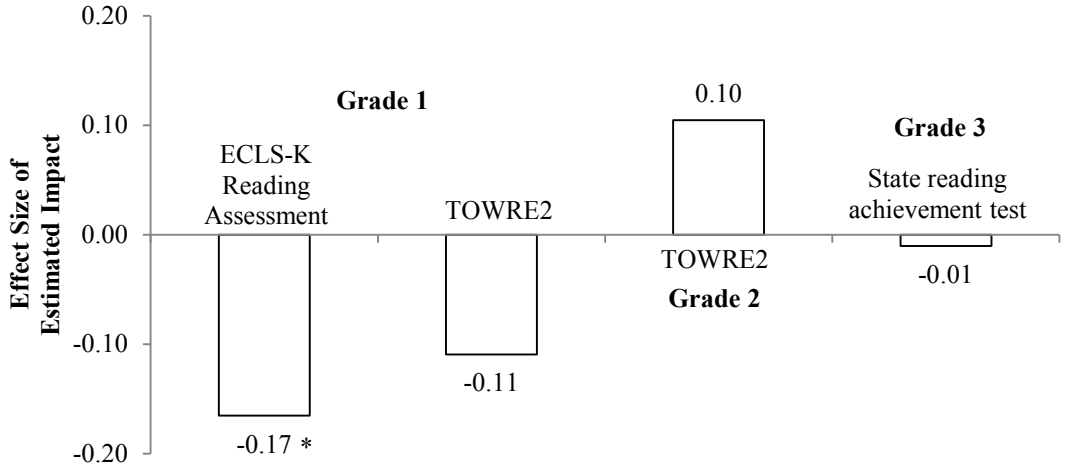
- **Assignment to Tier 2 or Tier 3 intervention services in impact sample schools had a negative effect on performance on a comprehensive reading measure for first-graders just below the Tier 1 cut point on a screening test. The estimated effects on reading outcomes in Grades 2 and 3 are not statistically significant.**

Figure ES.4 presents the estimated effects across four outcomes and three grade levels. The height of each bar in the figure represents the magnitude of the estimated effect, and an asterisk indicates that an estimated effect is statistically significant at the 5 percent level. The study-administered tests were the ECLS-K:2011 comprehensive reading measure, used in Grade 1, and the TOWRE2 measure of decoding fluency, used in Grades 1 and 2. Data from state reading tests provided outcomes for Grade 3 students. Figure ES.4 shows that the estimate for the effect of assignment to Tier 2 or Tier 3 intervention on the ECLS-K Reading Assessment measure is –0.17 standard deviation and is statistically significant (p-value = 0.002). For students who were close to the cut point and were assigned to receive intervention, a negative effect of this magnitude is equivalent to approximately one-tenth of a year less learning than what they would have achieved had they not been assigned to intervention. The estimate for the effect of treatment assignment on the TOWRE2 Sight Word Efficiency test for first-graders close to the cut point is –0.11 standard deviation and is not statistically significant (p-value = 0.057); for second-graders close to the cut point, the estimated impact is +0.10 standard deviation and is not statistically significant (p-value = 0.084). The estimated impact on the state reading achievement test for third-graders in the vicinity of the cut point is –0.01 standard deviation and is not statistically significant (p-value = 0.823).

The Response to Intervention (RtI) Evaluation

Figure ES.4

Estimated Impacts of Assignment to Tier 2 or Tier 3 Intervention Services for Students Within Optimal Bandwidth, by Grade and Outcome Measure



SOURCES: Study-administered ECLS-K Reading Assessment scores for Grade 1; study-administered TOWRE2 test scores for Grades 1 and 2; state reading achievement test scores from district records for Grade 3; fall screening scores and student tier placement data from schools in the sample; student demographic data from district records.

NOTES: The optimal bandwidth defines the sample of students to be used in the impact regression to best balance the trade-off between bias and precision. The optimal bandwidth for each grade and outcome measure was pre-selected using the algorithm described in Imbens and Kalyanaraman (2012). See Appendix E for more details.

Statistical significance at the $p \leq 0.05$ level is indicated as *.

ECLS-K Reading Assessment is a comprehensive reading measure; TOWRE2 is a decoding fluency exam; the state achievement test is a comprehensive reading measure.

- **The estimated impacts of reading interventions on reading outcomes vary significantly across schools. This is true for all four outcomes across three grade levels.**

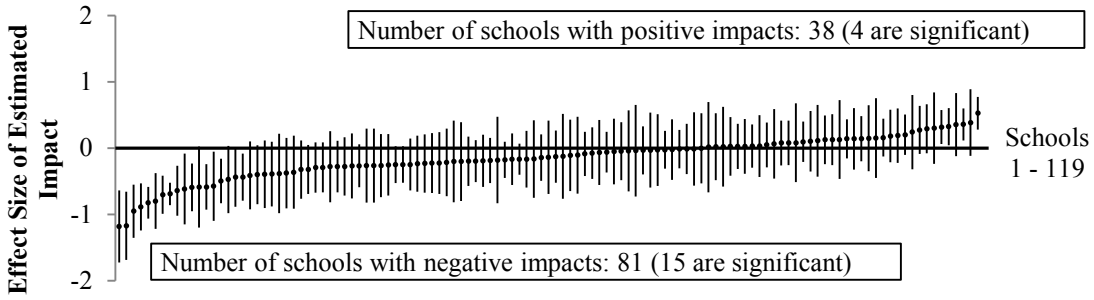
Figure ES.5 presents results for the Grade 1 ECLS-K Reading Assessment comprehensive reading measure to illustrate the extent and significance of impact variation across schools. The figure plots the estimated impact of assignment to intervention on Grade 1 students' ECLS-K Reading Assessment scores for every RtI school in the study sample. The estimates are ordered by their magnitude. A solid dot represents the impact estimate for each school, and a vertical line running through each solid dot represents the respective 95 percent confidence

The Response to Intervention (RtI) Evaluation

Figure ES.5

Distribution of School-Level Impact Estimates of Actual Assignment to Tier 2 or Tier 3 Intervention Services for Grade 1

ECLS-K Reading Assessment



SOURCES: Study-administered ECLS-K Reading Assessment scores for Grade 1; study-administered TOWRE2 test scores for Grades 1 and 2; state reading achievement scores from district records for Grade 3; fall screening scores and student tier placement data from schools in the sample; student demographic data from district records.

NOTES: The outcome was standardized to have a standard deviation of 1, so impact estimates are reported in effect-size units. A chi-squared test was used to test the statistical significance of the variation in the empirical Bayes impact estimates.

interval of the estimated impact. In this example, the estimated school-level impacts on the ECLS-K Reading Assessment score for Grade 1 range from –1.18 to +0.53 standard deviations in effect size. Of the 119 schools included in the impact analysis for Grade 1, there are 15 schools with significant negative findings and four schools with positive and significant findings. Similar patterns of variation were found for the estimated impacts on the other three reading outcomes. Statistical tests show significant variation in impact estimates across schools — for all four outcomes across three grade levels. This finding indicates that the estimated impact could be more negative or more positive in some schools than others, regardless of the overall average impact estimate.

- **The school-level features and student characteristics examined are not consistently associated with school impacts across grades and reading outcomes.**

Across grades or outcomes, there is no consistent association between the impact estimates and examined school features, which include measures of school-level RtI practices, school context, and composition of the student population. (See Box ES.1 for details.) Specifically, the analysis yielded no statistically significant associations between school features and the impact estimates for the two comprehensive reading measures: Grade 1 ECLS-K Reading Assessment scores and Grade 3 state achievement test scores. There are sporadic associations for the decoding-fluency measure for Grades 1 and 2.

Box ES.1

Exploratory Factors Examined in the RtI Evaluation

School-level RtI practices: Whether a school used single or multiple screening tests to assign students to tiers, the proportion assigned to Tier 2 or Tier 3 intervention services, whether the school provided intervention to at least one group at all reading levels, and the proportion of intervention groups served outside the core reading block

School context factors: Overall school reading performance in a baseline year, eligibility for Title I funds, and use of RtI practices for behavior-related interventions

Composition of the student population: Proportion of students who are male or who were English Language Learners, overage for grade, or low-income status or who had an Individualized Education Program (IEP) on account of a student disability

At the student level, for some outcomes and grades, students in specific learning circumstances (for example, those who were overage for grade or who had an Individualized Education Program [IEP]) appear to have been affected by the treatment more negatively. But this finding is not consistent across outcomes and grade levels, and it applies only to students in these circumstances who scored near the cut point on their fall screening test.

How to Interpret the Impact Findings and How This Study Differs from Prior Literature

The study uses a Regression Discontinuity (RD) design for its impact estimation. While this design demonstrates a causal relationship between assignment to receive intervention services and reading test outcomes in the impact sample, it also requires caution when interpreting the impact findings. In particular, the RD design estimates the impact of assignment to intervention by comparing outcomes of students just above or just below the cut point. Findings based on this design, therefore, cannot be generalized to all students receiving intervention services.

This is different from a randomized controlled trial (RCT), whereby similar eligible students are randomly assigned either to receive interventions or not to receive them. As a result, this design provides estimates of the average effect of intervention for students who would be added or dropped by marginally changing the eligibility criterion. In this sense, these results are relevant for decisions about expanding or reducing the scope of intervention but not, necessarily, for decisions about offering or not offering intervention. It would be misleading to conclude from these findings that providing increasing intensity of services to the students most at risk (for example, students whose screening test scores are far below the cut point) is inappropriate or ineffective.

In addition, this study is unique in the sense that it examines the RtI system as it operated in multiple states in a large sample of experienced schools that had implemented RtI on their own, without monitoring or support from researchers. This is different from most existing efficacy studies, in which the scale of the treatment is small (usually samples consist of fewer than 100 students and only a handful of schools) and the design and implementation of the RtI interventions are closely controlled by the researchers.

In order to understand the primary impact findings, the study explores the relationship between the impact estimates and school characteristics and RtI practices related to assignment to intervention. The key factors listed in Box ES.1 do not consistently explain the pattern of findings across grades. Unexplored but plausible factors that may be related to negative impacts of assignment to intervention on some Grade 1 students include (1) false or incorrect identification of students for intervention, (2) mismatch between reading intervention and the instructional needs of students near the cut point, and (3) poor alignment between reading intervention and core reading instruction.

This page has been left blank for double-sided copying.