



Comparing Tier 1 reading instruction with Tier 3 or special education intervention through an observational snapshot of school-implemented response to intervention across Grades 1–5

Stephanie Al Otaiba¹ · Jennifer Stewart³ · Wilhelmina van Dijk² · Carlin Conner³ · Dayna Russell Freudenthal^{1,4} · Brenna Rivas¹ · Paul Yovanoff¹ · Jill Allor¹

Accepted: 12 March 2024
© The Author(s) 2024

Abstract

There is limited research about Tier 3 interventions provided during typical school Response to Intervention (RTI) implementation. As part of a larger RTI exploration study designed to focus on students with the most intensive reading needs, our goal was to contrast their Tier 1 core reading instruction with their Tier 3 intervention. Schools identified participating students as receiving Tier 3 or special education for reading. We aimed to provide a snapshot of differences in Tier 1 and intensive interventions delivered to 264 students from 32 elementary schools. Our research team used the Instructional Content Emphasis-Revised (ICE-R; Edmond & Briggs, 2003) to observe differences in group size and the amounts and types of curricular content (e.g., code vs. meaning-focused) students received during both tiers. We explored whether these patterns were consistent across Grades 1–5 and if they differed in relation to students' disability (reading disability vs. other disability vs. no disability). We also examined whether the Tier 3 observation data was consistent with administrators' reports about RTI implementation within their school. Across the grades, we found significantly more small-group than whole-class instruction during Tier 3 than in Tier 1. We observed significantly lower proportions of code-focused instruction than meaning-focused instruction, particularly during Tier 1. Generally, code-focused instruction decreased across the grades in both tiers. Although we found a trend suggesting students with reading disabilities may have received higher proportions of code-focused instruction than other students, this was not significant after multiple comparisons. We discuss some similarities and differences between our observations and administrator data. We discuss implications, limitations, and directions for future research.

Extended author information available on the last page of the article

Keywords Response to intervention · Reading instruction · Reading intervention · Observation

Understanding how schools can accelerate students' reading performance is an important issue for educational research. Far too many students in the US have not learned to read proficiently; data suggest that in 2022, only about 33% of fourth graders could read proficiently, which represented a decline from both 2017 and 2019, when 37% and 35% of students could do so (National Assessment of Educational Progress; NAEP, 2017, 2019, 2022). Furthermore, NAEP data indicate that among students with disabilities, only about a third of fourth graders can read at even a basic level (NAEP, 2022). To prevent reading failure and provide additional intervention to accelerate the reading performance of students who are struggling to read (with or without an identified disability), US schools in all 50 states implement Response to Intervention (RTI) or Multi-tiered Systems of Support (MTSS) approaches for identifying students early and providing additional intervention (e.g., Berkeley et al., 2020; Vaughn & Fuchs, 2003).

An Institute of Education Sciences practice guide for RTI implementation for reading summarized the evidence base and specified key recommendations to support implementation (Gersten et al., 2009). These recommendations are central to the current study because they provide a framework for designing our study and interpreting the data we collected about RTI implementation, including observations and structured interviews with administrators. First, schools should screen all students' reading performance at least three times a year (beginning, middle, and end of the academic year). Second, school systems should provide adequate instructional time for systematic and explicit differentiated Tier 1 core instruction based on students' current reading levels. Third, schools should provide intensive, systematic Tier 2 standardized interventions to students who screen below benchmarks for grade level performance. Fourth, for students who receive intervention, progress monitoring should be implemented at least monthly to guide data-based decision-making. Finally, to accelerate reading growth for students with the most intensive needs, schools should provide daily intensive Tier 3 intervention that may include a standardized intervention program or be individualized based on student needs, referred to as *data-based individualization*.

Despite these recommendations, there is marked variability in how RTI and MTSS are implemented across states and even among districts. It is worth noting that intensive interventions within RTI and MTSS are intended to accelerate student learning and not to replace special education (National Center on Response to Intervention, 2010). This is important because researchers have cautioned about the potential blurring of boundaries between intensive services and special education within variable RTI implementation (e.g., Fuchs et al., 2010). Furthermore, many states passed dyslexia-specific legislation, particularly over the past decade, and such legislation includes many RTI components recommended by the IES practice guide (Gersten et al., 2009). For example, legislation may require schools to screen students for risk and provide reading instruction and intervention that is evidence-based or at least evidence-aligned (e.g., National Center on Improving Literacy, 2023; Petscher

et al., 2020; Youman & Mather, 2018). In some states, students with dyslexia may also receive support through a 504 plan, provided under Sect. 504 of the Rehabilitation Act (1973). Because of the variability in RTI and MTSS interpretation and implementation, there is a need for researchers to focus broadly on students with intensive reading needs, regardless of whether they have a disability label, and to include students who may receive intensive tiered services, dyslexia services, or special education.

Fortunately, there is a robust research base to inform effective literacy instruction and intervention for elementary students, including those with reading difficulties and disabilities (e.g., Castles et al., 2018; National Reading Panel, 2000). For example, explicit and systematic one-on-one or small group interventions help students increase their reading skills, especially in the early elementary grades (Al Otaiba et al., 2022). The effects of this instruction are more pronounced for foundational code-focused skills (i.e., phonemic awareness, phonics, word reading, spelling, and fluency) than for meaning-focused skills (vocabulary, comprehension, and written expression) Al Otaiba et al., 2022). The effects are strongest in early elementary grades. In the upper elementary grades, it is important to provide intervention for longer periods and to choose interventions addressing multiple components instead of just one skill or interventions with standardized protocols to help students optimize their growth (Donegan & Wanzek, 2021).

In contrast to this converging research on effective reading instruction and intervention, much less is known about the interplay and alignment between instruction at the different tiers of an RTI/MTSS system. Several large-scale studies focused on examining student outcomes but only focused on Tier 1 and Tier 2 (e.g., Burns et al., 2020; Coyne et al., 2018; Fien et al., 2021; for a review, see Russell Freudenthal et al., 2023). Across these studies, positive effects were only found when researchers supported the RTI implementation in schools. In more real-world conditions, without researcher supports, implementation does not always adhere to best and evidence-based practices (Balu et al., 2015; Ruffini et al., 2016), resulting in less growth for struggling readers (Fien et al., 2021). Notably absent from research is information about RTI implementation systems that include Tier 3, particularly regarding the alignment of Tier 1 core instruction and intensive interventions. Partly, this may be due to the variability with which researchers, policymakers, and schools define risk or response to intervention and partly to the limited research studies that have been conducted in Tier 3 settings (Al Otaiba et al., 2022).

Observation studies also inform our understanding of reading instruction and intervention within RTI implementation. A recent comprehensive review of studies published over the past two decades ($n=18$) synthesized findings of observational studies of literacy instruction and intervention within general and special education elementary classrooms (Stewart et al., 2024). The most common low inference observational tool was the Instructional Content Emphasis (ICE-R; Edmond & Briggs, 2003), which was used in four of the eighteen studies (Ciullo et al., 2019; Connor et al., 2009a, b; Swanson et al., 2012; Swanson & Vaughn, 2010). This tool provides a quantitative description of both grouping arrangements and the specific areas of literacy instruction focused on during the observation—two elements most relevant to the present study, which also used this tool. Notably, across these four studies,

students received relatively higher proportions of meaning-focused instruction than code-focused instruction. In addition, the lower dosages of code-focused instruction were generally less than research would recommend for accelerating growth in word reading skills for students with or at risk for learning disabilities.

Across the studies, only one study used the ICE-R to observe Tier 1 reading instruction in general education settings for Grades 1–3 (i.e., Connor et al. 2009a, b). Observations occurred in 30 classrooms once during half of the 90-minute language arts block, and findings revealed that meaning-focused instruction was more frequent than code-focused instruction. In contrast, the other three studies used the ICE-R to observe in resource rooms (Ciullo et al., 2019; Swanson & Vaughn, 2010) or RTI intensive intervention settings (Swanson et al., 2012). They did not observe Tier 1 settings but reported a similar pattern to Connor et al. (2009a, b) of more time spent on meaning-focused instruction than code-focused instruction (with 62%, 53%, and 66% of observed time reported as meaning-focused, respectively). Both resource room studies had an average of only four students, and only the earlier study (Swanson & Vaughn, 2010) described grouping arrangements; not surprisingly, whole group (45.8%) instruction was the most common grouping type. Less common were individualized (27.3%), independent (19.8%), small group or paired (4.6%).

Taken together, findings from these four studies using the ICE-R inform future research and underscore the need to contrast and compare Tier 1 to intensive interventions (in terms of grouping and curricular content) provided within school-implemented RTI to students identified as needing intensive interventions. Across the four studies, researchers observed relatively few classrooms (ranging between 10 and 30), suggesting a wider variety of classrooms is warranted to increase the generalizability of findings and to replicate Connor et al. (2009a, b) finding about the decline in code-focused instruction and increase in meaning-focused instruction across the grades.

Current study purpose within context of larger exploration study

The current study also used the ICE-R during observations as part of a larger exploration study about RTI implementation in 65 public elementary schools in 12 states. The research team recruited schools across three years (2016–2019) by contacting districts and schools proximal to their university settings or within their research partnership networks. Prior to participation, the research team contacted principals to confirm that the school was implementing RTI for reading and that an administrator (principal, assistant principal, or RTI leader) would participate in a structured interview about RTI implementation. After receiving university and district or school human subject approval, a subset of 32 of the total 65 administrators elected for their schools to participate in this current observation study. Because of the variability in RTI and MTSS interpretation and implementation and our focus on students with intensive reading needs, we asked schools to identify one to three target students for us to observe who received Tier 3 or special education services for reading. We extended prior research by providing an observational *snapshot* contrasting the core reading instruction with intensive reading interventions schools provided to students who qualified for Tier 3 or special education as part of schools' typical RTI

implementation. In other words, schools identified students who received intensive tiered services, dyslexia services, or special education services for reading. Our first research question asked, “What differences were observed between Tier 1 and Tier 3 groupings, and to what extent did groupings vary across Grades 1–5?” We extended prior research by observing whether Tier 3 was provided in smaller groupings than in Tier 1. We hypothesized that Tier 3 would involve relatively smaller groups of students to allow teachers to individualize based on student needs. Our second question asked, “To what extent did dosage and type of instruction (code-focused vs. meaning-focused) differ between core reading instruction and Tier 3 intervention?” We described the overall amount of code and meaning instruction. Given that the length or number of minutes we observed reading varied, we examined whether any differences in the *proportion* of content students received were consistent across elementary grades (Grades 1–5) and whether they differed relative to students’ disability (reading disability vs. other disability vs. no disability). We observed whether students eligible for Tier 3 or special education received relatively more code- or meaning-focused curricular content during Tier 3 or Tier 1. We hypothesized that students would receive relatively more code-focused instruction in the earlier grades while they were learning how to read (particularly in Tier 3). By contrast, we hypothesized that students in the upper grades might receive multi-component interventions focused relatively more on comprehension. We examined whether the type of students’ disability reported by schools (i.e., reading, other special education, or no disability) was associated with differential curricular content (indicating differentiation, potentially individualized for students’ needs). We hypothesized that students with reading difficulties or disabilities might receive relatively more code-focused instruction.

Our third question asked, “To what extent did the Tier 3 observation data collected align with school administrators’ reports about Tier 3 implementation in their schools?” We explored whether data from the observations of Tier 3 were consistent with administrators’ reports about how Tier 3 was implemented within RTI in their schools. Our intent was to contribute *uniquely* to the RTI research base by contextualizing our observation findings with administrators’ interviews about RTI implementation, particularly for students who received Tier 3. To date, there is a paucity of research exploring the alignment between administrator reports and observed implementation in classroom settings, so we had no a priori hypothesis about this question.

Method

Schools and participants

As mentioned, principals confirmed that their schools implemented RTI for reading. The sample included 32 schools in 9 states, but a majority ($n=18$) were in Texas. Similar proportions of schools were categorized as rural ($n=8$), urban ($n=11$), and suburban ($n=13$). The proportion of students in these schools who received free and reduced-priced lunch (FRPL) ranged from 22.9 to 97.8%; in a majority of schools, more than 50% of students received FRPL. Across the schools in the sample, rela-

tively fewer students were identified racially as Native American, Pacific Islander, or Asian than as Black, Hispanic, or White. Schools did not provide individual demographic information beyond grade level, tiered instruction, and special education status. Supplemental Table 1 details the demographics of the 32 participating schools and specifies the number of students observed within each school.

The nomination criteria we provided to administrators were intentionally broad; we asked them to identify 5–6 target students in grades 1–5 who were receiving intensive interventions (Tier 3, dyslexia services, or special education) in reading or literacy in addition to Tier 1 core instruction. We did not have permission to examine student intervention plans to review individual student literacy goals or to link any reading data collected by schools with our observations. We observed 264 students in Grades 1 through 5; due to scheduling conflicts, we were not able to observe Tier 1 for three students and Tier 3 for one student. Of these 264 students, we observed roughly the same proportion of students at each grade; in other words, the number of students we observed per grade level was not statistically significant from each other ($X^2_{[4]}=2.06, p=.725$). For the purposes of our study, we categorized disability status as a *reading disability* (i.e., dyslexia or specific learning disability), *other disability* (i.e., Visual Impairment, Autism, Cognitive Impairment, Emotional Disorders, Intellectual Disability, Speech or Language Impairment, or Multiple or Unspecified Special Education), or *no disability*.

Data sources and procedures

Prior to observation, participating schools assisted us in identifying students with intensive reading needs, recruiting these students, and contacting their families to secure consent. Schools nominated several target students in grades 1–5 who were receiving intensive reading interventions, either through special education or in Tier 3. Because our goal was to contrast intensive interventions with the core instruction in schools, the research team scheduled observations with each teacher to observe target students during their core classroom reading instruction (Tier 1) and their intensive intervention reading instruction (Tier 3 or special education). Observations occurred at one time point during either winter or spring to obtain a snapshot of the type and amount of reading instruction students received in their classroom and in intervention.

Observation tool and coding components

For our observation tool, our research team used the *Instructional Content Emphasis Instrument-Revised (ICE-R; Edmonds & Briggs, 2003)* to record what reading content was taught, for how many minutes, and for what size instructional group. Consistent with the guidelines of the ICE-R, specific instructional activities were coded only if they lasted for at least 1 min. ICE-R distinguishes between several literacy curricular content categories, including phonological awareness, phonics/word recognition, fluency, vocabulary/oral language development, grammar, comprehension, spelling, writing, and independent text reading (separate from teacher-led instruction). The tool also includes non-literacy curricular content categories: other academic instruc-

tion (e.g., art or specials), non-instructional time (e.g., transitions, restroom breaks), and behavior management (e.g., teacher setting expectations or providing redirections). Observers also coded instructional groupings as whole class, small-group, pairs, independent activity/assignment, individualized instruction, or tutoring. Whole class denoted an entire class working on the same assignment. This scenario could have occurred in Tier 1 or Tier 3 if the class size in intensive intervention included seven or more students. Small groups could occur when the class was divided into two or more groups or if the class size was less than seven. Independent instruction occurred when all students were working by themselves on the same assignment; individualized or tutoring occurred when the assignments were differentiated for each student or when the teacher was providing one-to-one instruction.

For the purposes of this study, we followed the convention of the *Simple View of Reading* (Gough & Tunmer, 1986) and separated the literacy content into code-focused (phonological awareness, phonics/word recognition, fluency, and spelling) and meaning-focused (vocabulary, language, text reading, grammar, comprehension, and writing) curricular categories. To clarify, fluency was coded when students were reading for accuracy and prosody, whereas text reading was coded when teachers or students were reading aloud within connected text with an emphasis on reading for understanding.

Observer training and reliability

We recruited 24 observers and trained them to use the ICE-R (Edmonds & Briggs, 2003). These observers were recruited from ten universities located near elementary schools participating in the larger study. Most observers were doctoral students, some were faculty, and some were research staff with advanced education degrees. Yearly, before we conducted observations, we established inter-rater reliability to a gold standard with all observers. This reliability training included in-person and virtual sessions that (a) provided details about the observation form and coding manual and (b) modeled and shared an example of the gold-standard coding sheet that was paired with a reading instructional video. Upon completion of the training session, observers coded another reliability video; reliability was met when the observer obtained 90% agreement with the gold-standard coder. We required continuing observers to reestablish reliability before the start of each observation year. Across the years, reliability ranged from an average of 91.4–92.9%.

Data analysis plans

In the first part of the data analysis, we focused on providing descriptive statistics for the data. First, we examined the raw minutes of instruction observed. Due to the high variability in raw observation minutes, we converted raw minutes to a proportion of time for both instructional grouping and curricular content (see Figure 1). We then used these proportions as the basis for the remaining parts of the quantitative analysis. We described the overall percentage of target students who received any code- or meaning-focused instruction to provide context for the observations.

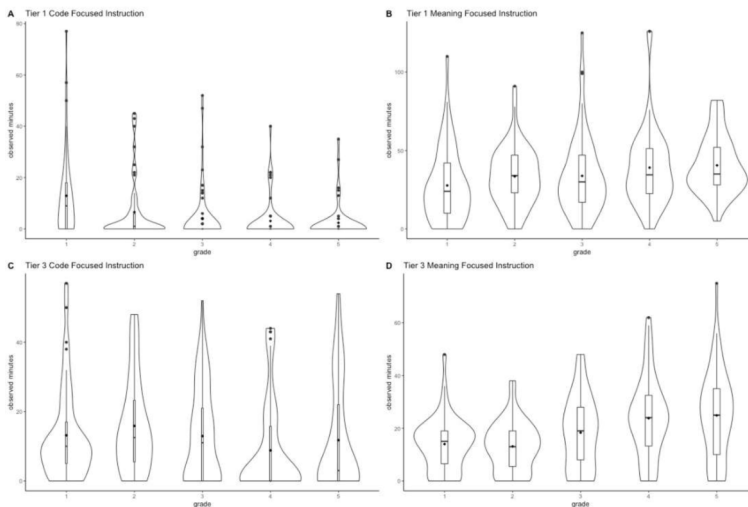


Fig. 1 Violin Plots of Observe Minutes Across Tiers and Grades. *Note* Violin plots with boxplots for observed minutes. For the violin plots, a wider waist indicates more observations. Panel A shows the violin plots for Tier 1 code-focused instruction. Panel B shows the violin plots for Tier 1 meaning-focused instruction. Panel C shows the violin plots for Tier 3 code-focused instruction. Panel D shows the violin plots for Tier 3 meaning-focused instruction. Code-focused instruction included phonological awareness, phonics/word recognition, spelling, and fluency. Meaning-focused instruction included vocabulary and oral language development, comprehension, text reading, grammar, and writing. Other instruction included other academic instruction (i.e., instruction not included in code- or meaning-focused instruction), non-instructional time, and non-instructional time that focused on behavior management

To examine differences in the proportions of time students spent in a type of instruction or grouping, depending on the instructional tier, grade level, and disability status, we conducted three separate split-plot ANOVAs. The two tiers represented the within-subjects factor in each ANOVA. We added the type of instruction, grade level, instructional grouping, and disability status as between-subjects factors. For this study, we categorized disability status as (1) reading disability (i.e., dyslexia or specific learning disability), (2) other disability, or (3) no disability. All data analyses were conducted in the *R* statistical environment using the *lme4* (Bates et al., 2015) and *emmeans* (Lenth, 2022) packages.

To explore consistency in what we observed with school administrators' descriptions of Tier 3 interventions, we examined data from structured interviews conducted in the larger study. Trained research staff interviewed administrative staff (including principals, assistant principals, and RTI/MTSS coordinators) at all participating schools using a standard interview protocol, the *RTI Essential Components Worksheet* (Center on Response to Intervention; AIR, 2014). We used this protocol to elicit administrators' descriptions of their campus-wide implementation of essential RTI/MTSS practices. While this interview protocol included multiple components, for the present study, we focused on the component *Tier 3 interventions*.

Results

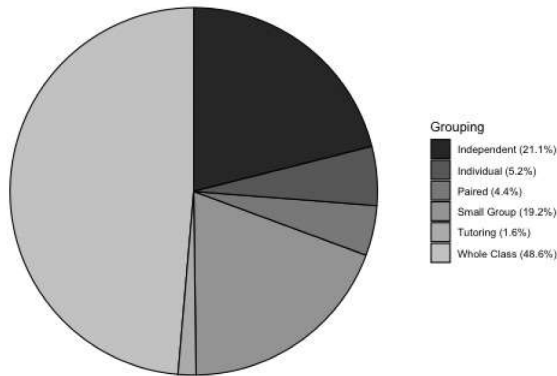
Observed differences in grouping between Tier 1 and Tier 3 across the grades

First, we examined differences in the proportions of time students spent in different instructional groupings in both tiers across grades. We conducted a split-plot ANOVA with tier as the within-subject factor and grade and instructional grouping as between-subjects factors. The results from this ANOVA suggest that the proportion of time spent in a particular instructional grouping depended only on tier and not on grade. The pie charts in Fig. 2 show the proportions of time students spent in each instructional grouping per tier. The three-way interaction between tier, grade, and grouping was not statistically significant, nor were the interactions between tier and grade and grouping and grade. The model did yield a statistically significant interaction between tier and grouping ($F_{[5,2712]}=150.18, p<.001$). Figure 3a depicts this interaction in a marginal means plot. The differences between small group and whole class instruction were both statistically significant with large effect sizes. In Tier 3, students spent most of their time in small group instruction (roughly 62% of the observed time), compared to time spent in small group within Tier 1 (roughly 20% of the observed time) ($\Delta M=0.42, t_{[2498]}=18.30, p_{holm}<.001, d=1.72$). While participating in Tier 1, students spent most of their time in whole class instruction (roughly 50% of the time); conversely, during Tier 3, they spent very little time in this grouping (roughly 5%) ($\Delta M=-0.45, t_{[2498]}=-19.66, p_{holm}<.001, d=-1.84$). The difference for independent work was also statistically significant with a medium effect size. Students spent proportionally less time in independent work in Tier 3 compared to Tier 1 ($\Delta M=-0.08, t_{[2498]}=-3.71, p_{holm}=.006, d=-0.35$). There were no statistically significant differences between the tiers in the proportion of time spent in individual instruction, pair work, or tutoring.

Differences in dosage and types of code- and meaning-focused instruction between core reading instruction and Tier 3 intensive intervention

Second, we examined differences in the dosage and types of code- and meaning-focused instruction and intervention students received during their core reading instruction compared to their Tier 3 intervention. Our initial descriptive analyses of observations, as shown in Table 1, reveal that relatively few students received any code-focused content during their Tier 1 instruction (i.e., only 35% of the 264 target students). By contrast, nearly all students (95%) received meaning-focused instruction. During their Tier 3, more than half (65%) of the target students were observed receiving code-focused intervention, and more (85%) received meaning-focused instruction. However, Table 1 reveals slightly different observed trends by grade level, with relatively less code-focused instruction and intervention received in the upper grades. For example, in Grade 1, during Tier 1 observations, 65% and 86% of target students received code- and meaning-focused instruction, respectively; whereas, during Tier 3 observations, 83% of target students received code and meaning-focused interventions. However, in Grade 4, during Tier 1 observations, only 19% of target students received code-focused instruction, while 98% received mean-

A Tier 1



B Tier 3

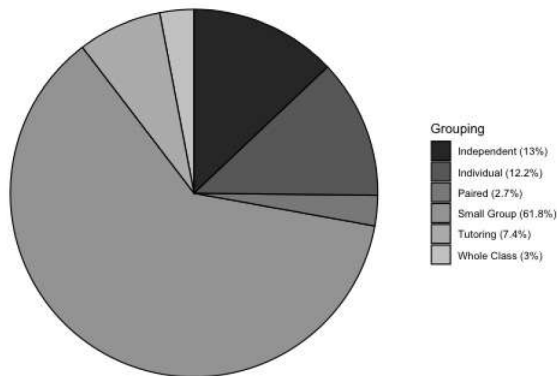


Fig. 2 Pie Charts of Proportions of Time Students Spent in Each Instructional Grouping. *Note.* Whole class instruction occurred when the entire class was involved in the same activity or assignment. Small group instruction occurred with at least 2 students and a teacher or at least three students without a teacher. Paired instruction occurred with 2 students where one student acted as a tutor. Independent instruction occurred when students were working individually. Individualized instruction occurred when students were working individually but on differentiated assignments. Tutoring occurred when a teacher worked with only one student for the entire instructional period. Panel A shows the proportions of time for Tier 1. Panel B shows the proportions of time for Tier 3

ing-focused instruction. During Tier 3 observations, 48% and 90% of target students received code- and meaning-focused interventions.

Generally, the length of the observations for Tier 1 and Tier 3 varied, even though the overall means were relatively consistent, as is shown in Table 2. On average, across the grades, we observed students in Tier 1 for 45.87 (range 9–195) minutes and then observed the same students during Tier 3 for an average of 33.42 min (range 6–96). The violin plots in Fig. 1a-d represent this variability in raw observed minutes. Each violin plot includes a boxplot with the median and quartiles along with outliers; the shape of the plot represents the distribution of the observed minutes (i.e., a tradi-

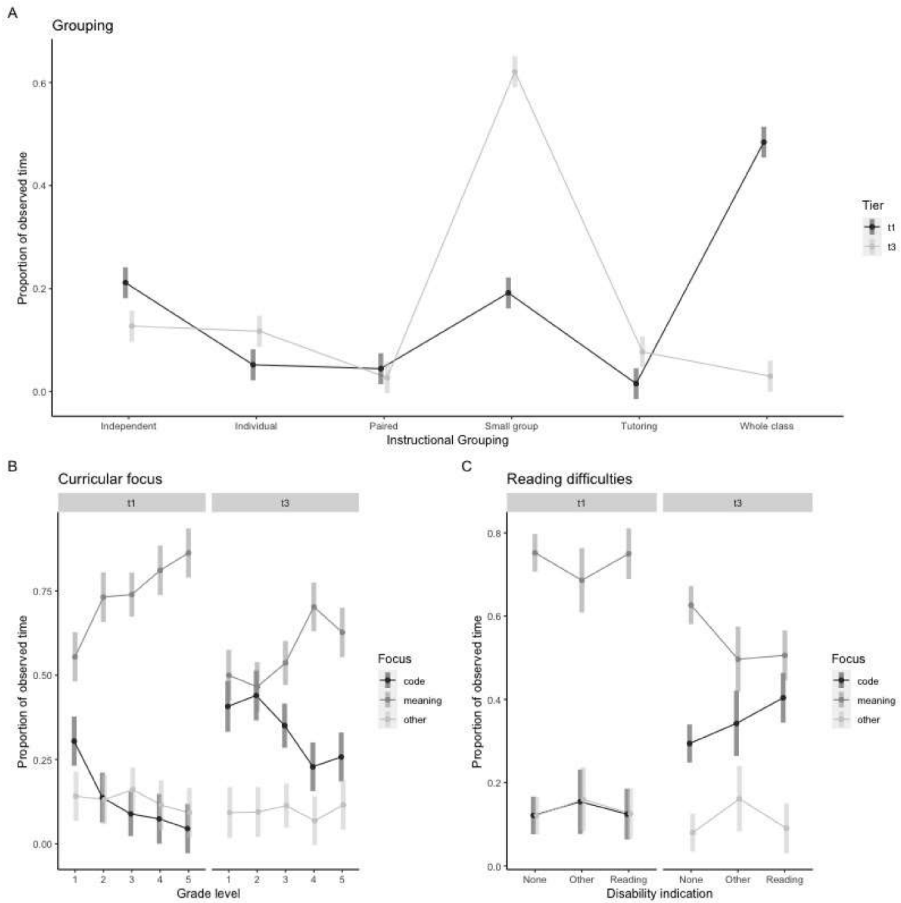


Fig. 3 Marginal Means Plots. *Note.* Panel A shows the marginal means for the interaction between instructional grouping and Tier. Panel B shows the marginal means for the three-way interaction between grade level, curricular focus (i.e., code-focused, meaning-focused, and other instruction), and Tier. Panel C shows the marginal means for the three-way interaction between disability status (i.e., no disability, a different disability, or a reading disability), curricular focus, and Tier

tional distribution plot rotated 90 degrees). Code-focused instruction in Tier 1 had the least variability, with plots heavily concentrated around zero observed minutes, with several high outliers. Meaning-focused instruction in Tier 1 showed heavier *waists* in the violins, indicating most students received this instruction around the median number of minutes. In Tier 3, the violin plots for meaning-focused instruction show increased variability across grades. While Grade 1 has a smaller range with a normal distribution, Grade 5 shows a platykurtic shape (wider and flatter).

Given this variability within and between tiers and grades, our analyses focused on the proportions of time rather than the raw number of minutes. Tables 3 and 4 provide the proportion of observed minutes for instructional grouping and curricular content, respectively, per tier and grade. We describe the specific proportions of time for each subtype of code- and meaning-focused curricular content by grade in Supplemen-

Fig. 4 Pie Charts Representing the Change in Proportions of Curricular Content Across Grades and Tiers. *Note* Columns represent the two Tiers (1= Tier 1; 3= Tier 3). The rows represent the different grades (1=Grade 1; 2=Grade 2; 3=Grade 4; 4=Grade 4; 5=Grade 5). Code-focused instruction included phonological awareness, phonics/word recognition, fluency, and spelling. Meaning-focused instruction included vocabulary and oral language development, comprehension, text reading, grammar, and writing. Other instruction included other academic instruction (i.e., instruction not included in code- or meaning-focused instruction), non-instructional time, and non-instructional time that focused on behavior management

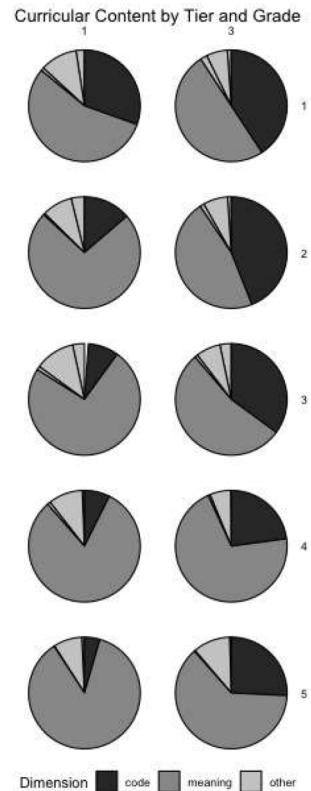


Table 1 Percentage of students observed receiving code vs. Meaning-focused instruction by Tier

Grade	Tier 1		Tier 3	
	Code	Meaning	Code	Meaning
Grade 1	65	86	83	83
Grade 2	51	96	85	85
Grade 3	25	93	61	82
Grade 4	19	98	48	90
Grade 5	16	100	53	88
Total	35	95	65	85

Note Code-focused instruction included phonological awareness, phonics/word recognition, spelling, and fluency. Meaning-focused instruction included vocabulary and oral language development, comprehension, text reading, grammar, and writing

tal Table 2. To examine whether these differences in content were consistent across grades, we conducted a split-plot ANOVA with the two tiers representing the within-subjects factor and the three types of instruction and grade representing the between-subjects factor. There was a significant three-way interaction between tier, grade, and instructional focus ($F_{[8,1355]}=2.60, p=.008$). See Fig. 3b for the marginal means. In general, the proportion of code-focused instruction decreased across the grades in both tiers, whereas the proportion of meaning-focused instruction increased.

Table 2 Length of observation in minutes by Tier

Grade	Tier 1				Tier 3			
	<i>M</i>	<i>SD</i>	<i>Range</i>	Total	<i>M</i>	<i>SD</i>	<i>Range</i>	Total
1	47.38	28.00	9-150	2,369	28.04	15.34	6-72	1,402
2	45.66	22.56	20-148	2,283	30.60	13.34	10-61	1,530
3	44.98	29.93	11-195	2,789	34.50	10.99	13-60	2,139
4	46.00	28.24	10-141	2,300	34.74	13.11	10-74	1,737
5	45.56	24.98	19-98	2,369	38.73	18.56	10-96	2,014
Total	45.87	26.81	9-195	12,110	33.42	14.73	6-96	8,822

This pattern is represented by a significant two-way interaction between grade and instructional focus ($F_{[8,1355]}=10.299, p<.001$). Additionally, the interaction between tier and instructional focus was also significant ($F_{[2,1355]}=66.36, p<.001$). Across the five grades, students received proportionally more code-focused instruction in Tier 3 compared to Tier 1 ($\Delta M=0.22, t_{[1141]}=8.76, p_{holm}<.001, d=0.88$) and less meaning-focused instruction ($\Delta M=-0.18, t_{[1140]}=-7.63, p_{holm}<.001, d=-0.74$). Despite the overall trend of increase in meaning-focused instruction across grades and tiers, as Figs. 3b and 4 show, we found different patterns in Grades 1 and 2. In Grade 1, the amounts of code- and meaning-focused instruction were similar across Tier 1 and Tier 3. Uniquely, in Grade 2, the proportion of code-focused instruction increased in Tier 3, and the proportion of meaning-focused instruction decreased relative to Tier 1. In Tier 3, the amount of time students received code- and meaning-focused instruction in Grades 1 and 2 was about the same. There were no differences across grades or tiers in the proportion of instructional time spent on other instruction.

Then, we examined whether instruction varied across the tiers relative to special education status. We conducted a split-plot ANOVA, with tier as the within-subjects factor and both special education status (i.e., reading disabilities, other disabilities, and no disabilities) and instructional focus as the between-subjects factors. The three-way interaction was not statistically significant. The interaction between instructional focus and special education status was significant ($F_{[4, 1508]}=4.83, p<.001$), as well as the interaction between instructional focus and tier ($F_{[2,1508]}=59.47, p<.001$). Figure 3c shows the marginal means for the split-plot ANOVA. As the figure indicates, instruction in Tier 1 was generally similar for students regardless of their disability status. In Tier 3, the proportion of code-focused instruction increased for students with disabilities, but none of the pairwise comparisons were statistically significant. Students without disabilities spent proportionally more time on meaning-focused instruction than students with other disabilities ($\Delta M=0.13, t_{[1508]}=2.81, p_{holm}=.23, d=0.48$) and students with reading disabilities ($\Delta M=0.12, t_{[1508]}=3.14, p_{holm}=.09, d=0.45$), but these differences were not statistically significant after correction for multiple comparisons.

Exploring consistency and alignment between observations of Tier 3 and school administrators' reports about Tier 3 implementation

As we described earlier in our Method section, we conducted structured interviews with administrators in the 32 schools, asking them about how their campuses imple-

Table 3 Proportion of observed time for instructional grouping per Tier and Grade

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5		Total		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	
Whole class	T1	45.78	32.56	44.34	33.66	52.66	38.54	51.55	36.38	47.79	36.62	48.63	35.64
	T3	0.71	3.75	1.92	8.58	2.15	12.94	5.35	18.66	4.89	18.00	3.01	13.71
Small Group	T1	23.95	32.39	14.13	24	18.54	27.24	18.92	28.68	20.32	29.22	19.16	28.34
	T3	64.37	39.33	59.65	40.88	55.6	39.39	59.07	43.07	71.77	35.48	61.76	39.80
Paired	T1	4.85	10.05	3.89	9.79	3.28	7.91	7.19	17.61	3.15	7.46	4.41	11.03
	T3	3.74	15.5	1.84	5.79	1.85	8.59	5.02	20.01	1.15	6.97	2.68	12.46
Independent	T1	17.38	20.56	25.8	25.43	19.59	26.97	19.73	27.29	23.24	27.72	21.06	25.75
	T3	14.3	23.3	10.51	19.15	17.73	26.45	11.81	21.2	9.29	18.23	12.95	22.15
Individual	T1	8.04	16.32	5.88	14.5	4.07	18.17	2.61	11.55	5.46	17.95	5.17	16.00
	T3	5.04	17.36	15.97	32.6	21.06	37.32	8.75	25.46	7.84	21.07	12.22	28.78
Tutoring	T1	0	0	5.95	23.34	1.86	12.23	0	0	0.03	0.24	1.57	11.88
	T3	11.83	30.67	10.11	30.03	1.61	12	10	30.3	5.07	20.95	7.39	25.41

Note Whole class instruction occurred when the entire class was involved in the same activity or assignment. Small group instruction occurred with at least 2 students and a teacher or at least three students without a teacher. Paired instruction occurred with 2 students where one student acted as a tutor. Independent instruction occurred when students were working individually. Individualized instruction occurred when students were working individually but on differentiated assignments. Tutoring occurred when a teacher worked with only one student for the entire instructional period

Table 4 Proportion of observed time for curricular content per Tier and Grade

Curricular Content	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5		Total		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	
Code focused	T1	30.44	31.74	13.70	22.78	8.85	21.78	7.39	20.05	4.48	12.35	12.82	24.18
	T3	40.76	31.38	43.99	31.25	35.06	34.58	22.87	32.83	25.75	34.14	33.59	33.69
Meaning focused	T1	55.45	32.69	73.18	26.51	73.93	30.18	81.13	23.71	86.31	18.18	773.97	28.62
	T3	50.00	32.06	46.58	31.58	53.65	38.02	70.32	35.75	62.72	36.64	56.68	35.87
Other	T1	14.10	13.95	13.12	14.97	16.07	21.36	11.48	15.31	9.21	10.35	12.96	15.98
	T3	9.24	11.15	9.43	12.25	11.29	19.23	6.81	13.70	11.52	22.13	9.73	16.43

Note Code-focused instruction included phonological awareness, phonics/word recognition, spelling, and fluency. Meaning-focused instruction included vocabulary and oral language development, comprehension, text reading, grammar, and writing. Other instruction included other academic instruction (i.e., instruction not included in code- or meaning-focused instruction), non-instructional time, and non-instructional time that focused on behavior management

mented RTI. Overall, administrator reports of Tier 3, or intensive intervention, were consistent with our observations. When we asked them about a typical student receiving Tier 3 services, administrators reported dosages of intensive intervention ranging from 15 to 60 min per session. Most commonly, principals reported that 30-minute intervention sessions were delivered a minimum of four times per week. Thus, the administrator reports of dosage align with our mean observed time of 33.42 min for Tier 3 instruction.

Most administrators reported that students received intensive interventions mostly in a one-on-one setting; although some indicated the interventions could be delivered in a small group setting (e.g., up to three students). However, we found discrepancies between administrator interviews in terms of Tier 3 grouping and instructional content. In contrast to administrator reports, our observations of Tier 3 indicated that the majority (61.8%) of time was within a small group setting. One-to-one instruction (i.e., tutoring) was observed for 7.4% of the time, and individualized instruction (i.e., students working alone on an individualized assignment) was observed for 12.2% of the time. Initially, administrators gave minimal descriptions of data-based decision-making to develop individualized interventions. However, when prompted to give more detail, principals typically described those intensive interventions following a “case-by-case” approach. Regarding a discrepancy in the observed instructional content, many administrators reported that interventions targeted fluency specifically. However, while generally fluency was included in our combined analysis of code-focused instruction, individually it only accounted for a small proportion (5.03%) of observed Tier 3 intervention.

Discussion

In this observational study, we provided a snapshot to contrast Tier 1 and Tier 3, or intensive reading, interventions within RTI implementation delivered to targeted students identified by schools as receiving Tier 3 or special education. As we hypothesized, we observed differences in grouping across the two tiers, with significantly more small group instruction during Tier 3 and mostly whole group instruction during Tier 1. We found large effect sizes corresponding to these differences between Tier 1 and Tier 3 for small and whole group ($d=1.70$ and $d=1.84$, respectively). These findings suggest it was feasible for teachers to provide small group instruction, which is an important dimension of individualized and intensive intervention (e.g., Al Otaiba et al., 2022; Gersten et al., 2009; Hall et al., 2022). Prior research (though provided in Tier 2 settings) has demonstrated that the effectiveness of interventions may be even stronger in one-to-one versus small group settings in the early grades, particularly in primary grades (e.g., Gersten et al., 2020; Vaughn et al., 2010). While we observed more individualized interventions during Tier 3 than Tier 1, one-to-one intervention occurred for only 7% of students who received Tier 3. High caseloads and other resource demands may limit teachers’ ability to provide more one-to-one intervention.

Our findings about Tier 1, or core instruction, differ considerably from the teacher reports of core instruction reported by Balu et al.’s (2015) evaluation of RTI imple-

mentation. In that study, schools were recruited that had implemented RTI for at least three years; for our study, we recruited schools that reported currently implementing RTI. Balu et al. found that teachers reported spending around 33% of the core reading block in whole class instruction and 25% in small groups. In contrast, we observed relatively more whole group instruction (about 47% of the time) and relatively less small group instruction (only 19% of the time). Additionally, there was a difference in peer activities; teachers in the Balu et al. study noted using this instructional mode around 17% of the time, while we observed this mode for only 4% of the time. Possible explanations for the lack of convergence across our study findings could relate to differences between the studies' measures and procedures (e.g., Balu et al. used teacher reports, whereas we observed actual instruction).

Our second aim was to observe whether students eligible for intensive interventions in Tier 3 or special education received relatively more code- or meaning-focused curricular content during Tier 1 or Tier 3. Almost all students received meaning-focused instruction in Tier 1 (95%) and Tier 3 (85%). Notably, few students (35%) received any code-focused instruction in Tier 1. In contrast, during Tier 3, about 65% of students were observed receiving code-focused instruction. The amount of code-focused instruction in both tiers appeared limited, with dosages less than recommended by the research base on effective instruction and intervention (e.g., Al Otaiba et al., 2022; Austin et al., 2017; Hall et al., 2022). We also observed relatively more code-focused intervention in Tier 3 than in Tier 1. Although we do not have specific recommended amounts or types of instruction or intervention based on individual students' needs, we noted that students in the earlier grades received relatively more code-focused instruction (consistent with a learning-to-read phase), with decreasing amounts across the upper grades. The pattern of decreasing code-focused instruction is consistent with Connor and colleagues' observation study of Tier 1 during Reading First (Connor et al., 2009a, b). An interesting deviation from this pattern occurred in Grade 2; as shown in Fig. 3b, we observed that students who received Tier 3 were exposed to more code-focused intervention than expected (based on the data). One possibility for this deviation could be that schools prioritized code-focused intervention because of the impending curriculum shifts to reading-to-learn (with a focus on comprehension) in the upper grades. For example, Connor et al. (2007) referred to students needing a second chance in second grade when they had not mastered word reading and foundational skills in first grade.

Our third aim was to observe whether the type of students' disability status (i.e., reading disability, other disabilities, or no disability) was associated with differential curricular content. We found that students without disabilities received a significantly greater proportion of meaning-focused instruction in Tier 3 than students with disabilities. We had hypothesized that students with reading difficulties would receive relatively more code-focused instruction during Tier 3; although we did observe this trend (see Fig. 3c), it was not statistically significant. Overall, only about a third of observed intervention time was code-focused during Tier 3. This proportion is within the range of code-focused observations in prior observation research conducted in resource rooms or RTI settings, ranging from a high of 47% of the observed time (Swanson & Vaughn, 2010) to a low of 28% (Ciullo et al., 2019). By contrast, slightly more than half of their observed Tier 3 time had a meaning-focused intervention

(57%), suggesting that students may have been receiving multi-component interventions. The proportion of meaning-focused instruction we observed was lower than in the Swanson et al. (2012) study (66%) or the Ciullo et al. study (62%) but slightly higher than in the Swanson and Vaughn (2010) study (53%). We observed that less than 10% of Tier 3 intervention time was non-instructional, similar to the amount reported by Ciullo et al. at 9%. Additional research that includes IEP or intervention goals is warranted; we do not intend to suggest that all students needed only more code-focused instruction, but rather that they may have needed differentiated amounts based on their own skill levels. In addition, a majority of students with poor comprehension skills benefit from word level instruction as well as comprehension supports (e.g., Donegan & Wanzek, 2021; Hall et al., 2022).

Our final aim was to contextualize our observation findings with administrators' interviews about RTI implementation, particularly for students who received Tier 3. We found some consistent alignment between what we observed and their reports about Tier 3 implementation. However, we also identified some inconsistencies; administrators perceived fluency as a more prominent component of Tier 3 interventions, whereas we observed that only 5% of intervention time focused on fluency specifically. The administrators also reported that Tier 3 was more individualized than we observed. For example, they described their schools as providing a more intensive, individualized, one-to-one instructional setting for students who received Tier 3, but the snapshots we observed revealed that small group settings were more common during Tier 3. It may be that the administrators were focused on systems support for RTI and may not have had a detailed view of what was currently being provided during Tier 3 implementation during the time of our observations.

Limitations and directions for future research

As with all research conducted in school settings, our research had several limitations. First, we had limited information about student demographics given the nature of our consent process with schools and districts, which precluded us from accessing student identification numbers that would have allowed us to identify their race, ethnicity, or gender or to link observations with their individual reading outcomes. Relatedly, we could not triangulate our observation data with either instructional plans or individualized education plans. Although we specifically relied on school administrators to nominate target students who had intensive reading needs, our research team did not have permission to evaluate each individual's disability or to access their specific IEP goals or their Tier 3 instructional plans. It is possible that some students we classified as having a reading disability due to their *specific learning disabilities* label had mathematics or writing as primary need areas, even though we specifically asked administrators to identify students receiving intensive reading interventions. Thus, a limitation, which is relevant to many large-scale studies involving school-provided data, is that researchers may not have all the details about individual student needs. Future research could examine the alignment between observed instruction and student responsiveness data.

Second, our specific focus was to contrast Tier 1 with Tier 3 for students with the most intensive reading needs; while we observed each target student during core

reading instruction (Tier 1) and intensive intervention (Tier 3 or special education), we observed only once. Future research is needed to conduct longitudinal observations that would be sensitive to changes across the year, include more detailed information about core reading and intervention programs, and document whether progress monitoring or data-based decision-making informed Tier 3 instruction. In addition, we did not observe any Tier 2 intervention. Given that research has shown the effect of tiered instruction across Tiers 1 and 2 (Fien et al., 2021), future longitudinal research should also observe Tier 2. This research might include a focus on alignment of instruction across tiers and a focus on students' movement across tiers (cf. Al Otaiba et al., 2014). For example, it would be pertinent to confirm whether students receiving the most intensive services for Tier 3 or special education also receive some Tier 2 programs across the school year. Third, we did not seek information about teacher or administrator training for RTI, generally, or for reading interventions specifically. Future research could explore the relations between training and observed instruction.

Implications for research and practice

Despite our limitations, our study adds some novel findings with important implications for research and practice. Given that we found that Tier 1 was mostly whole group and was rarely individualized or differentiated, administrators and teachers should carefully consider instructional arrangements to support differentiation of instruction. For example, systems should be in place that provide data to help teachers identify which students need to practice with which skills so brief small group instruction can take place. One example of a web-based system to guide differentiation of Tier 1 with a substantial evidence base is the Assessment to Inform Instruction (A2I) within the individualized student instruction (ISI) intervention (Connor et al., 2009b). Another evidence-based practice for differentiating Tier 1 is peer tutoring, which increases the intensity of instruction in several ways, including (a) additional opportunities to read aloud and receive immediate feedback from a peer, (b) text selected to meet the needs of pairs of students, and (c) opportunities for the teacher to provide brief corrective feedback during oral reading (Mathes & Fuchs, 1994).

We also found that it was feasible for schools to provide significantly more small group instruction during Tier 3 than during Tier 1. It was noteworthy that few students received any code-focused instruction, and when they did, it was for less time (in either tier) than would be recommended by the research base about effective instruction and intervention. Given our research design and focus, we did not have access to students' individual skills and goals in the code- or meaning-focused components; however, based on our findings and those of earlier observation studies (Ciullo et al., 2019; Swanson et al., 2012; Swanson & Vaughn, 2010), we urge administrators and practitioners to provide appropriate amounts of code- and meaning-focused instruction, particularly in the early years (Connor et al., 2009b). Teachers need to be aware that most students needing Tier 3 intervention, particularly those with word reading difficulties, typically need intensive instruction to develop skilled and automatic word reading skills. Even as these students advance through the elementary grades, they typically struggle to identify longer, polysyllabic words and to read with fluency

(Hudson et al., 2022; Kearns et al., 2022). Students with word reading difficulties also need support for comprehension. Strategies for developing these skills should be routine parts of both Tier 1 and Tier 3 instruction. Again, data systems similar to the system designed by Connor et al. (2009b) could support teachers in identifying appropriate amounts of code-focused versus meaning-focused instruction needed by specific students.

Further, target students received relatively more code- than meaning-focused content during Tier 3 than during their Tier 1 instruction; thus, there is a need to ensure alignment between the core and supplemental curricular content. Our observations also suggest that content was not significantly different in relation to school-reported students' disability status (reading disability vs. other disabilities vs. no disability), in contrast to our expectations; therefore, an important implication is for interventions and changes to instruction to be guided by an individual student's educational need.

Finally, we found that administrator reports about RTI implementation were consistent with our observations about differences in grouping across tiers, but they were not always aligned with the specific observed curricular content provided to students during Tier 3. Therefore, administrators need to monitor data carefully to ensure that students' code- and meaning-focused skills are adequately developing and instruction is aligned with these needs. In addition, school administrators could conduct walk-throughs, observing Tier 1 and Tier 3 within their schools. In conclusion, we caution that our study provided only one snapshot of instruction and intervention and that further research is needed to guide school practice about RTI implementation.

Funding Open access funding provided by SCEL, Statewide California Electronic Library Consortium. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A160132 to Southern Methodist University and Grant R324B180034 to the University of Virginia. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To wait in Tier 1 or intervene immediately: A randomized experiment examining first-grade response to intervention in reading. *Exceptional Children*, *81*(1), 11–27. <https://doi.org/10.1177/0014402914532234>.
- Al Otaiba, S., McMaster, K., Wanzek, K., & Zaru, M. (2022). What we know and need to know about literacy interventions for elementary students with reading difficulties and disabilities, including dyslexia. *Reading Research Quarterly*, *58*(2), 313–332. <https://doi.org/10.1002/rrq.458>.

- Austin, C. R., Vaughn, S., & McClelland, A. M. (2017). Intensive reading interventions for inadequate responders in grades K–3: A synthesis. *Learning Disability Quarterly*, 40(4), 191–210. <https://doi.org/10.1177/0731948717714446>.
- Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). *Evaluation of response to intervention practices for elementary school reading (NCEE 2016–4000)*. National Center for Education Evaluation and Regional Assistance.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Berkeley, S., Scanlon, D., Bailey, T. R., Sutton, J. C., & Sacco, D. M. (2020). A snapshot of RTI implementation a decade later: New picture, same story. *Journal of Learning Disabilities*, 53(5), 332–342. <https://doi.org/10.1177/0022219420915867>.
- Burns, M. K., Maki, K. E., Brann, K. L., McComas, J. J., & Helman, L. A. (2020). Comparison of reading growth among students with severe reading deficits who received intervention to typically achieving students and students receiving special education. *Journal of Learning Disabilities*, 53(6), 444–453. <https://doi.org/10.1177/0022219420918840>.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19(1), 5–51. <https://doi.org/10.1177/1529100618772271>.
- Center on Response to Intervention (2014). RTI fidelity of implementation rubric. *American Institutes of Research*. <http://https://mtss4success.org/resource/essential-components-mtss-rubric>.
- Ciullo, S., Ely, E., McKenna, J. W., Alves, K. D., & Kennedy, M. J. (2019). Reading instruction for students with learning disabilities in grades 4 and 5: An observation study. *Learning Disability Quarterly*, 42(2), 67–79. <https://doi.org/10.1177/0731948718806654>.
- Connor, C. M., Morrison, F. J., & Underwood, P. (2007). A second chance in second grade? The independent and cumulative impact of first and second grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading*, 11(3), 199–233. <https://doi.org/10.1080/10888430701344314>.
- Connor, C. M., Lara, J., Crowe, J., E. C., & Meadows, J. G. (2009a). Instruction, student engagement, and reading skill growth in reading first classrooms. *The Elementary School Journal*, 109(3), 221–250. <https://doi.org/10.1086/592305>
- Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., Underwood, P., & Morrison, F. J. (2009b). Individualizing student instruction precisely: Effects of child × instruction interactions on first graders' literacy development. *Child Development*, 80(1), 77–100. <https://doi.org/10.1111/j.1467-8624.2008.01247.x>.
- Coyne, M. D., Oldham, A., Dougherty, S. M., Leonard, K., Koriakin, T., Gage, N. A., & Gillis, M. (2018). Evaluating the effects of supplemental reading intervention within an MTSS or RTI reading reform initiative using a regression discontinuity design. *Exceptional Children*, 84(4), 350–367. <https://doi.org/10.1177/0014402918772791>.
- Donegan, R. E., & Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: A look at recent research. *Reading and Writing*, 34(8), 1943–1977. <https://doi.org/10.1007/s11145-021-10123-y>.
- Edmonds, M., & Briggs, K. L. (2003). The instructional content emphasis instrument: Observations of reading instruction. In S. Vaughn, & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 31–52). Brookes.
- Fien, H., Nelson, N. J., Smolkowski, K., Kosty, D., Pilger, M., Baker, S. K., & Smith, J. L. M. (2021). A conceptual replication study of the enhanced core reading instruction MTSS-reading model. *Exceptional Children*, 87(3), 265–288. <https://doi.org/10.1177/0014402920953763>.
- Fuchs, D., Fuchs, L. S., & Stecker, P. M. (2010). The blurring of special education in a new continuum of general education placements and services. *Exceptional Children*, 76(3), 301–323. <https://doi.org/10.1177/001440291007600304>.
- Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2009). *Assisting students struggling with reading: Response to intervention and multi-tier intervention for reading in the primary grades. A practice guide (NCEE 2009–4045)*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *Journal of Research on Educational Effectiveness*, 13(2), 401–427. <https://doi.org/10.1080/19345747.2019.1689591>

- Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10. <https://doi.org/10.1177/074193258600700104>.
- Hall, C., Dahl-Leonard, K., Cho, E., Solari, E. J., Capin, P., Conner, C. L., Henry, A. R., Cook, L., Hayes, L., Vargas, I., Richmond, C. L., & Kehoe, K. F. (2022). Forty years of reading intervention research for elementary students with or at risk for dyslexia: A systematic review and meta-analysis. *Reading Research Quarterly*. <https://doi.org/10.1002/rrq.477>. Advance online publication.
- Hudson, R. F., Anderson, M., & McGraw, E. M. (2022). Structured literacy interventions for reading fluency. In L. Spear-Swerling (Ed.), *Structured literacy interventions: Teaching students with reading difficulties, grades K-6* (pp. 43–64). Guilford.
- Individuals with Disabilities Education Improvement Act, 20 U.S.C. § 1400 et seq (2004).
- Kearns, D. M., Lyon, D. P., & Kelley, S. L. (2022). Structured literacy interventions for reading long words. In L. Spear-Swerling (Ed.), *Structured literacy interventions: Teaching students with reading difficulties, grades K-6* (pp. 43–64). Guilford.
- Lenth, R. (2022). *_emmeans*: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.5, <<https://CRAN.R-project.org/package=emmeans>.
- Mathes, P. G., & Fuchs, L. S. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review*, 23(1), 59–80. <https://doi.org/10.1080/02796015.1994.12085695>.
- National Center on Improving Literacy (2023). State of Dyslexia. Retrieved from <https://improvingliteracy.org/state-of-dyslexia>.
- National Reading Panel (US), National Institute of Child Health, & Human Development (US) (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*.
- National Center on Response to Intervention. (2010). Essential components of RTI – A closer look at response to intervention. *Office of Special Education Programs, National Center on response to intervention*. U.S. Department of Education. https://mtss4success.org/sites/default/files/2020-07/rti-essentialcomponents_042710.pdf.
- Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., Lonigan, C. J., Phillips, B. M., Schatschneider, C., Steacy, L. M., Terry, N. P., & Wagner, R. K. (2020). How the science of reading informs 21st-century education. *Reading Research Quarterly*, 55(S1), S267–S282. <https://doi.org/10.1002/rrq.352>.
- Ruffini, S. J., Miskell, R., Lindsay, J., McInerney, M., & Waite, W. (2016). *Measuring the implementation fidelity of the response to Intervention Framework in Milwaukee Public Schools*. REL 2017–192. Regional Educational Laboratory Midwest.
- Russell Freudenthal, D., Zaru, M. W., & Al Otaiba, S. (2023). Early literacy and multi-tiered systems of support. In S. Cabell, S. Neuman, & N. Patton Terry (Eds.), *Handbook on the Science of Early Literacy*.
- Section 504 of the Rehabilitation Act of 1973 Regulations, 34 C.F.R. § 104.1 et seq.
- Stewart, J., Otaiba, A., S., & Ciullo, S. (2024). *Observing reading instruction for elementary students with or at-risk for learning disabilities: A synthesis. [Manuscript in preparation]*. Department of Teaching and Learning. Southern Methodist University.
- Swanson, E. A., & Vaughn, S. (2010). An observation study of reading instruction provided to elementary students with learning disabilities in the resource room. *Psychology in the Schools*, 47(5), 481–492. <https://doi.org/10.1002/pits.20484>.
- Swanson, E., Solis, M., Ciullo, S., & McKenna, J. W. (2012). Special education teachers' perceptions and instructional practices in response to intervention implementation. *Learning Disability Quarterly*, 35(2), 115–126. <https://doi.org/10.1177/0731948711432510>.
- U.S. Department of Education Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Reading Assessment.
- U.S. Department of Education Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Reading Assessment.
- U.S. Department of Education Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2022 Reading Assessment.
- Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*, 18(3), 137–146. <https://doi.org/10.1111/1540-5826.00070>.







Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools*, 47(5), 432–444. <https://doi.org/10.1002/pits.20481>.

What Works Clearinghouse (2017). Procedures and standards handbook (version 4.0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf.

Youman, M., & Mather, N. (2018). Dyslexia laws in the USA: A 2018 update. *Perspectives on Language and Literacy*, 44(2), 37–41. <https://dyslexiaida.org/perspectives/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Stephanie Al Otaiba¹  · Jennifer Stewart³  · Wilhelmina van Dijk²  ·
Carlin Conner³  · Dayna Russell Freudenthal^{1,4}  · Brenna Rivas¹ ·
Paul Yovanoff¹ · Jill Allor¹ 

✉ Stephanie Al Otaiba
salotaiba@smu.edu

¹ Southern Methodist University, Dallas, US

² University of Virginia, Charlottesville, VA, US

³ Utah State University, Logan, US

⁴ Mount St. Joseph University, Cincinnati, OH, US