

Delivered as Designed: Implementation Adherence, Responsiveness, and Reading Growth in Tiered Literacy Intervention

Jolanta Jonak, Ed.D.

Department of Literacy, Leadership, Development and Special Education, Northeastern Illinois University,
Chicago, Illinois, USA

Corresponding Author: Jolanta Jonak

ABSTRACT: This retrospective practice-based study examined reading growth, observed implementation adherence, and instructional responsiveness within school-based Tier 2 small-group literacy intervention. Archival assessment and implementation records were available for 20 students in Grades 1–3. The primary outcome was AIMSweb curriculum-based oral reading fluency, expressed as words read correctly per minute (WRC), measured at fall, winter, and spring benchmarks. Complete benchmark data were available for nine students in each intervention cohort. The SLANT cohort increased from 23.67 WRC (SD = 27.93) in fall to 52.33 (SD = 40.02) in spring, a mean gain of 28.67 WRC, $t(8) = 5.80$, $p < .001$, $d_z = 1.93$, 95% CI [17.27, 40.07]. The My Sidewalks cohort increased from 31.67 WRC (SD = 24.74) to 81.44 (SD = 19.68), a mean gain of 49.78 WRC, $t(8) = 7.60$, $p < .001$, $d_z = 2.53$, 95% CI [34.67, 64.88]. Across 15 implementation observations and 209 scorable checklist ratings, overall observed adherence was 92.3%, with all scored teacher-instruction and materials-use indicators observed. Field notes and intervention records further documented responsiveness through instructional scaffolding, corrective feedback, changes in dosage and grouping, and data-based movement between levels of support. Together, the findings show how reading growth, strong procedural adherence, and responsive adaptation operated concurrently within Tier 2 intervention.

KEYWORDS: MTSS; Tier 2 literacy intervention; implementation adherence; instructional responsiveness; oral reading fluency; progress monitoring

Date of Submission: 12-09-2026

Date of acceptance: 23-09-2026

I. INTRODUCTION

Reading proficiency provides a foundation for academic learning, yet some students experience persistent difficulty acquiring the foundational skills needed for accurate, fluent, and increasingly complex reading. Within multi-tiered systems of support (MTSS), students who demonstrate inadequate progress in core instruction may receive supplemental intervention of increasing intensity. Tier 2 is designed to provide timely, targeted small-group instruction for students who require more support than core classroom instruction alone can provide (Fuchs & Fuchs, 2006).

Research on reading development has established the importance of explicit and systematic instruction in foundational skills, including phonological awareness, sound-symbol relationships, decoding, word recognition, fluency, vocabulary, and comprehension (Castles et al., 2018; National Institute of Child Health and Human Development [NICHD], 2000). Contemporary models further emphasize that skilled reading develops through interacting word-recognition and language-comprehension processes rather than through any single instructional component (Duke & Cartwright, 2021). Structured literacy approaches place particular emphasis on explicit, systematic, and cumulative instruction in language structure and are frequently recommended for students experiencing persistent word-reading difficulty (Spear-Swerling, 2019).

Selecting an evidence-based instructional approach is only one part of the intervention process. Student outcomes also depend on whether critical intervention components are delivered consistently, whether students receive sufficient instructional opportunity, and whether educators recognize and respond when progress is inadequate. In Tier 2 systems, implementation adherence, progress monitoring, and instructional responsiveness therefore operate as interconnected parts of the same instructional process.

Tier 2 Literacy Intervention and Progress Monitoring

Tiered intervention systems organize instructional support according to student need. Tier 1 consists of core instruction provided to all students, Tier 2 adds targeted supplemental small-group intervention for students demonstrating elevated risk or insufficient progress, and Tier 3 provides increasingly intensive and

individualized support (Fuchs & Fuchs, 2006). Meta-analytic research indicates that supplemental reading intervention can produce meaningful improvements for students experiencing early reading difficulty, although response varies substantially across students and instructional contexts (Wanzek et al., 2016).

Because students respond differently to supplemental instruction, progress monitoring is central to Tier 2 practice. Curriculum-based measurement was developed to provide brief, repeatable indicators of academic performance that are sufficiently sensitive to change to support instructional decision-making (Deno, 2003). Oral reading fluency, commonly expressed as words read correctly per minute, is widely used within reading MTSS because it can be administered repeatedly and interpreted in relation to performance and growth over time. Repeated data help educators determine whether current intervention is producing adequate growth and whether instruction should be maintained, modified, or intensified (Stecker et al., 2005). Progress monitoring thus serves both a measurement function and a decision-making function within Tier 2.

Implementation Adherence Within a Multidimensional Fidelity Framework

Progress-monitoring data become more informative when teams also know how the intervention was delivered. Implementation fidelity, also referred to as treatment or intervention integrity, concerns the extent to which an intervention is implemented as intended. The construct is multidimensional. Dane and Schneider (1998) identified adherence, exposure, quality of delivery, participant responsiveness, and program differentiation as distinct aspects of program integrity. Carroll et al. (2007) similarly distinguished adherence from contextual and moderating factors affecting implementation, while broader implementation research has emphasized the relationship between implementation processes and intervention outcomes (Durlak & DuPre, 2008; O'Donnell, 2008).

School-based treatment-integrity literature cautions against treating implementation as a single percentage of completed procedures. Hagermoser Sanetti and Kratochwill (2009) described integrity in terms of content, quality, quantity, and process, and Gresham (2009) distinguished treatment adherence from interventionist competence and treatment differentiation. Noell and Gansle (2006) further emphasized that adequate intervention implementation is essential to interpreting response within RTI and that treatment strength, progress monitoring, decision rules, and implementation adequacy are interrelated.

These concerns remain current in literacy intervention research. Dahl-Leonard et al. (2023) found that fidelity reporting in foundational reading-intervention studies remained disproportionately focused on adherence and dosage, with considerably less attention to quality, responsiveness, and differentiation. Van Dijk et al. (2023) likewise documented substantial variation in how reading-intervention studies conceptualize and measure fidelity. More recent reviews continue to identify gaps in reporting fidelity and alignment within Tier 2 literacy research (King et al., 2026; Young et al., 2026), while school-based implementation research demonstrates that contextual barriers and dosage influence how reading interventions operate under routine educational conditions (Mesa et al., 2024).

For the present study, direct observation provided a structured way to document whether specified instructional procedures were present during intervention sessions. This focus on observable adherence complements, rather than substitutes for, the broader multidimensional construct of implementation fidelity. It also creates a practical bridge between the implementation literature and the day-to-day instructional behaviors captured in the archival records.

Instructional Responsiveness and Data-Based Decision-Making

Consistency and adaptation are sometimes treated as competing demands, yet Tier 2 instruction requires both. Core intervention features need sufficient consistency to preserve the instructional model, while teachers must also respond to learner variability. Harn et al. (2013) framed this balance in terms of fidelity, flexibility, and contextual fit.

Data-based individualization provides one framework for this process. Progress-monitoring information can be used to identify inadequate response and guide deliberate changes in dosage, grouping, instructional focus, scaffolding, or other dimensions of intervention intensity (Fuchs et al., 2017; Jung et al., 2018). Practice-based studies similarly emphasize the organizational processes needed to use data for increasingly intensive reading support (Kearns et al., 2022; Murdoch et al., 2024), and recent work demonstrates the feasibility of systematic data-based intensification when student response is insufficient (Filderman et al., 2025).

Seen this way, responsiveness operates at two interconnected levels. Within a lesson, teachers can respond to errors through additional modeling, corrective feedback, segmentation, or guided practice. At the system level, accumulated progress-monitoring data can lead to changes in instructional time, group size, intervention approach, or movement between levels of support. Contemporary observational work likewise underscores the importance of examining what students actually receive in tiered reading systems (Al Otaiba et al., 2025).

Present Study

Against this backdrop, the present study brought student outcome data and implementation records together to examine Tier 2 small-group literacy intervention as it operated in practice. The analysis focused on three connected dimensions of the intervention process: oral reading fluency growth, observed implementation adherence, and instructional responsiveness.

The study addressed three research questions:

1. What patterns of oral reading fluency growth were observed within each Tier 2 intervention cohort across the academic year?
2. What level and pattern of implementation adherence were documented through direct observation of intervention delivery?
3. What forms of instructional responsiveness and intervention adjustment were documented during Tier 2 implementation?

II. METHODOLOGY

Research Design

This study used a retrospective, practice-based pre-post design combining quantitative student outcomes with structured observational and documentary implementation evidence. The underlying assessment and implementation records were collected contemporaneously during routine school-based intervention implementation and were later analyzed as an archival dataset. The term practice-based reflects the naturally occurring school intervention context. The archival records were originally collected during the 2011–2012 academic year. Students and teachers followed district-established intervention arrangements rather than researcher-assigned treatment conditions. The design therefore centered on within-cohort growth and the implementation processes surrounding that growth.

Setting and Tier 2 Intervention Context

The study was conducted within elementary school settings in one suburban public school district in the Midwestern United States. At the time of the study, the district implemented reading supports through a Response to Intervention (RTI) framework, using universal screening, benchmark assessment, progress monitoring, and problem-solving procedures to guide student placement and movement across levels of support. These practices are consistent with the broader MTSS framework commonly used today.

All participants in the present analytic sample received Tier 2 small-group literacy intervention. District procedures used AIMSweb and MAP benchmark data, program assessments, teacher information, and problem-solving-team review to inform placement and instructional decisions. Students receiving Tier 2 support were monitored repeatedly, and inadequate progress could lead to increased instructional intensity, regrouping, or movement to a different level or approach of intervention.

Participants

The archival dataset included 20 students in Grades 1–3 receiving Tier 2 small-group reading intervention delivered by four elementary reading interventionists. Students had been identified for supplemental support through existing district screening and progress-monitoring procedures. The SLANT cohort included nine students, and the My Sidewalks cohort included 11 students. Complete fall, winter, and spring oral reading fluency benchmark data were available for all nine SLANT students and nine My Sidewalks students.

Table 1
Grade Composition and Availability of Benchmark Oral Reading Fluency Data

Characteristic	SLANT	My Sidewalks
Total students	9	11
Grade 1	5	5
Grade 2	1	6
Grade 3	3	0
Complete Fall-Winter-Spring ORF data	9	9

Note. Cohorts were analyzed separately because expected oral reading fluency performance and growth vary by grade.

Tier 2 Intervention Programs

Two Tier 2 supplemental literacy programs were represented in the archival records: the SLANT System for Structured Language Training and My Sidewalks. SLANT provided systematic, sequential instruction emphasizing phonological awareness, sound-symbol relationships, decoding, spelling, fluency, vocabulary, and comprehension. Archival program materials describe teacher modeling, guided practice,

cumulative review, grapheme-phoneme work, blending, word reading, spelling, decodable text, and sentence-level application.

My Sidewalks served as the district's strategic Tier 2 literacy program. Observed sessions included explicit sound and word work, decoding of unfamiliar and nonsense words, guided and choral reading, vocabulary instruction, comprehension discussion, partner reading, and teacher modeling. Both programs were delivered in small-group settings. Within the district's Tier 2 problem-solving framework, instructional intensity could be adjusted in response to student progress.

Measures

AIMSweb oral reading fluency was the primary student outcome and was expressed as words read correctly per minute (WRC). Benchmark scores were available across fall, winter, and spring assessment periods. Fall and spring values were used for primary within-cohort pre-post analyses, while winter values were retained as intermediate indicators of growth trajectory. Complete benchmark series were available for nine students in each cohort. Missing outcomes were not imputed.

AIMSweb-generated rate-of-improvement (ROI) values from the district's progress-monitoring records were also summarized as descriptive indicators of growth. Measures of Academic Progress (MAP) reading RIT scores were available for a subset of participants and were treated as secondary indicators of reading growth. Complete fall-to-spring MAP data were available for eight students in the SLANT cohort and seven students in the My Sidewalks cohort.

Implementation Adherence Procedures

Implementation was assessed through repeated direct observations using a structured reading-study observation checklist. The retained adherence dataset comprised 15 observation records: 10 associated with SLANT instruction and five with My Sidewalks. The checklist contained 14 indicators organized into three domains: classroom environment, teacher instruction, and materials use.

Classroom-environment indicators addressed seating for engagement, program visuals, student-work display, posted rules or expectations, and materials organization. Teacher-instruction indicators addressed pacing, adjustment of instruction based on learner response, corrective feedback and positive reinforcement, active engagement, and direct instruction. Materials-use indicators addressed use of program materials, use of teacher materials, alignment with the designated lesson, and use of supplemental materials supporting the base program.

Two university observers conducted the implementation observations. Archival documentation indicates that observer preparation included training with videotaped instruction and classroom examples and discussion of checklist indicators. A calibration exercise using a videotaped SLANT lesson was reported to produce complete agreement among participating coaches. Because session-specific inter-rater coefficients were not retained for the live observations analyzed here, this calibration result is treated as evidence of observer preparation rather than as a reliability estimate for the complete observational dataset. Because the checklist focused on directly observable components of intervention delivery, it served as the study's measure of procedural adherence across three domains: classroom environment, teacher instruction, and materials use.

Instructional Responsiveness Evidence

Field notes accompanying the structured observations documented instructional interactions, student errors, teacher feedback, instructional scaffolding, grouping, intervention time, and placement changes. District implementation materials provided additional information concerning progress-monitoring expectations and decision rules. The records were reviewed for two complementary forms of responsiveness. Within-session responsiveness captured teacher adjustments to immediate student performance, including additional modeling, error correction, segmentation, cueing, guided practice, and other scaffolding. System-level responsiveness captured changes in dosage, group composition, intervention approach, or placement based on accumulated student-performance information. The analysis emphasized recurring patterns documented across the observation notes and intervention records.

Data Collection Procedures

Student assessment information was obtained from existing district benchmark and intervention records. Fall, winter, and spring AIMSweb benchmark scores formed the primary longitudinal outcome series, while MAP data provided a secondary measure where complete scores were available. Implementation information was obtained through completed observation checklists, field notes, intervention records, and district documentation describing Tier 2 decision procedures. Student records were maintained using participant identification numbers, and the present analysis used de-identified archival records. The original data collection

was conducted under institutional review procedures in effect at the time. The present retrospective analysis used de-identified archival records and involved no new contact with students, families, or school personnel.

Data Analysis

Descriptive statistics were calculated separately for each intervention cohort. For the primary oral reading fluency analysis, paired-samples *t* tests examined fall-to-spring change among students with complete paired benchmark scores. Cohen's *d_z* was calculated as the mean paired difference divided by the standard deviation of the paired difference scores. Ninety-five percent confidence intervals were calculated for mean fall-to-spring change.

Given the small number of complete cases, exact Wilcoxon signed-rank tests were conducted as sensitivity analyses to confirm the fall-to-spring pattern identified by the paired-samples *t* tests. Winter oral reading fluency scores and AIMSweb-generated ROI values were summarized descriptively. MAP fall-to-spring change was examined descriptively for students with complete data.

Each cohort was analyzed separately, consistent with the study questions and the cohorts' differing grade composition. Implementation adherence was summarized using frequencies and percentages overall and by checklist domain. Observational field notes and intervention records were reviewed to characterize documented forms of instructional responsiveness and data-based intervention adjustment.

III. FINDINGS

Oral Reading Fluency Growth

Complete fall-to-spring AIMSweb WRC data were available for nine students in each intervention cohort. Both cohorts demonstrated statistically significant increases in oral reading fluency across the academic year.

Table 2

AIMSweb Oral Reading Fluency Across Benchmark Periods

Cohort	n	Fall WRC M (SD)	Winter WRC M (SD)	Spring WRC M (SD)	Mean Gain M (SD)
SLANT	9	23.67 (27.93)	41.11 (38.93)	52.33 (40.02)	28.67 (14.83)
My Sidewalks	9	31.67 (24.74)	48.33 (21.27)	81.44 (19.68)	49.78 (19.65)

Note. WRC = words read correctly per minute. Winter values are presented descriptively; inferential analyses compare fall and spring benchmark scores within each cohort.

Students in the SLANT cohort increased from a fall mean of 23.67 WRC (SD = 27.93) to a spring mean of 52.33 WRC (SD = 40.02), representing a mean increase of 28.67 WRC (SD = 14.83), $t(8) = 5.80$, $p < .001$, $d_z = 1.93$, 95% CI [17.27, 40.07]. Students in the My Sidewalks cohort increased from a fall mean of 31.67 WRC (SD = 24.74) to a spring mean of 81.44 WRC (SD = 19.68), representing a mean increase of 49.78 WRC (SD = 19.65), $t(8) = 7.60$, $p < .001$, $d_z = 2.53$, 95% CI [34.67, 64.88]. Exact Wilcoxon signed-rank sensitivity analyses confirmed the pattern identified by the paired-samples *t* tests, with significant positive fall-to-spring change in both cohorts (both exact $p = .004$).

Growth Trajectory and Rate of Improvement

The intermediate benchmark data showed a clear upward trajectory in both cohorts. SLANT mean WRC increased from 23.67 in fall to 41.11 in winter and 52.33 in spring. My Sidewalks mean WRC increased from 31.67 in fall to 48.33 in winter and 81.44 in spring. AIMSweb-generated ROI values were also positive, averaging 0.82 WRC per week (SD = 0.37) for SLANT and 1.33 WRC per week (SD = 0.32) for My Sidewalks. Individual trajectories added another layer to the group-level pattern. Some students entered intervention with very low fluency levels, whereas others began at higher baseline levels, and the magnitude of improvement varied across students. This variability underscores the value of considering benchmark growth together with ongoing progress-monitoring and instructional-response information.

Secondary Reading Outcomes

Complete fall-to-spring MAP data were available for eight students in the SLANT cohort and seven students in the My Sidewalks cohort. Students in the SLANT cohort with complete MAP data demonstrated a mean gain of 19.1 RIT points (SD = 9.7), while students in the My Sidewalks cohort demonstrated a mean gain of 17.1 RIT points (SD = 10.4). These secondary outcomes paralleled the positive growth pattern observed in oral reading fluency.

Observed Implementation Adherence

Across the 15 retained observations, 209 of 210 possible checklist ratings were scorable. Of these, 193 were rated Observed, yielding an overall observed-adherence rate of 92.3%. All scored teacher-instruction and

materials-use indicators were observed. Variability occurred within classroom-environment indicators, particularly items involving display of student work and program visuals.

Table 3
Observed Implementation Adherence Across 15 Observations

Domain	Scorable Ratings	Observed	Not Observed	Observed Adherence
Classroom environment	74	58	16	78.4%
Teacher instruction	75	75	0	100.0%
Materials use	60	60	0	100.0%
Overall	209	193	16	92.3%

Note. One classroom-environment rating was not scorable in the archived checklist data. Percentages therefore reflect scorable ratings only.

Taken together, the observation data show consistently high procedural adherence, with the strongest consistency in teacher instruction and materials use. Classroom-environment indicators were more variable, creating a useful domain-level picture of how intervention delivery looked across observed sessions.

Instructional Responsiveness and Data-Based Decision-Making

The adherence findings were accompanied by clear evidence of instructional responsiveness. Within individual sessions, teachers responded to student errors through additional modeling, cueing, segmentation, repeated practice, explicit correction, and other forms of scaffolding. Field notes described, for example, immediate correction during decoding, phonological cues for unfamiliar words, and tactile or visual supports when students had difficulty applying previously taught patterns.

Responsiveness also extended beyond the lesson itself. Students demonstrating insufficient progress could receive increased instructional time, smaller or reorganized groups, or movement to a different intervention approach, while students demonstrating adequate progress could return to less intensive support. These decisions drew on multiple sources of performance information, including AIMSweb, MAP, curriculum assessments, and problem-solving-team review.

Table 4
Documented Forms of Instructional Responsiveness

Responsiveness Domain	Illustrative Evidence in Archival Records
Error-responsive instruction	Immediate correction, additional modeling, segmentation, and word-part analysis
Scaffolding	Cueing, repeated guided practice, tactile or visual supports, tracing and sounding
Instructional emphasis	Additional attention to difficult sound patterns or unfamiliar words
Dosage	Increased instructional time when progress was insufficient
Grouping	Division or reorganization of groups and movement to smaller instructional settings
Intervention change	Movement to another intervention approach following inadequate response
Movement between support levels	Return to less intensive support after adequate progress or consideration of more intensive support after insufficient progress
Data-based placement	Use of AIMSweb, MAP, curriculum assessments, and problem-solving-team review

Note. Examples illustrate recurring forms of responsiveness documented across observation and intervention records.

IV. DISCUSSION

Reading Growth Within Tier 2 Intervention

Across both cohorts, oral reading fluency increased significantly from fall to spring, with winter benchmarks showing continued upward movement across the year. Available MAP scores showed a similar positive pattern, adding a second indicator of reading growth. The value of these outcomes becomes clearer when they are considered alongside the implementation data. Student growth occurred in a context of high observed adherence to instructional and materials-related procedures and active use of progress-monitoring information to adjust support. In this sense, the study captures not only whether students improved, but also how Tier 2 intervention was delivered and adapted as that growth unfolded. This combined view of outcomes and implementation addresses a continuing need in Tier 2 literacy research to document both intervention delivery and the instructional context surrounding student response (King et al., 2026; Young et al., 2026).

Implementation Adherence Within a Broader Fidelity Construct

The adherence data add important context to the growth findings. Overall observed adherence reached 92.3%, and every scored teacher-instruction and materials-use indicator was observed. The domain pattern is especially informative. Classroom-environment adherence was 78.4%, compared with 100% observed adherence for teacher instruction and materials use. This contrast shows why implementation is more informative when examined by domain: the interventionists consistently demonstrated the observed instructional and materials-use practices even when environmental features were less uniformly present.

A single global adherence percentage can flatten these differences. Treatment-integrity scholarship has long noted that intervention components vary in substantive importance and that component-level information can clarify what was actually delivered during instruction (Gresham, 2009; Noell & Gansle, 2006). Viewed in this way, the checklist yielded more than an overall adherence score; it identified the areas in which implementation was most consistent. This domain-focused interpretation aligns with contemporary calls for richer, multidimensional reporting of fidelity in literacy intervention research (Dahl-Leonard et al., 2023; van Dijk et al., 2023).

Adherence and Responsiveness as Complementary Processes

More importantly, high adherence did not preclude adaptation. Teachers maintained the instructional routines represented on the observation checklist while adjusting support in response to student performance. Within sessions, this included additional modeling, corrective feedback, guided decoding, cueing, and other forms of scaffolding. Across time, responsiveness included changes in instructional time, group composition, intervention approach, and movement between levels of support.

These patterns reveal two complementary forms of responsiveness. Instructional responsiveness occurred during teaching as educators reacted to immediate evidence of understanding or difficulty. System-level responsiveness occurred as accumulated progress-monitoring information informed decisions about dosage, grouping, intervention approach, and placement. Together, these practices illustrate how fidelity and purposeful adaptation can operate within the same intervention system (Harn et al., 2013; Hagermoser Sanetti & Kratochwill, 2009).

The district framework gave this responsiveness an organizational structure. Progress monitoring was directly connected to decisions about continued Tier 2 instruction, increased intensity, regrouping, intervention change, and movement to less or more intensive support. Responsiveness was therefore embedded both in day-to-day instruction and in the broader problem-solving process.

Implications for MTSS Practice

For school-based MTSS practice, the findings highlight the value of examining implementation at the level of meaningful instructional components. A global adherence score is useful, but domain-level information shows more clearly where delivery is consistently strong and where implementation may vary. That level of detail can help teams connect implementation information to the student-performance data they are already using within problem-solving routines.

The findings also reinforce the practical value of linking implementation monitoring with progress monitoring. Dosage, quality of delivery, student engagement, instructional responsiveness, and differentiation help teams understand the instructional conditions surrounding student growth. In the present setting, student data were connected to decisions about continued Tier 2 support, additional instructional time, regrouping, intervention change, and movement between levels of support. Clearly identifying which intervention elements constitute the instructional core and which features can be adapted responsively allows schools to preserve consistent delivery of essential practices while adjusting pacing, scaffolding, grouping, or dosage according to student performance.

Limitations

The practice-based design places several boundaries around the findings. The sample was small and drawn from a single school district, and the cohorts differed in grade composition. Students were not randomly assigned to intervention programs, and the study did not include an untreated comparison group. The findings therefore characterize within-cohort growth and implementation in this setting rather than isolate a program-specific causal effect.

Outcome data were also incomplete for some participants. Complete oral reading fluency benchmark data were available for nine students in each cohort, and complete MAP data were available for smaller subsets. Complete-case analyses were used without imputation. In addition, raw WRC scores vary by grade and stage of reading development, which is particularly relevant given the cohorts' unequal grade composition.

The observation system captured procedural adherence effectively but did not represent every dimension of implementation fidelity. The checklist was a pragmatic implementation-monitoring tool rather than a validated program-specific fidelity instrument for both interventions, and its indicators were not weighted by theoretical importance. Although two observers participated in data collection and archival documentation describes observer training and calibration, session-specific inter-rater reliability statistics were not retained for the live observations. Direct observation also sampled selected sessions rather than implementation across the entire year.

Instructional responsiveness was reconstructed from field notes, intervention records, and district documentation rather than from a standardized responsiveness measure. These sources provided concrete

evidence of responsive practices but did not support frequency estimates or direct linkage of specific adaptations to individual student outcomes. Finally, the records come from an earlier period of RTI implementation. Terminology and implementation frameworks have continued to evolve, but the central questions examined here of how intervention is delivered, how student response is monitored, and how support is adapted when needed, remain central to contemporary Tier 2 literacy practice. Future studies can extend this work with larger samples, repeated multidimensional implementation measures, and designs that link specific implementation features to individual response.

V. CONCLUSION

Taken together, this study documents a Tier 2 system in which students demonstrated significant reading growth while interventionists maintained high procedural adherence and adjusted instruction in response to student performance. The combined outcome and implementation data provide a view of intervention as it unfolded in everyday school practice. Adherence and responsiveness emerged as complementary features of that process. Teachers and intervention teams preserved core instructional routines while adapting scaffolding, dosage, grouping, intervention approach, and levels of support as student needs changed. This combination of consistent delivery, ongoing progress monitoring, and purposeful adjustment is central to responsive tiered intervention. Examining these dimensions together provides a fuller account of Tier 2 practice than student outcomes or global adherence percentages considered in isolation.

REFERENCES

- [1]. Al Otaiba, S., Stewart, J., van Dijk, W., Conner, C., Russell Freudenthal, D., Rivas, B., Yovanoff, P., & Allor, J. (2025). 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. *Reading and Writing*, 38(4), 1129–1151. <https://doi.org/10.1007/s11145-024-10534-7>
- [2]. Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2, Article 40. <https://doi.org/10.1186/1748-5908-2-40>
- [3]. 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>
- [4]. Dahl-Leonard, K., Hall, C., Capin, P., Solari, E. J., Demchak, A., & Therrien, W. J. (2023). Examining fidelity reporting within studies of foundational reading interventions for elementary students with or at risk for dyslexia. *Annals of Dyslexia*, 73(2), 288–313. <https://doi.org/10.1007/s11881-023-00279-3>
- [5]. Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23–45. [https://doi.org/10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3)
- [6]. Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37(3), 184–192. <https://doi.org/10.1177/00224669030370030801>
- [7]. Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly*, 56(S1), S25–S44. <https://doi.org/10.1002/rrq.411>
- [8]. Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>
- [9]. Filderman, M. J., Stewart, A. A., Cramer, A. M., Hughes-Berheim, S. S., & Swanson, E. (2025). Intensification of a reading comprehension intervention using student data: A pilot study. *Remedial and Special Education*, 46(5), 382–394. <https://doi.org/10.1177/07419325241249767>
- [10]. Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41(1), 93–99. <https://doi.org/10.1598/RRQ.41.1.4>
- [11]. Fuchs, L. S., Fuchs, D., & Malone, A. S. (2017). The taxonomy of intervention intensity. *TEACHING Exceptional Children*, 50(1), 35–43. <https://doi.org/10.1177/0040059917703962>
- [12]. Gresham, F. M. (2009). Evolution of the treatment integrity concept: Current status and future directions. *School Psychology Review*, 38(4), 533–540. <https://doi.org/10.1080/2372966X.2009.12527793>
- [13]. Hagermoser Sanetti, L. M., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, 38(4), 445–459. <https://doi.org/10.1080/2372966X.2009.12527788>
- [14]. Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children*, 79(2), 181–193. <https://doi.org/10.1177/0014402913079002051>
- [15]. Jung, P.-G., McMaster, K. L., Kunkel, A. K., Shin, J., & Stecker, P. M. (2018). Effects of data-based individualization for students with intensive learning needs: A meta-analysis. *Learning Disabilities Research & Practice*, 33(3), 144–155. <https://doi.org/10.1111/ldrp.12172>
- [16]. Kearns, D. M., Walker, M. A., Borges, J. C., & Duffy, M. E. (2022). Can reading practitioners and researchers improve intensive reading support systems in a large urban school system? *Journal of Research in Reading*, 45(3), 488–516. <https://doi.org/10.1111/1467-9817.12406>
- [17]. King, S., Young, M. K., McFadden, K., Abdelnaby, H., Yoon, H., & Datchuk, S. M. (2026). Missed connections: Fidelity and tier alignment in single-case research design studies of Tier 2 literacy interventions. *Learning Disability Quarterly*. Advance online publication. <https://doi.org/10.1177/07319487261449488>
- [18]. Mesa, M. P., Hall, C., Zucker, T., Dahl-Leonard, K., Oh, Y., & Denton, C. (2024). Unpacking implementation: Fidelity and barriers to implementation of a kindergarten reading intervention by teachers. *The Elementary School Journal*, 125(1), 77–105. <https://doi.org/10.1086/731257>
- [19]. Murdoch, A., Morrison, J. Q., & Strickler, W. (2024). Process evaluation of a problem-solving approach for analyzing literacy practices within a multi-tiered system of supports framework. *Behavior and Social Issues*, 33, 479–503. <https://doi.org/10.1007/s42822-024-00166-5>

- [20]. National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). U.S. Government Printing Office.
- [21]. Noell, G. H., & Gansle, K. A. (2006). Assuring the form has substance: Treatment plan implementation as the foundation of assessing response to intervention. *Assessment for Effective Intervention*, 32(1), 32–39. <https://doi.org/10.1177/15345084060320010501>
- [22]. O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84. <https://doi.org/10.3102/0034654307313793>
- [23]. Spear-Swerling, L. (2019). Structured literacy and typical literacy practices: Understanding differences to create instructional opportunities. *TEACHING Exceptional Children*, 51(3), 201–211. <https://doi.org/10.1177/0040059917750160>
- [24]. Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 42(8), 795–819. <https://doi.org/10.1002/pits.20113>
- [25]. van Dijk, W., Lane, H. B., & Gage, N. A. (2023). How do intervention studies measure the relation between implementation fidelity and students' reading outcomes? A systematic review. *The Elementary School Journal*, 124(1), 56–84. <https://doi.org/10.1086/725672>
- [26]. Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of Tier 2 type reading interventions in Grades K–3. *Educational Psychology Review*, 28(3), 551–576. <https://doi.org/10.1007/s10648-015-9321-7>
- [27]. Young, M. K., King, S., Datchuk, S., McFadden, K., & Yoon, H. (2026). Fidelity of implementation and instructional alignment in Tier 2 interventions: An updated review of experimental studies. *Learning Disabilities Research & Practice*. Advance online publication. <https://doi.org/10.1177/09388982261476719>