

The Reading Horizons Implementation Integrity Rubric (RHIIR): Validation Study Summary

Shelby Danks, Ph.D.

Founder and Principal Advisor
ARKEN RESEARCH

Suggested citation:

Danks, S. (2023). *The Reading Horizons Implementation Integrity Rubric: Validation Study Summary*. ARKEN RESEARCH: Layton, UT.



TABLE OF CONTENTS

1. Introduction	3
2. About the Reading Horizons Implementation Integrity Rubric (RHIR)	3
Key Design Features.....	3
Progression of Practice.....	4
Indicators of Implementation Integrity.....	4
Sample Item.....	5
3. Development Process	6
4. Validation Studies	7
Research Questions	7
Participants	7
Data Collection Methods.....	7
Key Results	8
Summary and Implications	10
5. References	11

VALIDATION STUDY SUMMARY

1. INTRODUCTION

The purpose of this paper is to describe the key attributes of the Reading Horizons Implementation Integrity Rubric (RHIR) and to summarize the steps taken to develop and validate the tool for research and improvement use.

The Reading Horizons Implementation Integrity Rubric (RHIR) was developed, in collaboration with the curriculum and implementation team at Reading Horizons, by Arken Research to serve as an empowering framework for coaching, feedback, and reflection. The tool was initially intended to be used by grade-level teams for school-based self-assessment as an educative support for individual reflection. However, the tool proved useful as an observation protocol for implementation research studies conducted during the 2021–2022 and 2023–2023 academic years and was therefore validated for that particular use.

The development phase of the tool used a *Define-Build-Refine* model to describe the theory of action concerning the progression of practice, articulate the practices that produce desired literacy outcomes, and refine the tool based on four key design features: *rubric-based*, *educative*, *rigorous*, and *adaptable*.

The validation phase of the project involved exploring evidence of the utility and predictive validity of the tool across four implementation studies during the 2021–2022 academic year. Findings confirmed the tool's usefulness for leaders in offering coaching and support to educators. The findings also demonstrated evidence of predictive validity for key early literacy outcome measures.

2. ABOUT THE READING HORIZONS IMPLEMENTATION INTEGRITY RUBRIC (RHIR)

KEY DESIGN FEATURES

The Reading Horizons Implementation Integrity Rubric (RHIR) features a rubric-based design that is educative, rigorous, and adaptable:

- **Rubric-based:** When educators learn a new approach to teaching and learning, it is seldom that all of the desired principles, practices, and behaviors that are observable during the planning stage or during lesson facilitation are equally easy to integrate into their existing habits and routines. Rubrics enable the educator or a coach to see which practices may be “easier” than others or identify which practices could be prioritized based on the strategic goals for the educator, leaders, or district. Rubrics better articulate the natural progression of practices that educators demonstrate as they integrate the new approach into their existing habits and practices.
- **Educative, but non-prescriptive:** While rubrics are typically used for evaluative purposes (i.e., to score educator performance), their illustrative capacities offer a way to describe practices in a non-judgmental, non-evaluative manner. The RHIR illustrates what desirable practices may look like for key components of the instructional model, using examples where they are appropriate. These practices intentionally remain abstract and non-prescriptive in instances where creative adaptations are most appropriate, while retaining program integrity. Key practices or behaviors **marked with bold font** make the qualitative distinctions between the levels easier to visualize and achieve.

VALIDATION STUDY SUMMARY

- **Rigorous:** Rubrics should be designed with the journey toward instructional excellence in mind, and communicate that some practices take longer than others to integrate or improve upon. Educators who are new to the curriculum (in their first year of adoption) may exhibit the practices and behaviors found in the *Emerging* and *Exploring* levels, and that is not only acceptable but completely to be expected. Coaches can inspire educators to grow in their learning when they explicitly communicate the rigor of their work together.
- **Adaptable:** Not all indicators may be as relevant to all curricular implementations, depending on their desired use of the intended program. Coaches, educators, and implementation leaders can easily identify which indicators on this tool seem most aligned with their school goals and use those selected indicators to self-assess their progress. Taking time to prioritize which indicators to focus on and align with can in itself be a useful exercise for schools to improve their conditions for effective instruction.

PROGRESSION OF PRACTICE

Drawing from recent advances in the implementation science, each of the indicators on the RHIIR is measured using a progression of practice. The progression prompts educators to deeply reflect and assess their implementation using a growth framework, consisting of four levels: 1 - *Emerging*, 2 - *Exploring*, 3 - *Engaging*, and 4 - *Empowering*.

- **Level 1. Emerging:** The educator is still learning about this element, or is in the **early stage of planning** how to do this in the classroom.
- **Level 2. Exploring:** The educator has begun to **experiment** with this element in the classroom but has yet to form daily habits that directly impact student learning.
- **Level 3. Engaging:** The educator **consistently implements** this element, has identified lessons learned, and made minor improvements. Consistent practices translate to impacts on student learning.
- **Level 4. Empowering:** The educator is able to **leverage the full intent** of the program, making meaningful adaptations that equitably serve all students.

INDICATORS OF IMPLEMENTATION INTEGRITY

The Reading Horizons Implementation Integrity Rubric (RHIIR) contains five indicators that describe an effective Reading Horizons lesson implementation. The indicators listed below (Table 2) apply to the classroom-level implementation and were designed with the intention that the rubric could be used to observe and rate a 30–40-minute lesson.

Table 2. RHIIR List of Indicators

Indicator	Brief Description
1. Effective Communication of the RH Method for Meaningful Connections	The educator provides a clear and concise summary and visuals of key lesson concepts and helps students make meaningful connections throughout the lesson.
2. Instructional Routines	The educator uses a variety of instructional routines to ensure all students contribute to the learning community.

VALIDATION STUDY SUMMARY

3. Questioning and Monitoring	The educator embeds opportunities for students to meaningfully demonstrate mastery throughout the lesson and monitors student work to check for evidence of student mastery.
4. Feedback for Deeper Understanding	The educator provides meaningful feedback to all to deepen learning and understanding.
5. Pacing and Lesson Structure	The educator effectively chunks the lesson to focus on discrete skills and facilitates all four phases of the Daily Core 4 lesson structure.

SAMPLE ITEM

The following sample from the RHIR illustrates the progression of practice for *Indicator 1. Effective Communication of the RH Method for Meaningful Connections*. As teachers progress from *Emerging* to *Empowering* in their practice, they enhance their ability to provide clear and concise summaries of key concepts, use key resources to visualize the Reading Horizons® method, and help students make meaningful connections between past and new skills, as well as between their new skills and the transfer of those skills.

Indicator 1. Effective Communication of the RH Method for Meaningful Connections

The educator provides a clear and concise summary and visuals of key lesson concepts and helps students make meaningful connections throughout the lesson.

Emerging	Exploring	Engaging	Empowering
<p>The educator shares information but may miss key lesson concepts or replace them with materials that do not align with key concepts.</p> <p>The educator may provide too much direct instruction outside of the Skill Review and Instruction components of the Daily Core 4.</p>	<p>The educator shares key lesson concepts during the Skill Review and Instruction components of the Daily Core 4, using suggested instructions from the curriculum.</p> <p>The educator may struggle to keep it clear and concise (e.g., provides too much explanation).</p>	<p>The educator provides a clear and concise summary of key lesson concepts during the Skill Review and Instruction components of the Daily Core 4.</p> <p>The educator visualizes the RH method effectively (e.g., models how to prove words using the slide decks, cards, posters, or other tools).</p> <p>The educator references past skills at the beginning of the lesson but may not make clear connections between past and new skills throughout the lesson.</p>	<p>The educator provides a clear and concise summary of key lesson concepts during the Skill Review and Instruction components of the Daily Core 4.</p> <p>The educator visualizes the RH method effectively (e.g., models how to prove words using slide decks, posters, or other tools).</p> <p>The educator helps students make meaningful connections between past and new skills, and between new skills and Transfer.</p>

3. DEVELOPMENT PROCESS

Arken Research adapted the instrument development process recommended by Danks and Allen (2014) to design a performance-based rubric to measure the extent to which educators implement the Reading Horizons curriculum with integrity. The key steps that were used to design the rubric included a clear definition of the RH method and instructional approach, collaborative design, instrument build based on the progression of practice, iterative feedback, refinement, and the development of materials to support use (see Table 2).

Table 2. Instrument Development Process

Summary of Activities
Phase 1 - Define
1. Initiation: Summary of the RH method, instructional approach to reading instruction, theory of change, and deep document review.
2. Collaborative design meetings: Description of key practices hypothesized to produce desired outcomes.
Phase 2 - Build
1. Instrument development: Translation of the RH method and goals into clear success criteria at different points along the progression of practice.
2. Iterative feedback and review meetings: Strawman design and feedback loops.
Phase 3 - Refine
1. Final instrument design: Refinement of instruments based on feedback.
2. Front matter design and user resources: Development of materials to support leaders, instrument researchers, partners, and collaborators.

After completing the design process, a series of studies were conducted to validate the utility and predictive validity of the instrument. Additional refinements and improvements were made between the completion of each individual study, based on feedback from district leaders and users. This served as a way to continually improve the design, clarity, and discrimination between and across indicators within the rubric.

VALIDATION STUDY SUMMARY

4. VALIDATION STUDIES

RESEARCH QUESTIONS

As a part of a larger effort to explore the relationship between the implementation of Reading Horizons at the classroom level and student learning outcomes in Grades K–3, four correlational studies were conducted during the 2021–2022 academic year. These studies provided the additional ability for Reading Horizons, in collaboration with its external research partners (Arken Research, Elite Research, LLC, and McREL International), to evaluate not only the correlation between implementation integrity and student outcomes but also to validate the use of the RHIR tool for coaching and support purposes. Each of these studies included the following research questions as part of their larger design:

1. To what extent do teachers implement RH with integrity, as measured by the Reading Horizons Implementation Integrity Rubric (RHIR)? Does the rubric effectively discriminate between various levels of classroom implementation integrity?
2. What is the relationship between teacher implementation of RH and student performance and growth, as measured by the DIBELS 8 assessment composite scores and the relevant individual subscales?

Note: The full reports for each of these studies can be found at <https://www.readinghorizons.com/research/>.

PARTICIPANTS

Reading Horizons recruited four school districts that were implementing Reading Horizons in each of their K–3 classrooms during the 2021–2022 school year. The school districts that participated in these collaborative studies, as well as the third-party evaluators who conducted the analyses for each of the studies, are shown in Table 3.

Table 3. District Participants and Third-Party Evaluators

District	Grade Levels	# of Classrooms Observed	Outcome Measure	Third Party Evaluators
Cleveland County Schools	K–3	49	DIBELS 8	Elite Research
Oklahoma City Public Schools	K–4	115	iStation	Elite Research
Opelika City Schools	K–3	80	aimsWeb	Elite Research
Tyler Independent School District	K–2	48	DIBELS 8	McREL Intl.
TOTAL		292		

DATA COLLECTION METHODS

A diverse group of RH curriculum experts, facilitators, and researchers were recruited to complete the data collection and observation process for the 292 classrooms across all four school districts. The research team at RH led a series of training sessions to train observers on the purpose and non-purpose of the data collection process and rubric, calibrate expectations to ensure consistency in observations, and identify how to pinpoint meaningful “glows and grows” for both individual educators and as campus themes.

VALIDATION STUDY SUMMARY

Site visits took place between October 2021 and April 2022. After each site visit, observers provided verbal reports to the district leaders and campus principals, as well as formal reports of district-level trends. Feedback about the usefulness of the rubric and its associated findings was gathered after each reporting period. Iterative improvements were made to the rubric following the first two site visits, reaching a saturation point of recommended enhancements. The rubric’s content was finalized by December 2022, after which the remaining site visits were conducted using the final version of the rubric.

KEY RESULTS

Research Question 1. Implementation Integrity and Discrimination

Given that three of the four districts were in their first year of implementing Reading Horizons, it was not surprising to observe significant variation in teachers’ implementation practices, as evidenced by the distribution of classroom observation scores across all four studies (see Table 4). Both collectively and within each district, the majority of teachers scored in the *Exploring* and *Engaging* levels of implementation.

Table 4. Implementation Integrity - Score Distributions

Percent of Educators at Each Progression of Practice Level (all four studies)				
Indicator	<i>Emerging</i>	<i>Exploring</i>	<i>Engaging</i>	<i>Empowering</i>
1. Effective Communication of the RH Method for Meaningful Connections	9%	25%	37%	29%
2. Instructional Routines	10%	50%	34%	6%
3. Questioning and Monitoring	14%	40%	38%	8%
4. Feedback for Deeper Understanding	22%	43%	31%	5%
5. Pacing and Lesson Structure	19%	46%	30%	5%

Inter-item reliability estimates were conducted for each of the four studies. This was to ensure that the reliability scores were substantive enough to correlate the implementation integrity variable to other outcomes without demonstrating evidence of item redundancy. Reliability estimates of 0.764, 0.854, and 0.842, respectively (see Table 5 below), were identified for three of the four studies for which an estimate was computed. These results fell within desirable thresholds, indicating that the RHIR demonstrated evidence of discrimination between educators who implemented with integrity and those who did not.

VALIDATION STUDY SUMMARY

Research Question 2. Relationship Between Implementation Integrity and Outcomes

To explore the relationship between implementation integrity and student outcomes for three of the four studies, Elite Research, LLC first applied a simple composite score analysis to classify educators into one of three groups. (Note: Due to the qualitative nature of rubric-based instruments, creating composite scores for cross-indicator comparisons is not recommended. However, mean composite scores were created to estimate scores for individual educators as a means to identify meaningful clusters of educators and run the correlational analysis.) Table 5 lists the composite scores used to create meaningful groups or cohorts of educators for correlational purposes for three of the studies.

Table 5. Item Reliability and Mean Scores Thresholds for Groups (Min = 1.0 and Max = 4.0)

District	Inter-item Reliability <i>r</i>	Thresholds by Group		
		Low, x_1	Med, x_2	High, x_3
Cleveland County Schools	0.764	1.0	2.16	3.4
Oklahoma City Public Schools	0.854	1.0	2.02	3.6
Opelika City Schools	0.842	1.83	2.8	3.83

Once appropriate groups were formed, researchers applied a multiple regression analysis to investigate the relationships between implementation integrity on student growth.

Cleveland County Schools. For the time frame established in the preliminary study (Reading Horizons, 2022a), average student growth was significantly higher for students on the DIBELS 8 assessment in classrooms where teachers scored higher on the rubric than in classrooms where teachers scored lower on the rubric:

- For kindergarten students, higher teacher classroom ratings were associated with higher composite scores ($d = .124$), as well as with *Phonemic Segmentation Fluency* ($d = .230$), *Nonsense Word Fluency* ($d = .842$), *Word Reading Fluency* ($d = .368$) scale scores.
- For first-grade students, higher teacher classroom ratings were associated with higher *Phonemic Segmentation Fluency* ($d = .295$) scale scores.
- For second-grade students, higher teacher classroom ratings were associated with higher composite scores ($d = .140$), as well as with *Nonsense Word Fluency* ($d = .258$), *Word Reading Fluency* ($d = .337$), and *Oral Reading Fluency* ($d = .177$) scale scores.

Oklahoma City Public Schools. For the time frame established in the preliminary study (Reading Horizons, 2022b), average student growth on the iStation assessment was significantly higher for students in classrooms where teachers scored higher on the rubric than in classrooms where teachers scored lower on the rubric:

- For kindergarten and third-grade students, higher teacher classroom ratings were associated with higher *Decoding* scores ($d = .944$, and $.445$, respectively)—a very large effect size for the kindergarten students.

VALIDATION STUDY SUMMARY

Opelika City Schools. For the time frame established in the preliminary study (Reading Horizons, 2022c) average student growth on the aimsweb (NCS Pearson Inc.) assessment was significantly higher for students in classrooms where teachers scored higher on the rubric than in classrooms where teachers scored lower on the rubric:

- For kindergartners, higher teacher classroom ratings were associated with *Early Literacy* scores ($d = .417$ to $.525$)—a large effect size.
- For second graders, higher teacher classroom ratings were associated with higher *Oral Reading Fluency* scores ($d = .402$ to $.714$)—a very large effect size.

SUMMARY AND IMPLICATIONS

The results from these studies demonstrate that the Reading Horizons Implementation Integrity Rubric (RHIIR) is a valuable tool for measuring and describing variation in implementation integrity among teachers, regardless of their level of experience with the curriculum. The results also indicated that higher classroom ratings on the RHIIR were associated with higher student growth across multiple assessments of early literacy skills.

Such results illustrate the potential value of more meaningful measurement of implementation integrity to support coaching and feedback, as well as their predictive power for anticipating student growth. Similar measures of implementation integrity, particularly those built on the principles of *rubric-based*, *educative*, *rigorous*, and *adaptable*, may benefit additional researchers and practitioners seeking to improve the implementation of high-quality instructional materials or predict the likelihood of student growth.

5. REFERENCES

Danks, S., & Allen, J. (2014). Performance-Based Rubrics for Measuring Organizational Strategy and Program Implementation. *Performance Improvement Quarterly*, 27(1), 33–49.

Danks, S., Naylor, C. & Burwell, R. (2021). The Reading Horizons Implementation Rubric. Kaysville, UT: Reading Horizons.

Reading Horizons (2022a). The impact of the Reading Horizons implementation on student learning: Cleveland County Schools. Retrieved from <https://www.readinghorizons.com/research/>

Reading Horizons (2022b). The impact of the Reading Horizons implementation on student learning: Oklahoma City Public Schools. Retrieved from <https://www.readinghorizons.com/research/>

Reading Horizons (2022c). The impact of the Reading Horizons implementation on student learning: Opelika City Schools. Retrieved from <https://www.readinghorizons.com/research/>

Reading Horizons (2022d). The impact of the Reading Horizons implementation on student learning: Tyler Independent School District. Retrieved from <https://www.readinghorizons.com/research/>

© 2023 ARKEN RESEARCH. All rights reserved.

