Stated Briefly Measuring the implementation fidelity of the Response to Intervention framework in Milwaukee Public Schools



Stephen J. Ruffini WestEd

Jim Lindsay American Institutes for Research Ryan Miskell WestEd

Amy Proger American Institutes for Research

In collaboration with the Midwest Urban Research Alliance

Regional Educational Laboratory Midwest assisted Milwaukee Public Schools in developing a fidelity monitoring system for measuring schools' progress in implementing Response to Intervention (RTI). This study examined the ratings produced by that system to determine the system's reliability, schools' progress in implementing RTI, and whether ratings were related to school characteristics. Two years after rolling out RTI, 53 percent of participating schools were implementing it with adequate fidelity. Average implementation ratings suggest that priority schools need more coaching and professional development on RTI than do focus schools and other schools and that most schools could benefit from coaching and professional development on tier 3 instruction and instruction for diverse students. Implementation fidelity ratings were related to the percentage of teachers with advanced credentials, retention of licensed staff, percentage of economically disadvantaged students, and percentage of students suspended during the school year.

This brief summarizes the findings of Ruffini, S. J., Lindsay, J., McInerney, M., Waite, W., & Miskell, R. (2016). *Measuring the implementation fidelity of the Response to Intervention framework in Milwaukee Public Schools*. (REL 2017–192), Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. That report is available at http://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=1460.



Institute of Education Sciences U.S. Department of Education



Why this study?

Response to Intervention (RTI; defined in box 1) has been widely adopted as a framework for meeting the instructional needs of students and as a school improvement strategy. In a 2011 survey 68 percent of U.S. public school districts indicated that they had implemented or were in the process of implementing RTI as a strategy to improve learning and achievement among all students, including high-need students struggling with basic math and reading skills (Detgen, Yamashita, Davis, & Wraight, 2011; Global Scholar et al., 2011 [as cited by Shah, 2011]; National Center on Response to Intervention, 2010).

Research suggests that RTI works best when implemented with fidelity, meaning that schools are implementing the RTI framework as intended (Gersten et al., 2008; Rolfhus et al., 2012). However, RTI is a complex framework involving multiple components, and school and district administrators often lack valid and reliable information about schools' progress implementing them (Keller-Margulis, 2012). This brief describes a system for measuring the implementation fidelity of the RTI framework and uses information generated by that system to describe the implementation fidelity of RTI in Milwaukee Public Schools.

During the 2013/14 school year a technical assistance team consisting of Regional Educational Laboratory (REL) Midwest staff and staff of the former National Center on Response to Intervention¹ worked with Milwaukee Public Schools to develop a formative feedback system on the implementation fidelity of the RTI framework. That system consists of a customized rubric for assessing whether schools were implementing the six components of RTI with fidelity and a data dashboard that collects rubric-based ratings by trained district staff, aggregates the data for the overall RTI rubric and its components, displays average ratings for components, and identifies components that need improvement (the rubric is described in box 2).

The implementation fidelity monitoring system was rolled out in 2014/15. REL Midwest staff who developed the rubric and dashboard successfully trained 22 of the district's school improvement coaches to collect implementation-related information during school visits, to rate schools' success at implementing RTI using the rubric, and to enter the ratings into the dashboard. The coaches visited 70 schools between November 2014 and June 2015. Two coaches visited each school within one week of each other.

This study examined the ratings made by school improvement coaches during school visits. Ratings for 2 schools were incomplete, leaving 68 schools in the sample. The study team analyzed the data to determine whether the school improvement coaches had been successfully trained to understand RTI components and use the RTI implementation fidelity monitoring system reliably. The study also examined how well Milwaukee schools were implementing RTI generally, and which components of RTI should be emphasized

Box 1. What is Response to Intervention?

Response to Intervention (RTI) integrates assessment and intervention within a multilevel prevention system to maximize student achievement and reduce behavioral problems. Schools identify students at risk for poor learning outcomes, monitor student progress, provide evidence-based interventions, and adjust the intensity and nature of those interventions based on a student's responsiveness (National Center on Response to Intervention, 2010, pp. 1–2).

RTI's distinctive feature is its data-based decisionmaking approach (Fuchs, Fuchs, & Compton, 2012) combined with tiers of support for students, depending on their needs. Tier 1 involves core classroom instruction for all students in the focus subjects (such as math and reading). Students who test poorly in these subjects then become eligible for supplemental small-group instruction in the subjects under tier 2 (Deno, 1985; Vaughn, Denton, & Fletcher, 2010). Tier 3 provides more individualized and intensive instruction for students who do not respond to tiers 1 and 2.

Box 2. Description of the implementation fidelity rubric of the Response to Intervention framework

The implementation fidelity rubric for monitoring the Response to Intervention framework measures the degree to which schools are implementing the overall framework and its six components as intended (referred to as implementation fidelity). The six components are data-based decisionmaking, balanced assessment, multitiered instruction, leadership, collaboration, and evaluation. Two components, balanced assessment and multitiered instruction, also had subcomponents representing processes that make up the key components. Overall, the rubric contains 33 indicators. During the school visits the school improvement coaches rated each indicator on a five-point scale, from 1, indicating little progress toward implementation, to 5, indicating full implementation. The rubric provides clear definitions for ratings of 1, 3, and 5; ratings of 2 or 4 can also be assigned if the evidence indicates that the rating should fall between two of the three defined points (the rubric and scale value descriptions are in Ruffini, Lindsay, McInerney, Waite, and Miskell, 2016). The study found the implementation fidelity monitoring system to be reliable (that is, the ratings across school improvement coaches were consistent and ratings on indicators within components were also consistent), with agreement greater than .85 and Cohen's kappa greater than .70. Interitem reliability was good for the overall rubric and for each of the components (alphas of .70 or greater).

Box 3. Methodology

Ratings were completed for 68 of the 70 schools. The dashboard system calculated the average of ratings for individual components and for the overall rubric. Partners formerly affiliated with the National Center on Response to Intervention recommended cutpoints for classifying implementation fidelity after two years: little fidelity, an average rating of less than 2.00; inadequate fidelity, average rating of 2.00–3.49; adequate fidelity, average rating of 3.50–4.99; and full fidelity, average rating of 5.00. Three types of schools were considered: priority schools (Title I schools in which overall student achievement was in the lowest 5 percent of Title I schools in the state), focus schools (Title I schools in which overall student achievement was in the lowest 10 percent of Title I schools and either subgroup performance was very low or achievement gaps between subgroups were the most significant), and other schools (schools that either were not a Title I school or were not classified as a priority or focus school). Schools' ratings and fidelity of implementation classification were used to determine which types of schools were implementing the Response to Intervention framework adequately and whether schools had more difficulty implementing particular components. The study team also calculated correlation coefficients to determine whether certain school factors were related to schools' implementation ratings. These characteristics included the percentage of highly qualified teachers, percentage of teachers with advanced degrees, percentage of students who were English learner students, and students' academic proficiency.

in professional development or coaching sessions (see box 3 for information on the methodology used for this study). Finally, correlation analyses were conducted to determine whether several school factors were related to stronger implementation fidelity of RTI.

While this report presents findings that are specific to Milwaukee Public Schools, the findings may be of interest to administrators in other districts and states. The findings illustrate how such a system can show not only how well schools are implementing RTI generally, but also which components and subcomponents of the framework should be the focus of professional development for staff. The correlation analysis can inform administrators about factors that might contribute to implementation fidelity, bearing in mind that correlation findings are suggestive at best and cannot be used to infer causal relationships.

What the study found

The findings of this study show that schools' implementation fidelity was mixed, with a little more than half of schools implementing the framework adequately. For some components of RTI, such as multitiered intervention, most schools have yet to implement the components adequately. Schools' average implementation fidelity ratings were related to several school factors.

Some 53 percent of schools visited showed adequate implementation fidelity of the Response to Intervention framework, but fidelity varied by school type

Schools' fidelity classifications indicate that 53 percent had attained a level of implementation fidelity considered adequate (figure 1). The implementation of two components was considered inadequate districtwide: multitiered instruction (69 percent of schools had average ratings for this component that fell below the cutpoint for adequate fidelity) and evaluation (49 percent of schools had average ratings below the cutpoint for adequate fidelity).

However, average ratings for the overall rubric and ratings for three of six components (balanced assessment, multitiered instruction, and evaluation) varied by type of school (priority schools, focus schools, and other schools). Schools classified as other had higher average implementation fidelity ratings than priority schools, and focus schools fell between the two (table 1). These findings suggest that schools that are struggling academically also struggle with implementing RTI.

The same pattern was evident in the percentage of schools at each level of implementation fidelity. Schools classified as other were more likely to be implementing the components and rubric adequately, compared with priority and focus schools (figure 2).

Figure 1. More than half the Milwaukee Public Schools sample showed adequate implementation fidelity overall for the Response to Intervention framework, 2014/15



Note: The sample consists of 68 schools serving students in grades K–5. The cutpoints for categories of implementation fidelity are based on those used by the National Center on Response to Intervention. No components were implemented with little fidelity. **Source:** Authors' analysis of implementation fidelity ratings made by Milwaukee Public Schools staff in 2014/15.

All schools			ls	Priority schools			Focus schools			Other schools		
Component	Range	Mean	Standard deviation	Range	Mean	Standard deviation	Range	Mean	Standard deviation	Range	Mean	Standard deviation
Data-based	2.50-			3.00-			2.50-			3.00-		
decisionmaking	5.00	4.12	0.66	5.00	3.98	0.66	5.00	4.12	0.73	5.00	4.27	0.57
Balanced	2.67–			2.67–			2.67–			2.67–		
assessment	4.67	3.75	0.48	4.50	3.61*	0.43	4.67	3.73	0.43	4.50	3.90^{*}	0.51
Multitiered	2.07-			2.33-			2.07–			2.27–		
instruction	4.87	3.32†	0.59	4.33	3.12*†	0.55	4.33	3.30†	0.56	4.87	3.53*	0.61
Leadership	2.67–			3.17–			3.17–			2.67–		
	5.00	4.05	0.54	4.83	3.91	0.48	5.00	4.06	0.54	5.00	4.17	0.59
Collaboration	2.33–			2.67–			2.33-			2.67–		
	5.00	3.81	0.57	4.33	3.70	0.50	5.00	3.78	0.63	4.67	3.95	0.55
Evaluation	2.00-			2.00-			2.00-			2.00-		
	5.00	3.25†	0.73	4.00	3.09†	0.65	4.50	3.08*†	0.70	5.00	3.59^{*}	0.73
Overall rubric	2.41-			2.88-			2.41-			2.47–		
	4.82	3.61	0.49	4.47	3.45*†	0.42	4.38	3.59	0.48	4.82	3.79*	0.51

Table 1. Average implementation fidelity ratings for the Response to Intervention framework forthe Milwaukee Public Schools sample, by school type, 2014/15

* Difference in means across school types (rows) is significant at p < .05.

+ Average rating falls below the 3.5 cutpoint for adequate fidelity set by the National Center on Response to Intervention.

Note: The sample consists of 68 schools (22 priority schools, 24 focus schools, and 22 other schools) serving students in grades K–5. Priority schools are Title I schools in which overall student achievement was in the lowest 5 percent of Title I schools in the state. Focus schools are Title I schools in which overall student achievement was in the lowest 10 percent of Title I schools and either subgroup performance was very low or achievement gaps between subgroups were the most significant. Other schools either were not Title I schools or were not classified as priority or focus schools.

Source: Authors' analysis of data collected by Milwaukee Public Schools staff in 2014/15.

Drilling deeper into multitiered instruction indicates implementation fidelity was most problematic for tier 3

To get a better understanding of aspects of multitiered instruction that schools may be struggling with, the study team examined school ratings on the four subcomponents of multitiered instruction (tier 1, tier 2, tier 3, and instruction for culturally and linguistically diverse students). This process helped identify the root of the problem, which was tier 3 instruction (figure 3). Some 63 percent of the schools visited had made little progress in implementing tier 3 of the RTI framework. Schools of all types also had difficulty implementing instruction appropriate for culturally and linguistically diverse students (see figure 3).

Schools' average implementation ratings on the overall rubric and on the six components were related to other school factors

To gain more insight into other factors that might be related to a school's ability to implement the RTI framework with fidelity, the study team calculated Pearson correlation coefficients² between several school factors and average implementation ratings. The results indicate that schools with greater percentages of teachers with advanced credentials (such as a master's degree or National Board Certification) and schools with higher teacher retention rates tended to have higher average implementation ratings (table 2). Schools serving higher percentages of disadvantaged students (as proxied by eligibility for the federal school lunch program) had lower implementation ratings. Indicators of schools with challenging student populations, such as low proficiency rates in math and reading and high percentages of students suspended for disciplinary reasons, were negatively related to implementation fidelity.

Figure 2. Adequacy of implementation fidelity for the Response to Intervention framework in the Milwaukee Public Schools sample was lowest among priority schools, followed by focus schools and other schools, 2014/15



Note: The sample consists of 68 schools (22 priority schools, 24 focus schools, and 22 other schools) serving students in grades K–5. Priority schools are Title I schools in which overall student achievement was in the lowest 5 percent of Title I schools in the state. Focus schools are Title I schools in which overall student achievement was in the lowest 10 percent of Title I schools and either subgroup performance was very low or achievement gaps between subgroups were the most significant. Other schools either were not Title I schools or were not classified as priority or focus schools. No components were implemented with little fidelity.

Source: Authors' analysis of ratings made by Milwaukee Public Schools staff in 2014/15.

Figure 3. Most schools in the Milwaukee Public Schools sample had not implemented tier 3 instruction or appropriate instruction for culturally and linguistically diverse students with adequate fidelity, 2014/15



Note: Tier 1 involves core classroom instruction for all students, tier 2 provides supplemental instruction for students who perform poorly on subject-matter screening assessments, and tier 3 involves more individualized and intensive instruction for students who do not respond to tiers 1 and 2. The sample consists of 68 schools (22 priority schools, 24 focus schools, and 22 other schools) serving students in grades K-5. Priority schools are Title I schools in which overall student achievement was in the lowest 5 percent of Title I schools in the state. Focus schools are Title I schools in which overall student achievement was in the lowest 10 percent of Title I schools and either subgroup performance was very low or achievement gaps between subgroups were the most significant. Other schools either were not Title I schools or were not classified as priority or focus schools.

Source: Authors' analyses of ratings made by Milwaukee Public Schools staff.

Implications for Milwaukee Public Schools and other districts

The RTI fidelity monitoring system was designed for formative purposes: to give stakeholders in Milwaukee Public Schools trustworthy data that they can use to improve the implementation of the complex framework.

Specifically, the data suggest that most schools face challenges when implementing two subcomponents of multitiered instruction (tier 3 instruction and instruction appropriate for culturally and linguistically diverse students) and evaluation. Thus schools could use additional professional development and coaching on those parts of the RTI framework.

Though this study offers proof of concept for the system of monitoring the implementation fidelity of the RTI framework and shows how the system can be used to identify parts of the framework that most need improvement, administrators in other districts will likely want to know more about the costs to run the system. First, other districts will incur no costs associated with developing the rubric, which is appended to the full report for this study (Ruffini et al., 2016). The National Center on Response to Intervention's key components integrity rubric—on which the district's rubric was based—is also publicly available (National Center on Response to Intervention, 2010). However, districts that want to use this implementation fidelity monitoring system will incur costs associated with training staff to use the rubric for rating implementation (the school improvement coaches whose ratings were analyzed for this study received three days of training), assigning staff (school improvement coaches or trained staff holding other roles in the district) to visit Table 2. Correlations between school factors and implementation fidelity ratings for the response to intervention framework for Milwaukee Public Schools, 2014/15 (Pearson product-moment correlation coefficients)

	Response to Intervention component								
School factor	Data based decision making	Balanced assessment	Multitiered instruction	Leadership	Collaboration	Evaluation	Overall score		
Teacher characteristics									
Percentage of teachers with an advanced credential ^a	.28*	.22	.39***	.24	.25*	.20	.36**		
Percentage of teachers with five or more years of teaching experience	.04	.20	.10	04	.06	04	.09		
Percentage of teachers meeting federal highly qualified teacher requirements	05	.07	.12	.07	.07	08	.09		
Student–licensed staff ratio	02	.10	.11	.09	.10	.18	.12		
Teacher retention	.31*	.40***	.29*	.27*	.13	.11	.33**		
Student characteristics									
Percentage of students who are English learner students	.17	.23	.15	03	10	.02	.12		
Percentage of students with a disability	08	21	13	14	15	15	17		
Percentage of students eligible for the federal school lunch program	11	28*	23	26*	21	26**	28*		
Student enrollment	04	.02	.03	06	01	04	03		
Other factors									
Percentage of students proficient in math	.21	.35**	.33**	.31**	.23	.36**	.37**		
Percentage of students proficient in reading	.11	.29*	.20	.25*	.21	.23	.26*		
Percentage of students suspended	23	27*	25*	25*	27*	18	32*		

* Significant at p < .05; ** significant at p < .01; *** significant at p < .001.

Note: The sample consisted of 68 schools serving students in grades K–5. Positive coefficients indicate positive relationships (higher values on school or teacher characteristics have higher implementation fidelity ratings), and negative coefficients indicate negative relationships (higher values on school or teacher characteristics have lower implementation fidelity ratings). More consistent relationships have larger coefficients. Benjamini–Hochberg corrections for multiple comparisons did not affect determinations of statistical significance of the correlations (Benjamini & Hochberg, 1995).

a. Advanced credentials are advanced academic degrees (for example master's degree or higher) or credentials (for example, National Board Certification).

Source: Authors' analysis of ratings made by Milwaukee Public Schools staff in 2014/15 and school data from the Wisconsin Department of Public Instruction website.

schools and rate their implementation, and checking the accuracy of school improvement coaches (or other trained staff who visit schools to make ratings using the rubric) twice a year. Helping schools and districts understand how to use the data for formative purposes will also incur some costs.

Limitations of the study

The primary limitation of the study is that it was conducted in just one district. As this was the first study to use the rubric and data dashboard, there is no evidence that this system can produce trustworthy data about implementation fidelity of RTI in other settings. However, this study was conceived during

discussions with members of the Midwest Urban Research Alliance, whose school districts are all implementing RTI and which may therefore be interested in adopting this system or a customized version of it. Should one or more districts choose to do so, additional studies can be performed to provide more information on the system's reliability and average implementation ratings in those sites (reliability information for the Milwaukee RTI implementation fidelity monitoring system is given in Ruffini et al., 2016).

Another limitation is that the study findings are limited to aggregated ratings from the system. No qualitative information was gathered from potential system users, nor did the study team explore the extent of the system's use among district staff.

Also, the correlations between school factors and implementation fidelity ratings provide no information about the direction of causality between variables. Future research can investigate the direction of causality using an appropriate research design.

Finally, this study was not designed to determine whether fidelity of implementation of the RTI framework could improve student achievement test scores. Future studies can investigate the impact of implementation fidelity on student outcomes.

Notes

- 1. The National Center on Response to Intervention was a technical assistance center that supported states and school districts' efforts at establishing RTI. It ran from 2007 to 2012 through a grant from the U.S. Department of Education's Office of Special Education Programs.
- 2. For Pearson correlation coefficients, positive coefficients indicate positive relationships (higher values on school or teacher characteristics have higher implementation ratings), and negative coefficients indicate negative relationships (higher values on school or teacher characteristics have lower implementation fidelity ratings). More consistent relationships have larger coefficients.

References

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 289–300.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219–232.
- Detgen, A., Yamashita, M., Davis, B., & Wraight, S. (2011). State policies and procedures on response to intervention in the Midwest Region. (REL 2011–116). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. http://eric.ed.gov/?id=ED520326
- Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RTI: A next-generation approach to multilevel prevention. Exceptional Children, 78(3), 263–279. http://eric.ed.gov/?id=EJ970680
- Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., et al. (2008). Assisting students struggling with reading: Response to intervention and multi-tier intervention for reading in the primary grades: A practice guide (NCEE No. 2009–4045). National Center for Education Evaluation and Regional Assistance Working Paper. Washington, DC: U.S. Department of Education, Institute of Education Sciences. http://eric.ed.gov/?id=ED504264
- Keller-Margulis, M. A. (2012). Fidelity of implementation framework: A critical need for response to intervention models. *Psychology in the Schools*, 49(4), 342–352. http://eric.ed.gov/?id=EJ989963
- National Center on Response to Intervention. (2010). Essential components of RTI—A closer look at Response to Intervention. Washington, DC: U.S. Department of Education, Office of Special Education Programs, and the National Center on Response to Intervention. Retrieved August 20, 2014, from http://www.RTI4success.org/sites/default/files/RTIessentialcomponents_042710.pdf.
- Rolfhus, E., Gersten, R., Clarke, B., Decker, L., Wilkins, C., & Dimino, J. (2012). An evaluation of Number Rockets: A Tier-2 intervention for grade 1 students at risk for difficulties in mathematics. Final report. Washington, DC: U.S. Department of Education, Institute of Education Sciences, Regional Educational Laboratory Southwest. http://eric.ed.gov/?id=ED529429
- Ruffini, S., Lindsay, J., McInerney, M., Waite, W., & Miskell, R. (2016). Measuring the implementation fidelity of the response to intervention framework in Milwaukee Public Schools (REL 2016–192), Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies. ed.gov/ncee/edlabs
- Shah, N. (2011, August 19). Survey of school, district workers shows wider use of RTI. Education Week: Spotlight on Special Education. Retrieved on July 15, 2016, from http://blogs.edweek.org/edweek/speced/2011/08/yet_another_study_shows_the.html
- Vaughn, S., Denton, C., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading disabilities. *Psychology in the Schools*, 47(5), 432–444. http://eric.ed.gov/?id=EJ896142

REL 2017-193

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

November 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0004 by Regional Educational Laboratory Midwest administered by American Institutes for Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Ruffini, S. J., Lindsay, J., Miskell, R., & Proger, A. (2016). Stated Briefly: Measuring the implementation fidelity of the Response to Intervention framework in Milwaukee Public Schools (REL 2017–193). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

The Regional Educational Laboratory Program produces 7 types of reports

