The Impact of Multi-Tiered Systems of Support (MTSS) on Student Attendance and Behavior

Nicholas Gage
Kim Salomonson
Tori Ballew
Beth Clavenna-Deane
Nicolette Grasley-Boy

May 2024

Gage, N. A., Salomonson, K., Ballew, T., Clavenna-Deane, B., & Grasley-Boy, N. (2023). *The impact of multi-tiered systems of support (MTSS) on student attendance and behavior*. WestEd.

Table of Contents

LIST OF TABLES

Introduction

The success of all students in schools, including students with learning differences, is contingent on the ways schools operationalize universally designed instruction, positive behavior support, and data-based decision-making for individualized and group-level interventions. When schools have fully functional multi-tiered systems of support (MTSS) that center equity and pay necessary attention to academic conditions, student behavior, and social–emotional learning conditions, students are more likely to receive the instruction and interventions they need. Further, there is evidence of improvements in student outcomes, particularly those related to behavior, when MTSS is implemented effectively (Condliffe et al., 2022; Lee & Gage, 2020). However, in too many schools, systems of support are either not in place or not designed to effectively meet students' needs.

To address the need for effective MTSS implementation, WestEd engaged in a research and technical assistance partnership wherein WestEd subject matter experts provided guidance on designing and implementing effective MTSS; WestEd researchers simultaneously studied the short-, mid-, and long-term effects of these efforts on student outcomes. Starting in spring 2022, WestEd engaged directly with three schools from one district. At each site, WestEd provided coaching for MTSS design and implementation and collected baseline MTSS fidelity of implementation data and student-level data from the school district.

The goal of this study was to identify the ways in which district and school systems of general supervision and support for evidence-based practices can be strengthened to inform policies and procedures that support improvement efforts at scale. This interim report provides initial analyses on student-level impacts in one of the two school districts. Using a quasi-experimental design (QED), the study team compared data on students in three schools using the WestEd MTSS coaching model with data on students in schools receiving business-as-usual instruction. Final data collection and analyses will occur in summer 2024.

Methods

The evaluation team used a quantitative approach to evaluate WestEd's MTSS coaching model and the concomitant effects on student, school, and district outcomes. School and district recruitment and on-site coaching and support began in late spring 2022 and will continue through the 2023/24 school year. Specifically, the research questions were as follows:

1. Does the WestEd MTSS coaching model have a significant effect on behavior incidents for students in treatment schools compared to students in schools not receiving MTSS instruction/coaching?
2. Does the WestEd MTSS coaching model have a significant effect on the likelihood of having at least one behavior incident for students in treatment schools compared to students in schools not receiving MTSS instruction/coaching?
3. Does the WestEd MTSS coaching model have a significant effect on attendance for students in treatment schools compared to students in schools not receiving MTSS instruction/coaching?

Sample

WestEd partnered with Rock Island Milan School District (Rock Island). Rock Island Milan School District is in Western Illinois on the border of Iowa and includes nine elementary schools. Student-level data for all elementary students in the district were collected and included demographic information, academic achievement, behavior incidents, and attendance from the 2019/20 school year through the 2022/23 school year. Professional development and coaching began in fall 2022 and will continue through the end of the 2024 school year. It is important to note that not all data were available for each school year due to the COVID-19 pandemic. Specifically, for the 2020/21 school year, only beginning-of-the-year Northwest Evaluation Association Measures of Academic Progress (NWEA MAP) scores were collected, and no discipline data were recorded.

Measures

Behavior Incidents

The school district recorded each behavior incident referred to administration for disciplinary action. Behavior incidents are, essentially, office discipline referrals.

Northwest Evaluation Association Measures of Academic Progress

We used data from the 2020 beginning-of-year NWEA MAP English Language Arts (ELA) computer-adaptive subtest for both student matching and outcome model covariates. The K–2 Reading test contains 43 questions and takes students 15–30 minutes to complete, and it has the option to read aloud questions on the computer to increase accessibility. The grades 2–5 and 6–12 Reading tests contain 40–43 questions (45–60 minutes to complete). The K–2 Reading test focuses on phonological awareness, capitalization, punctuation, spelling, grammar, informational text, and vocabulary. The grades 2–5 and 6–12 tests focus on literary texts, summarization, conclusion drawing, informational texts, vocabulary, and context clues.

Attendance

The school district reported the number of absences from school each student experienced across the study period.

Data Analyses

Study Design

We used a QED approach to evaluate the efficacy of the WestEd MTSS model. This study is aligned with the Every Student Succeeds Act (ESSA) guidelines for Tier 2 Moderate Evidence (ESSA, 2015). Specifically, ESSA Tier 2 requires a study to (a) meet What Works Clearinghouse (WWC) Standards with Reservations, (b) indicate a statistically significant positive effect on a relevant outcome, (c) include at least 350 student participants, and (d) be conducted in at least two educational sites. The comparison student group, who received business-as-usual instruction, was created using propensity score matching of students in schools that did not receive WestEd MTSS support until a later school year, as outlined below.

Propensity Score Matching

We used propensity score matching (PSM) to identify a baseline-equivalent comparison group of students in schools that did not receive WestEd support until 2022/23. PSM is a statistical matching procedure that, in essence, identifies a one-to-one match for all students in the treatment schools to those in the comparison schools. We estimated propensity scores using

logistic regression following procedures outlined by Leite (2017). Specifically, treatment (WestEd MTSS) was coded as a dichotomous indicator and used as the dependent variable in a logistic regression model with all available student-level demographics, including gender, race/ethnicity, low-income status, disability status, English Learner (EL) status, students' fall 2020 NWEA MAP ELA scores, and 2021 discipline incidents. The propensity score is the predicted probability of a student being assigned to the treatment or comparison group based on the model covariates, and thus it reduces selection bias by establishing equivalence on the included model covariates.

Next, we used each treatment group student's estimated propensity score to match them with a control comparison group student using the one-to-one optimal matching method (Rosenbaum, 1989). This method minimizes global propensity score distance (i.e., predicted probability of being in the treatment or comparison group) to a student in a comparison school by finding the smallest average absolute distance across all the matched students. We used the one-to-one optimal matching algorithm using the MatchIt (Ho et al., 2011) and Optmatch (Hansen et al., 2016) packages in R (R Core Team, 2021). To confirm covariate equivalence, we calculated standardized mean difference effect sizes (*g*), where equivalence is defined as $g < .25$ standard deviations (WWC, 2020).

Multilevel Modeling

Next, we used multilevel modeling to estimate treatment effects. This modeling approach ensures accurate analysis of student-level discipline incidents and attendance, accounting for the nesting of students in classrooms and schools. We used linear models except for the likelihood a student received discipline, which was estimated using a multilevel logistic regression model. All multilevel models were estimated using lme4 (Bates et al., 2014) in R.

Results

We used PSM to identify a baseline equivalent comparison group at the student-level. Table 1 provides the descriptive statistics for the 653 students in the three treatment schools and the 653 PSM students. We also calculated standard mean difference effect sizes (*g*) for each baseline measure to confirm equivalence. The students were equivalent ($g < 0.25$) on all variables. Two important considerations should be noted. First, we used students' fall 2020 (beginning-of-the-year [BOY]) NWEA MAP ELA scores to ensure students were the same with regards to prior year ELA achievement. Thus, all kindergarten students who enrolled in the schools in fall 2021 were not included nor were the two kindergarten students in the sample (one in each treatment group) who were held back and who repeated kindergarten. Second, no behavior incidents were collected during the 2020/21 school year because of the pandemic. Therefore, we used students' 2021/22 behavior incident data to establish equivalence on behavior. This means that behavior incidents in 2021/22 could not be an outcome measure in that year, but we could use them the following year.

Table 1. Descriptive Statistics for Treatment and Propensity Score Matching Comparison Students

Baseline characteristic	PSM comparison	Treatment	Equivalence (*g*)
Grade			
K	0.2%	0.2%	0.00
1	16.5%	17.0%	0.05
2	16.1%	15.2%	0.04
3	13.8%	13.5%	-0.01
4	14.9%	15.9%	0.04
5	19.0%	18.2%	-0.03
6	19.6%	20.1%	0.02
Gender			
Female	49.3%	47.6%	0.04
Male	50.7%	52.4%	

Baseline characteristic	PSM comparison	Treatment	Equivalence (*g*)
Ethnicity			
Hispanic/Latino	14.2%	18.1%	0.16
Not Hispanic/Latino	85.8%	81.9%	
Race			
Amer Indian	0.0%	0.3%	0.00
Asian	5.5%	5.8%	0.03
Black	29.7%	38.6%	0.22
Multiracial	10.0%	8.0%	0.14
Pacific Isl	0.0%	0.0%	0.00
White	54.8%	47.3%	-0.17
Low-Income			
No	29.9%	26.3%	-0.10
Yes	70.1%	73.7%	
Disability			
No	84.4%	83.6%	-0.03
Yes	15.6%	16.4%	
English Learner			
No	89.7%	85.0%	-0.24
Yes	10.3%	15.0%	
2020 BOY MAP ELA			
M	175	174	-0.04
SD	25.7	25.3	
2021 Behavior Incidents			
M	0.161	0.211	0.07
SD	0.851	0.628	

***Note*. BOY MAP ELA is the beginning-of-the-year NWEA MAP English Language Arts Test.**

Behavior Incidents

We used behavior incidents from 2021/22 to establish baseline equivalence and, therefore, could not estimate treatment effects for that same year. Instead, we focused on the 2-year implementation effects from the end of the 2022/23 school year. Although we used PSM to establish equivalence, we included the covariates given their relation to the dependent variable and included the school as a fixed effect because of the small number of schools. The models are presented in Table 2.

We found a statistically significant treatment effect on the number of behavior incidents. Specifically, there were fewer behavior incidents in treatment schools after 2 years of

implementation when we controlled for the number of prior year behavior incidents. The covariate adjusted standardized mean difference effect size is *g* = -0.35.

Table 2. Multilevel Models Estimating Treatment Effects on Behavior Incidents at the End of 2022/23 School Year

Parameter	Estimate	Std. error	p-value
Intercept	0.14	0.79	0.862
Treatment	-0.35*	0.14	0.014

***Note*. *p < .05, , 942 students for spring 2023, 111 homerooms for spring 2023. We included all student characteristics and schools as a fixed effect.**

Next, we extended the model above by estimating the number of unique students with at least one behavior incident instead of the number of behavior incidents. Put differently, this model estimates the likelihood that a student in a treatment school had at least one behavior incident compared to a student in a comparison school when we controlled for the number of prior year incidents. The dependent variable was dichotomous (0 for no behavior incident and 1 for one or more behavior incidents); therefore, we used a multilevel logistic regression model. The models are presented in Table 3.

Again, **we found that there was a significant treatment effect, which suggests that students in treatment schools were significantly less likely to have a behavior incident** when we controlled for prior year incidents. The effect size was odd ratio (OR) = 0.32, or, when converted to standardized mean difference, the effect size is *g* = -0.63.

Table 3. Multilevel Logistic Regression Model Estimating Treatment Effects on the Likelihood of a Behavior Incident by the End of 2022/23 School Year

Parameter	Log(OR)	Std. error	p-value
Intercept	-16.18	>1.00	0.995
Treatment	-1.15**	0.39	0.004

***Note*. **p < .001, 942 students for spring 2023, 111 homerooms for spring 2023. We included all student characteristics and schools as a fixed effect.**

Attendance

We also examined the impact of MTSS coaching on days absent from school at the student level. We focused on only those students who were in schools prior to the pandemic and used PSM to establish equivalence for days absent. At baseline (2019/20 school year), the mean number of days absent was 12.6 (SD = 14.6) in PSM comparison schools and 14.8 (SD = 17.3) in treatment schools. Equivalence was *g* = 0.138, which is below the 0.25 threshold described

above. We then estimated multilevel models for each year after the treatment began and included all demographics and the baseline days absent in the models. The results are presented in Table 4.

We found a statistically significant treatment effect for both implementation years, suggesting that students in treatment schools had significantly fewer days absent than PSM students in comparison schools. The effect sizes were $g = 0.86$ for the 2021/22 school year and $g = 0.91$ for the 2022/23 school year.

Table 4. Multilevel Models Estimating Treatment Effects on Days Absent Using 2019/20 for Equivalence

Parameter	2021/22			2022/23		
	Estimate	Std. error	p-value	Estimate	Std. error	p-value
Intercept	66.10**	20.03	0.001	65.24**	20.20	0.001
Treatment	-12.68***	2.65	0.000	-13.19***	3.53	0.000

***Note*. **$p < .001$, ***$p < .000$. We included all student characteristics and schools as a fixed effect.**

Discussion of Results

This section of the report provides a discussion and contextualization of the research results.

Significant Reductions in Behavior Incidents

The results in Rock Island suggest that students in schools receiving WestEd's MTSS coaching have statistically significantly fewer behavior incidents and that students in those schools are less likely to have a behavior incident. Prior to schools receiving MTSS coaching, the results were consistent when students were matched prior to the pandemic and after the pandemic on demographics and prior behavior incidents across schools. Thus, the data suggest that WestEd MTSS-coached teachers and other school staff did a much better job of reducing problem behaviors for all students (universal prevention) and reduced the total number of behavior incidents, likely due to more intensive interventions (tiers 2 and 3).

There is evidence that suggests that schools that implement SWPBIS with fidelity have fewer ODR and suspensions (see Lee & Gage, 2020). Yet, most prior studies used school-level data.

Finding significant impacts using PSM at the student level is novel and important for several reasons. First, the WestEd coaching model does not focus exclusively on SWPBIS; instead, it focuses on coaching schools on both behavior and academic MTSS practices. This approach replicates the positive results of two similar MTSS studies that also combined behavior and academic MTSS (Algozzine et al., 2019; Scott et al., 2019). However, this study is more robust in design than those two studies because the data are at the student level and prior behavior incidents are controlled for. The current study meets WWC Standards with Reservations and finds significant reductions in behavior incidents. Yet, until we collect post-fidelity data, we cannot say for certain that it was likely an increase in MTSS practices that produced the reduction in behavior incidents. We will explore the causal links between MTSS implementation using spring 2024 fidelity data in the final report.

Significant Reductions in Student Absences

Chronic absenteeism has become a significant and vexing challenge for schools since returning from the pandemic (Blad, 2023). Thus, any program or intervention that can reduce the

likelihood of a student missing school is important. The results of our analyses reveal that students in schools participating in WestEd's MTSS coaching have fewer absences than do students in other schools. This finding is important because few studies have found a relationship between MTSS implementation and reductions in chronic absenteeism. The finding suggests that MTSS practices appear to either increase students' interest in school, parents' connection to school, or implementation of policies that increase attendance. Unfortunately, we do not have any data on how or why schools receiving MTSS coaching had fewer absences. Similar to behavior incidents, we need the post-fidelity data to determine whether schools in Rock Island increased their MTSS implementation overall and which subscales or items in particular increased. This will be a matter of focus of our analyses in the final report.

Conclusion

This report provides details about initial student-level impacts from one partner school district. As noted, we used a quantitative approach to evaluate WestEd's MTSS coaching model and the concomitant effects on student outcomes (i.e., behavior incidents, likelihood of a behavior incident, and attendance) and found significantly better outcomes for students in schools that received MTSS coaching. To reiterate, this is an interim report and does not reflect final impacts.

References

Algozzine, B., Barrett, S., Eber, L., George, H., Horner, R., Lewis, T., Putnam, B., Swain-Bradway, J., McIntosh, K., & Sugai, G. (2019). *School-wide PBIS tiered fidelity inventory*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. www.pbis.org

Algozzine, B., Wang, C., White, R., Cooke, N., Marr, M. B., Algozzine, K., Helf, S. S., & Duran, G. Z. (2012). Effects of multi-tier academic and behavior instruction on difficult-to-teach students. *Exceptional Children*, *79*(1), 45–64. https://doi.org/10.1177/001440291207900103

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4* (R package version 1.1–7) [Computer software].

Blad, E. (2023, October 12). *High absenteeism hits more schools, affecting students with strong attendance, too*. Education Week. https://www.edweek.org/leadership/high-absenteeism-hits-more-schools-affecting-students-with-strong-attendance-too/2023/10

Condliffe, B., Zhu, P., Doolittle, F., van Dok, M., Power, H., Denison, D., & Kurki, A. (2022). *Study of training in multi-tiered systems of support for behavior: Impacts on elementary school students' outcomes* (NCEE 2022-008). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Every Student Succeeds Act, S. 1177, 114th Cong., (2015). https://www.gpo.gov/fdsys/pkg/BILLS-114s1177enr/html/BILLS-114s1177enr.htm

Hansen, B. B., Fredrickson, M., Fredrickson, M. M. M., Rcpp, L., & Rcpp, I. (2016). *optmatch (R package version 0.9-6).* [Computer software].

Ho, D., Imai, K., King, G., & Stuart, E. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1–28. http://www.jstatsoft.org/v42/i08/

Lee, A., & Gage, N. A. (2020). Updating and expanding systematic reviews and meta-analyses on the effects of school-wide positive behavior interventions and supports. *Psychology in the Schools*, *57*(5), 783–804.

Leite, W. (2017). *Propensity score methods using R*. SAGE.

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Rosenbaum, P. R. (1989). Optimal matching for observational studies. *Journal of the American Statistical Association*, *84*(408), 1024–1032.

Scott, T. M., Gage, N. A., Hirn, R. G., Lingo, A. S., & Burt, J. (2019). An examination of the association between MTSS implementation fidelity measures and student outcomes. *Preventing School Failure: Alternative Education for Children and Youth*, *63*(4), 308–316.

What Works Clearinghouse. (2020). *Web-based Intelligent Tutoring for the Structure Strategy (ITSS)*. Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/WWC/Docs/InterventionReports/wwc_ALitss_IR_apr2020.pdf