

The Associated Effects of McGraw-Hill Education's Wonders Program on Measures of Achievement and Teacher Attitudes

Introduction

This study investigated the impact of McGraw-Hill Education's *Wonders* reading program on measures of reading achievement and teacher attitudes. Participants included 299 students in grades K-5 attending one elementary school in Wisconsin. *Wonders* was implemented at the start of the 2018-2019 school year and throughout the 2019-2020 school year, including a period of remote learning during the spring of 2020. Throughout the year, students were administered periodic benchmark assessments using the *Renaissance* STAR reading assessment. Fourth and fifth grade students additionally participated in Wisconsin's state-mandated test in English Language Arts, the *Wisconsin Forward*. At the end of the study, teachers completed an online survey to gauge perceptions of *Wonders*, including feasibility of use during a period of remote instruction.

Method

Site of Study

The site selected for this study included classrooms within one elementary school building located in a Wisconsin school district. In 2018, the school housed 469 students spanning Pre-Kindergarten through fifth grade. The majority of students were Caucasian (95%). School-wide, about 41% of students received free or reduced price meals, 20% received special education services, and 2% were considered English Language Learners.

Participants

Selected classrooms within the elementary school implemented *Wonders* during the 2018-2019 and 2019-2020 school years. Participants included 299 students spanning Kindergarten through fifth grade. As shown in Table 1, students largely identified as Caucasian (about 96%) and native English speakers (98%). About 46% of students received free/reduced lunch services and about 26% received special needs services.

Table 1

Student Demographics as a Percentage of the Sample

Demographic	(n)	(%)
Grade		
KG	46	15.4
1	59	19.7
2	50	16.7
3	41	13.7
4	56	18.7
5	47	15.7
Gender		
Female	153	51.2
Male	146	48.8
Ethnicity		
African American	4	1.3
Caucasian	288	96.3
Hispanic	7	2.3
ESL		
No	294	98.3
Yes	5	1.7
Free/Reduced Lunch		
No	162	54.2
Yes	137	45.8
Special Education		
No	222	74.2
Yes	77	25.8

Twenty-six teachers participated in the study and 25 teachers provided demographic information. Five teachers each instructed students in Kindergarten, second grade, and third grade; four teachers instructed students in first grade; and two teachers each instructed students in the fourth and fifth grades. Two teachers were Title I Reading teachers. The range in years of teaching experience was 1 to 27 years, with a mean of about 12 years. About half ($n = 12$) of teachers had Masters degrees while the remaining teachers ($n = 13$) had Bachelors degrees.

Measures

Achievement measures included fall (Beginning of Year, or *BOY*), winter (Middle of Year, or *MOY*) and spring (End of Year, or *EOY*) benchmark assessments on the Renaissance Learning *STAR Early Literacy* and *STAR Reading* assessments, as well as student scores on Wisconsin's state-mandated achievement test in English Language Arts, the Wisconsin *Forward*.

STAR Early Literacy and STAR Reading

STAR Reading assessments are computer-adaptive assessments, where the difficulty of test items is automatically adjusted based on the skill level of the student. *STAR Early Literacy* assesses proficiency with early reading skills, while *STAR Reading* assesses reading comprehension (Renaissance Learning, 2016). Internal consistency reliability estimates above .70 are considered at minimum, acceptable; reliability estimates for grades Kindergarten through second grade ranged from .79 to .84 (U.S. Department of Education, 2016). The district provided individual student percentile rank scores on the *STAR Early Literacy* and *STAR Reading* assessments. A percentile rank compares a student's performance relative to peers; that is, it indicates the percentage of individuals scoring lower than the student. For example, a student scoring at the 65th percentile scored higher than 65% of other students in the reference or comparison group. *Renaissance Learning* recommends using student percentile ranks as a benchmark measure to distinguish students likely to meet performance standards from students in need of intervention. On *STAR Early Literacy* and *STAR Reading* assessments, a percentile rank of 40 is a 'cut' score that serves as a proficiency benchmark. Students who earn a percentile rank of 40 or higher are likely to meet end-of-year performance standards in reading. Students scoring slightly below benchmark and within the 39th to 26th percentile are considered 'On Watch'. Students scoring below the 25th percentile but above the 10th percentile are considered in need of intervention. Students scoring below the 10th percentile are in need of 'urgent' intervention (Renaissance Learning, Inc., 2018 & 2020).

Students were administered *STAR* benchmark assessments in the fall at the beginning of the year (*BOY*), in the winter at the middle of the year (*MOY*), and in the spring at the end of the year (*EOY*). The district provided percentile ranks for participating students for the 2017-2018 school year and before implementation of *Wonders*; for the 2018-2019 school year and during the first year of *Wonders*; and for the 2019-2020 school year and during the second year of *Wonders*. Students were not administered *STAR* assessments in the spring of 2020 during the COVID-19 pandemic.

Wisconsin Forward Exam

The Wisconsin *Forward* Exam is a state-mandated assessment administered in the spring of each school year and designed to gauge how students perform in relation to expected academic standards in English Language Arts. Students in the third grade and above participate in the Wisconsin *Forward*. Scores are reported in a three-digit, scaled score format and are translated into performance categories: *Below Basic*, *Basic*, *Proficient*, and *Advanced*. A student scoring within the *Below Basic* category means the student demonstrates *minimal* understanding of and ability to apply grade level skills in English Language Arts. A student scoring within the *Basic* category means the student demonstrates *partial* understanding of and ability to apply grade level skills in English Language Arts. A student scoring within the *Proficient* category indicates *adequate* understanding and application of skills in English Language Arts, and a student scoring within the *Advanced* category demonstrates *thorough* understanding and application of skills in English Language Arts. The district provided individual student scaled scores, which were then translated into corresponding percentile ranks and proficiency categories based on the state of Wisconsin Department of Education guidelines. The exam was not administered in the spring of 2020, given the COVID-19 pandemic and resulting school closure.

For clarity, Table 2 provides a review of available scores by grade level and year. Here, *Grade Level* reflects student grade level as assigned during the 2019-2020 school year. For example, students identified as fifth graders during the 2019-2020 school year were in the fourth grade during the 2018-2019 school year, and in the third grade during the 2017-2018 school year. It should be noted that some cohorts of students (e.g., Grade 2 and Grade 3) were administered different assessments (*STAR Early Literacy* and *STAR Reading*) in subsequent years. Because *STAR Early Literacy* and *STAR Reading* emphasize different reading components, direct comparisons across assessments will not be made for some groups of students.

Table 2

Available Measures by Grade Level and Year

Grade Level in 2019-2020	2017-2018 (Before)	2018-2019 (Year 1)	2019-2020 (Year 2)
Kindergarten	--	--	<i>STAR Early Literacy</i>

Grade 1	--	<i>STAR Early Literacy</i>	<i>STAR Early Literacy</i>
Grade 2	<i>STAR Early Literacy</i>	<i>STAR Reading</i>	<i>STAR Reading</i>
Grade 3	<i>STAR Early Literacy</i>	<i>STAR Reading</i>	<i>STAR Reading</i>
Grade 4	<i>STAR Reading</i>	<i>STAR Reading</i> <i>Wisconsin Forward</i>	<i>STAR Reading</i>
Grade 5	<i>STAR Reading</i> <i>Wisconsin Forward</i>	<i>STAR Reading</i> <i>Wisconsin Forward</i>	<i>STAR Reading</i>

Results

STAR Reading and STAR Early Literacy

The following section presents student results on the *STAR Reading* and *STAR Early Literacy* assessments, by grade level. For all grade levels, *BOY* reflects the beginning of year administration in the fall, *MOY* reflects the middle of year administration in the winter, and *EOY* reflects the end of year administration in the spring. Because percentile ranks are not on an equal interval scale in which the distance between subsequent values is the same and therefore cannot be averaged, *median* percentile ranks were examined in this analysis. For tests of significance, the Wilcoxon Signed-Ranks Test was used to determine differences in the median ranks between two administrations.

Kindergarten

The Kindergarten cohort was administered *STAR Early Literacy* in the fall, at the beginning of the 2019-2020 school year (19BOY) and immediately before *Wonders* instruction. The Kindergarten cohort was again administered *STAR Early Literacy* in the winter and during the middle of the 2019-2020 school year (20MOY), after several months of *Wonders* instruction.

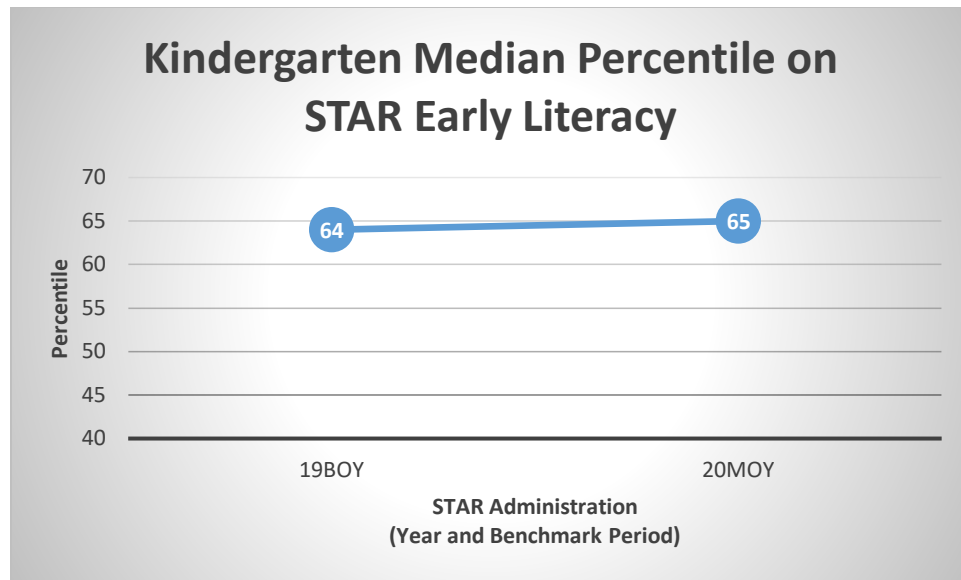


Figure 1. *Median Percentile on STAR Early Literacy, Kindergarten*

Figure 1 presents the median percentile score on *STAR Early Literacy* before *Wonders* (19BOY) and after several months of *Wonders* (20MOY). Although the difference in ranking between these two administrations was not statistically significant (Wilcoxon signed-ranks test $p = .364$), about 52% ($n = 23$) of students exhibited higher percentile scores after participating in several months of *Wonders*.

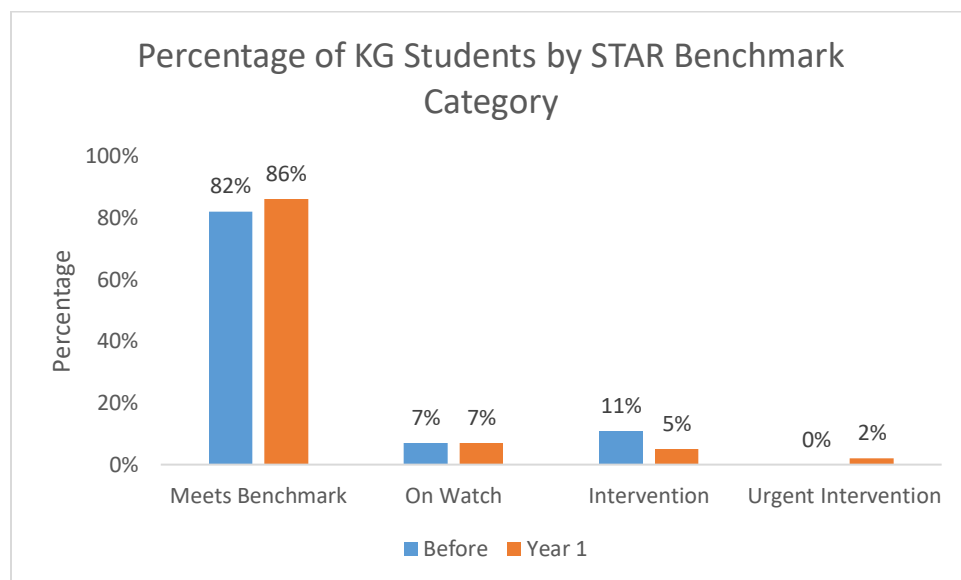


Figure 2. *Percentage of Kindergarten Students by Benchmark Category on STAR*

Figure 2 shows the percentage of Kindergarten students scoring within each benchmark category on the *STAR* assessment before exposure to *Wonders* and after several months exposure to *Wonders*. Before exposure to *Wonders*, about 82% of students were found to meet benchmark standards on *STAR* and about 18% were determined to score below benchmark standards (either *On Watch*, *Intervention*, or *Urgent Intervention*). Following several months participation in *Wonders*, about 86% of students were considered to meet benchmark standards and about 14% of students scored below benchmark standards. One student, identified with special needs, required *Urgent Intervention* on the middle of year benchmark assessment.

First Grade

First grade students were administered *STAR Early Literacy* during the 2018-2019 school year (during the school's first year of *Wonders* implementation) and during the 2019-2020 school year (during the school's second year of *Wonders* implementation).

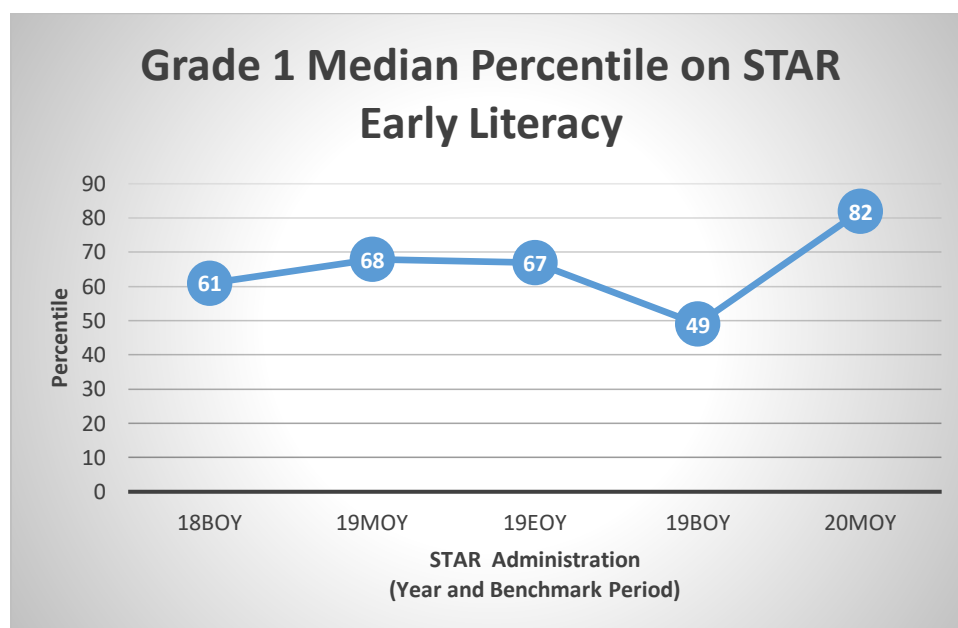


Figure 3. *Median Percentile on STAR Early Literacy, Grade 1*

As shown in Figure 3, before exposure to *Wonders* and in the fall of the 2018-2019 school year, the median percentile rank on *STAR Early Literacy* was about 61 percentile units (18BOY). Throughout the first year of exposure to *Wonders*, the median percentile score increased from 61 at the beginning of the year, to 68 in the middle of the year, and held steady at 67 by the end of the year. The overall gain from the beginning of the year to the end of the year was about 6 percentile units, overall. At the start of the second year of *Wonders* (19BOY) after the ‘summer slide’, the median percentile on *STAR Early Literacy* was 49. By the middle of the school year (20MOY), the median percentile increased to 82, reflecting a gain of about 33 percentile units from the beginning of the year to the middle of the year.

Using the Wilcoxon Signed-Ranks Test, the difference in ranking between the score on the 18BOY assessment before exposure to *Wonders* and the score on the 20MOY assessment immediately before school closure and during the second year of exposure to *Wonders* was statistically significant ($p = .000$). The overall gain in percentile rank was about 21 percentile units.

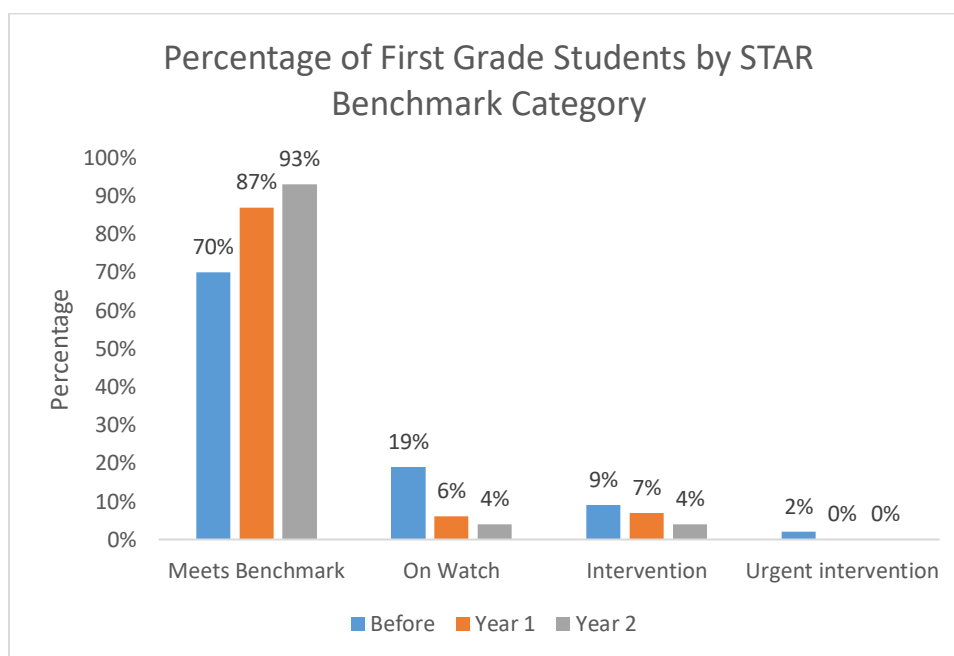


Figure 4. *Percentage of First Grade Students by Benchmark Category on STAR*

Figure 4 shows the percentage of first grade students scoring within each benchmark category on the *STAR Early Literacy Reading* assessment before exposure to *Wonders*, at the end of the first year of program participation (Year 1) and at the middle of the school year and during the second year of program participation (Year 2). Recall that Year 2 results reflect performance in the *middle* of the school year and immediately preceding the district-wide closure. As shown, before participating in *Wonders* the percentage of students meeting benchmark standards was about 70%; roughly 30% of students scored below benchmark (e.g., *On Watch*, *Intervention*, or *Urgent Intervention*). Following the first year of participation in *Wonders* and at the end of the school year, roughly 87% of students met benchmark standards, and about 13% of students scored below benchmark standards. During the second year of program participation, about 93% of students met benchmark standards and roughly 8% of students did not meet benchmark standards. Following the first and second years of program participation, none of the students were determined to require urgent intervention.

Second Grade

Second grade students were administered the *STAR Early Literacy* assessment during the 2017-2018 school year and before program implementation. These students were administered the *STAR Reading* assessments during the 2018-2019 school year during the first year of *Wonders* instruction, and again during the 2019-2020 school year during the second year of *Wonders* instruction. Because the *STAR Early Literacy* assessment and the *STAR Reading* assessment highlight different skills, student scores on the *STAR Reading* assessment will be shown here.

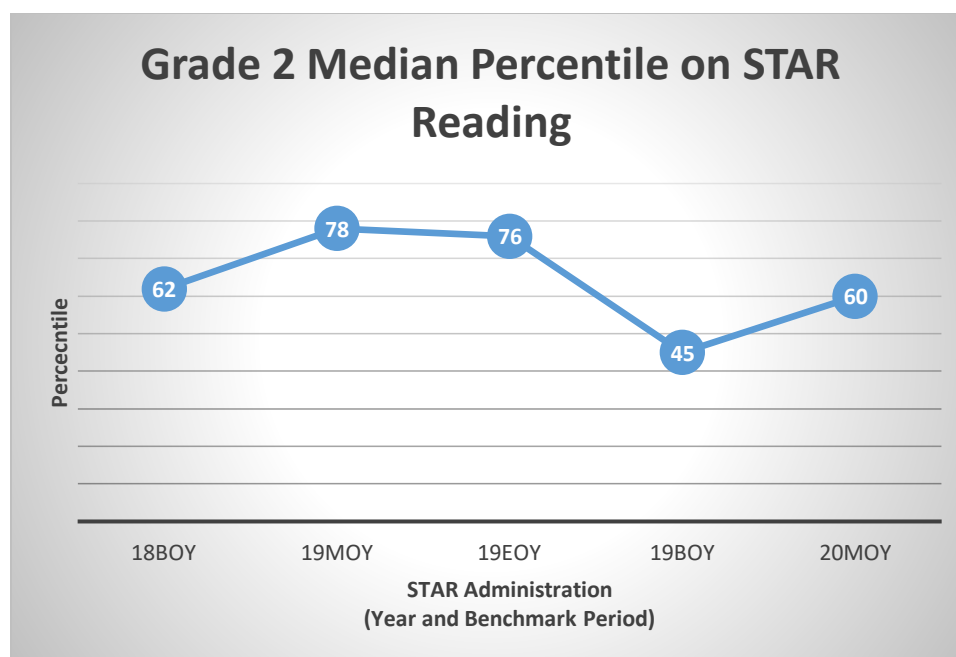


Figure 5. *Median Percentile on STAR Reading, Grade 2*

Figure 5 shows the second grade median score on the *STAR Reading* assessment immediately before exposure to *Wonders* (18BOY), throughout the first year (19MOY, 19EOY) and during the second year until the school closure in the spring (19BOY and 20MOY). At the start, the median percentile rank on the *STAR Reading* assessment was about 62 percentile units (18BOY), increased to 78 units (19MOY) during the middle of the first year, and dropped slightly to 76 by the end of the first year (19EOY). The gain in median percentile score was about 14 units for the first year and the difference in ranking was statistically significant for the first year of implementation (Wilcoxon Signed-Ranks Test $p = .001$).

At the start of the second year of exposure to *Wonders*, the median percentile score decreased at the start of the school year after summer break (19BOY). By the middle of the second year (20MOY), the median percentile increased to 60, a gain of 15 percentile units from fall to winter. The difference in ranking was statistically significant for the second year of implementation of *Wonders* instruction (Wilcoxon Signed-Ranks Test $p = .000$) but not longitudinally, across both years (e.g., 18BOY to 20MOY).

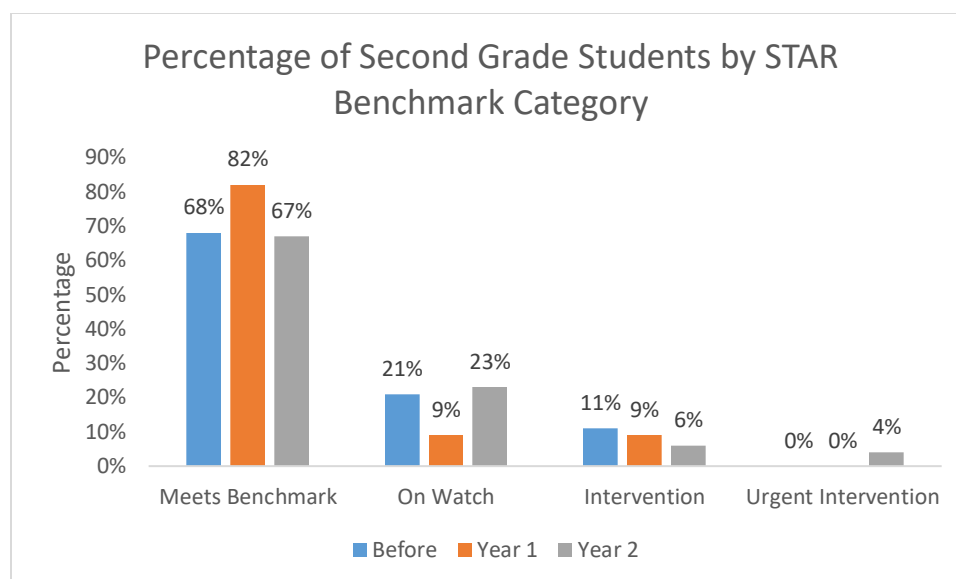


Figure 6. *Percentage of Second Grade Students by Benchmark Category on STAR Reading*

Figure 6 shows the percentage of second grade students scoring within each benchmark category on the *STAR Reading* assessment before exposure to *Wonders*, at the end of the first year of program participation (Year 1) and at the middle of the school year and during the second year of program participation (Year 2). Recall that Year 2 results reflect performance in the *middle* of the school year and immediately preceding the district-wide closure. Before participating in *Wonders* about 68% of students met benchmark reading standards, and about 31% did not meet benchmark standards (*On Watch*, *Intervention*, or *Urgent Intervention*). Following the first year of participation in *Wonders* and at the end of the school year, about 82% of students met benchmark standards, and about 18% of students scored below benchmark standards. During the second year of program participation, about 67% of students met benchmark standards and roughly 33% of students did not meet benchmark standards. Of students not meeting benchmarks, six students were identified with special needs.

Third Grade

Like second grade students, third grade students were administered the *STAR Early Literacy* assessment during the 2017-2018 school year (before program implementation) and changed to *STAR Reading* during the 2018-2019 and 2019-2020 school years. Student scores on the *STAR Reading* assessment will be shown here.

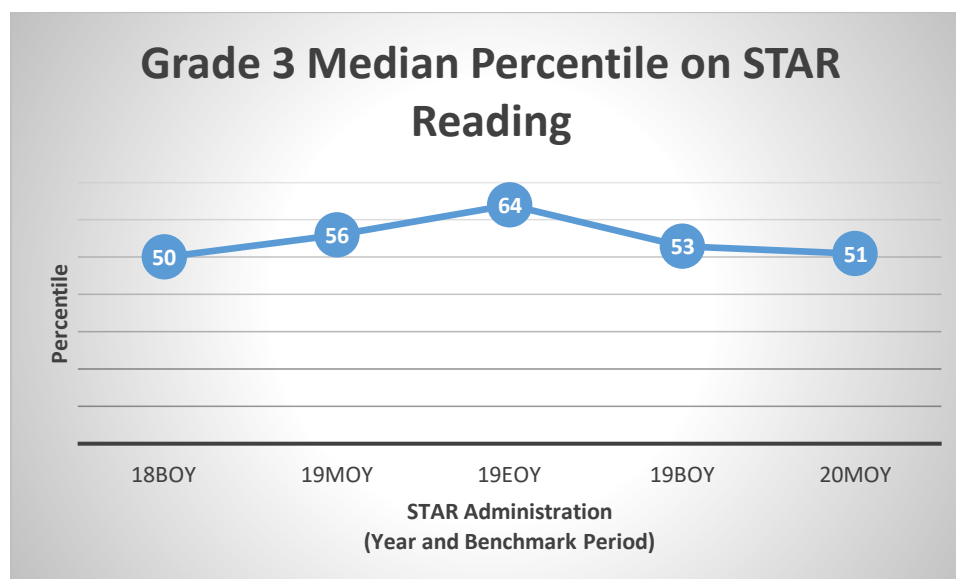


Figure 7. Median Percentile on STAR Reading, Grade 3

Figure 7 shows third graders' median score on the *STAR Reading* assessment immediately before exposure to *Wonders* (18BOY), throughout the first year (19MOY, 19EOY) and during the second year until the school closure in the spring (19BOY and 20MOY). At the start, the median percentile rank on the *STAR Reading* assessment was about 50 percentile units (18BOY), increased to 56 units (19MOY) during the middle of the first year, and increased further to 64 by the end of the first year (19EOY). The gain in median percentile score was about 14 units for the first year, and the difference in ranking on the end of year assessment following the first year of program implementation was statistically significant (Wilcoxon Signed-Rank Test $p = .000$).

At the start of the second year of exposure to *Wonders*, the median percentile score decreased at the start of the school year after summer break (19BOY) and decreased slightly to 51 during the middle of year administration immediately prior to the transition to remote learning. The difference in ranking was not significant during the second year and not significant longitudinally, across both years (e.g., 18BOY to 20MOY).

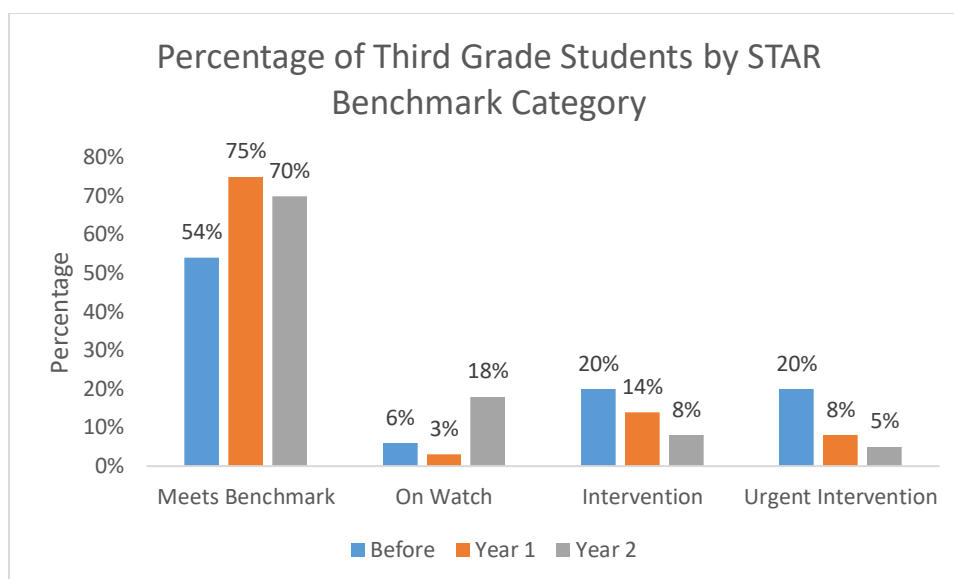


Figure 8. *Percentage of Third Grade Students by Benchmark Category on STAR Reading*

Third grade results by benchmark category and year are shown in Figure 8. Before participating in *Wonders* about roughly half of third graders (54%) met benchmark reading standards, while 46% did not meet benchmark standards (e.g., *On Watch*, *Intervention*, or *Urgent Intervention*). Following the first year of participation in *Wonders* and at the end of the school year, about 75% of students met benchmark standards, and about 25% of students scored below benchmark standards. During the second year of program participation, about 70% of students met benchmark standards and roughly 30% of students did not meet benchmark standards. Of those not meeting benchmark standards, half ($n = 6$) of students were identified with special needs.

Fourth Grade

Fourth grade students were consistently administered the *STAR Reading* assessment across multiple years: during the 2017-2018 school year (before program implementation), during the first year of *Wonders*, and during the second year of *Wonders*.

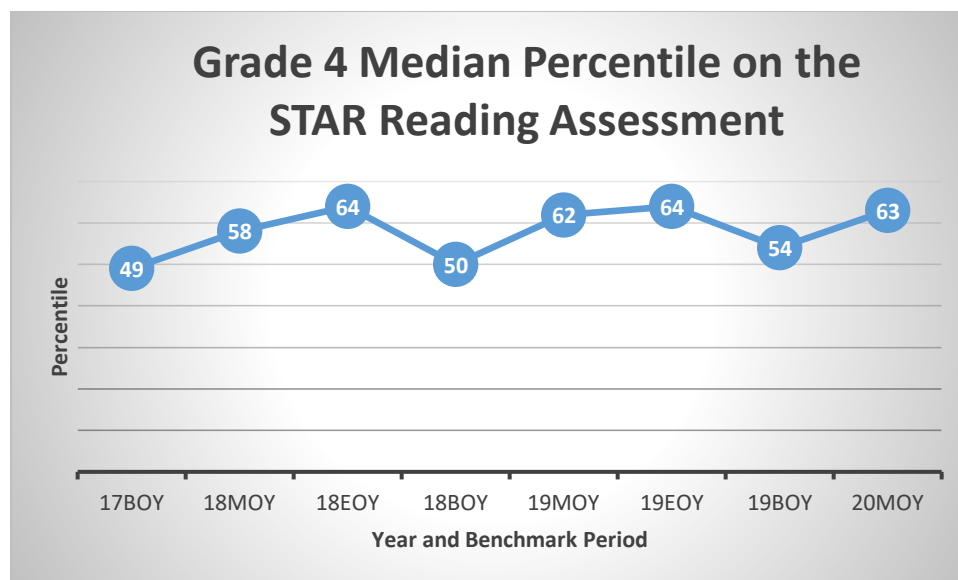


Figure 9. *Median Percentile on STAR Reading, Grade 4*

Fourth grade median percentile scores are presented in Figure 9. During the 2017-2018 school year prior to *Wonders*, students demonstrated a 15 unit gain in median percentile units (17BOY to 18EOY). Students demonstrated similar gains of about 14 percentile units during the first year, and about 9 percentile units across several months of use (fall to winter) during the second year. There were no significant differences in percentile gains, across years.

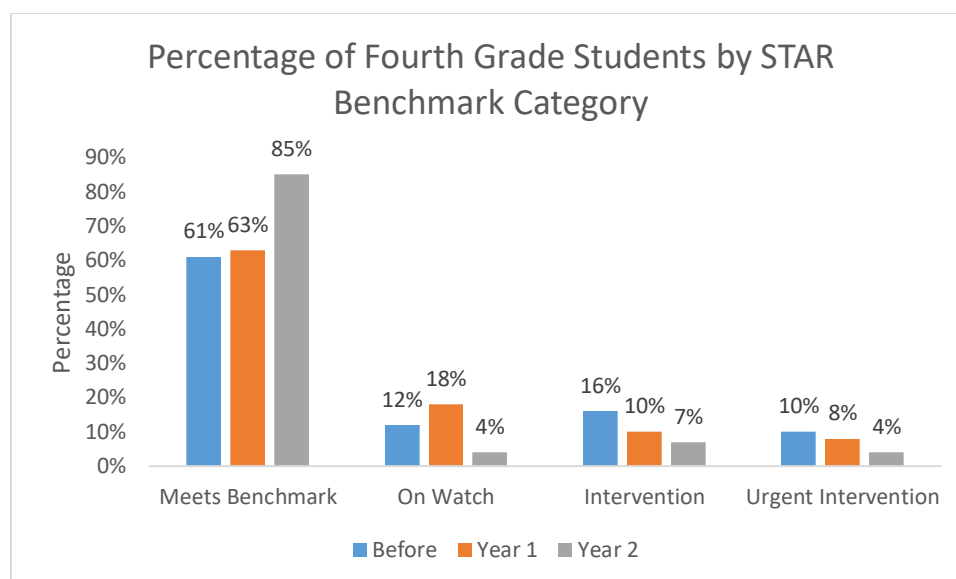


Figure 10. *Percentage of Fourth Grade Students by Benchmark Category on STAR Reading*

Before participating in *Wonders* about 61% of fourth grade students met benchmark reading standards, while 38% did not meet benchmark standards (Figure 10). Following the first year of participation in *Wonders* and at the end of the school year, about 63% of students met benchmark standards, increasing to 85% by the middle of the year during Year 2 implementation. The percentage of fourth graders who did not meet benchmark standards (e.g., *On Watch*, *Intervention* and *Urgent Intervention*) decreased from 36% following the first year of implementation to 15% during the second year of implementation. Of those not meeting benchmark standards, over half of students ($n = 5$) were identified with special needs.

Fifth Grade

Like fourth graders, fifth graders were consistently administered the *STAR Reading* assessment across multiple years: during the 2017-2018 school year (before program implementation), during the first year of *Wonders*, and during the second year of *Wonders*.

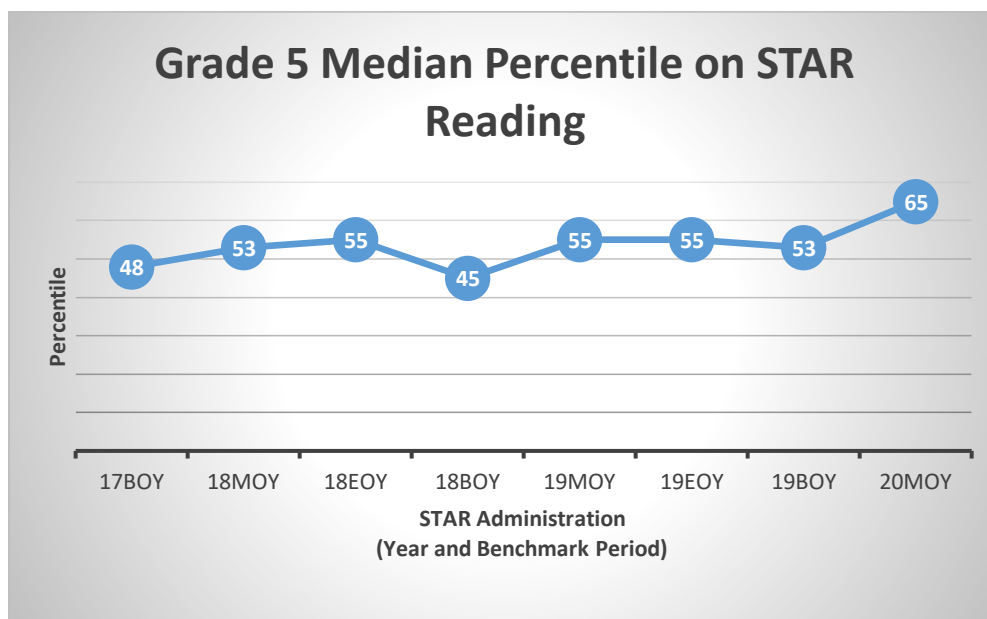


Figure 11. *Median Percentile on STAR Reading, Grade 5*

As shown in Figure 11, fifth grade median percentile scores increased by about 7 units the year before *Wonders* implementation (2017-2018) and increased by about 10 units the first year of *Wonders* implementation (2018-2019). During the first year of implementation, the difference in ranking was not statistically significant.

During the second year of *Wonders*, students demonstrated a gain of 12 percentile units from the beginning of the year (19BOY) to the middle of the year (20MOY). The difference in ranking on the beginning of year median percentile to the middle of year median percentile was significant (Wilcoxon Signed-Rank Test $p = .003$), as was the difference in ranking overall (17BOY to 20MOY).

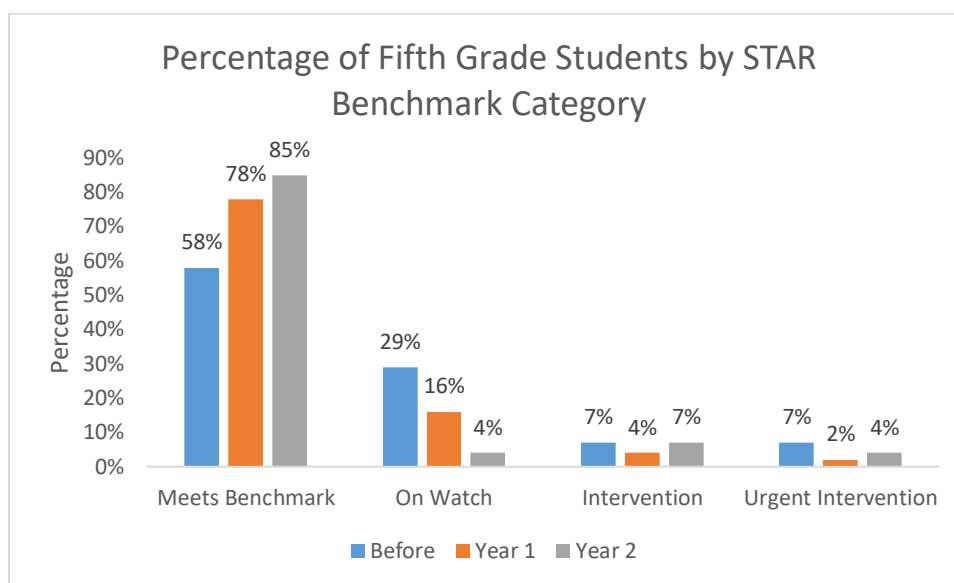


Figure 12. *Percentage of Fifth Grade Students by Benchmark Category on STAR Reading*

Fifth grade results by benchmark category and year are shown in Figure 12. Before participating in *Wonders* about 58% of students met benchmark reading standards. Following the first year of participation in *Wonders* and at the end of the school year, about 78% of students met benchmark standards. During the second year and preceding the school-wide closure, 85% of fifth graders met benchmark standards. Conversely, before *Wonders* about 43% of students did not meet benchmark standards. After the first year, 22% did not meet benchmark standards and during the second year about 15% did not meet benchmark standards. Of those students not meeting benchmark standards, 70% of students ($n = 5$) were identified with special needs.

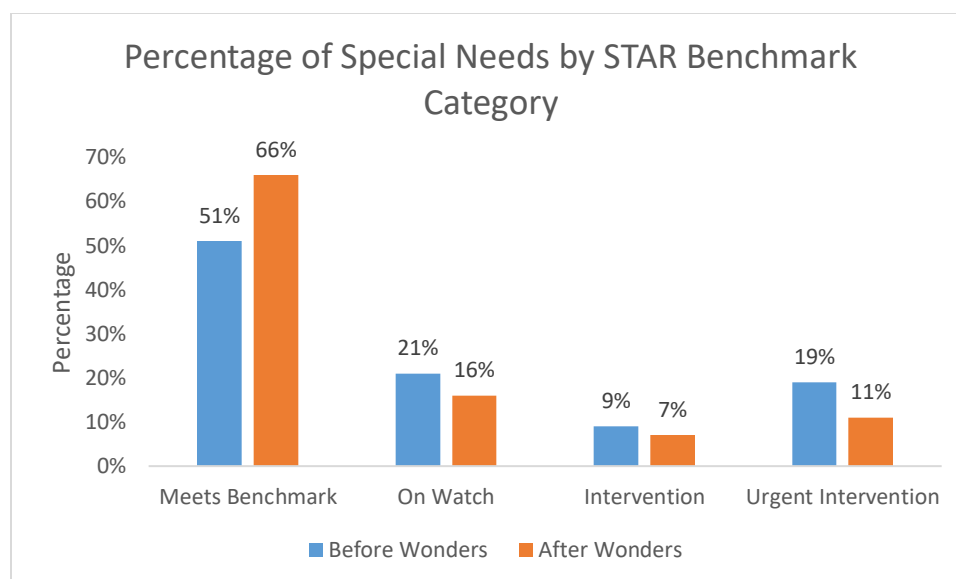


Figure 13. *Percentage of Students Identified with Special Needs by Benchmark Category on STAR Reading*

Figure 13 highlights the percentage of special needs students meeting benchmark standards immediately before exposure to *Wonders* and after exposure to *Wonders*. *Before Wonders* reflects the performance of students before any exposure to *Wonders*. *After Wonders* reflects performance of all students (grades K through 5) during the middle of the 2019-2020 school year and immediately before transitioning to remote instruction. Before participating in *Wonders*, roughly half of students (51%) met benchmark standards on STAR and about 49% did not meet standards. After exposure to *Wonders*, about two-thirds of students (66%) met benchmark standards. While 34% did not meet standards, a lower percentage required *Intervention* or *Urgent Intervention*.

Wisconsin Forward

Table 3 presents mean scores on the spring 2018 and 2019 administration of the Wisconsin *Forward*, Wisconsin's state-mandated achievement test in English Language Arts. Recall scores on the 2018 administration reflect student performance in the spring before implementation of *Wonders* and scores on the 2019 administration reflect student performance after one year of *Wonders*. The Wisconsin *Forward* was not administered in the spring of 2020. For this group of

students, the mean on the 2018 administration was about 548 units and the mean on the 2019 administration was about 592 units. The mean gain was about 44 units.

Table 3

Student Performance on the 2018 and 2019 administration of Wisconsin Forward, 5th Grade Students

Group	n	2018 Mean	Fall SD	2019 Mean	Spring SD	Gain Mean	Gain SD	Effect Size** <i>d</i>
<i>Wonders</i>	45	548.36	33.07	592.67	41.38	44.31*	24.27	1.18

*Statistically significant, $p = .000$

**Computed using Cohen's d

A paired-samples t-test indicated that the difference between student scores on the 2018 and 2019 administration of the Wisconsin *Forward* was statistically significant for those students in the *Wonders* group ($t_{(44)} = 12.25$, $p = .000$). Cohen's d was used as a measure of effect size, with .20 indicating a 'small' effect; .50 indicating a 'medium' effect; and .80 indicating a 'large' effect (Cohen, 1988).

Mean scores for students attending similar schools and not using *Wonders* and scores for students state-wide are provided for comparison (see Table 4). Here, the *Wonders* group includes the 5th grade cohort, who were 4th graders in 2019 and 3rd graders in 2018. Scores on the spring 2018 exam reflect student performance the year prior to implementation of *Wonders* and scores on the spring 2019 exam reflect the student performance after one year of implementation of *Wonders*. Scores from six comparable schools not using *Wonders* and scores state-wide are included as comparison. Before implementation of *Wonders*, the student mean score on the 2018 administration of the Wisconsin *Forward* was lower than that of students state-wide and students in the comparison schools with the exception of one (School C). Following one year of *Wonders*, the mean score on the Wisconsin *Forward* was highest for *Wonders* users and gains exceeded those in the comparison schools and students, state-wide.

Table 4

Mean Scores on the 2018 and 2019 administration of Wisconsin Forward, 5th Grade Cohort

Group	Mean 2018 (Before)	Mean 2019 (Year 1)	Mean Gain
<i>Wonders</i>	548	592	+ 44
School A	556	587	+ 31
School B	561	582	+ 21
School C	546	567	+ 21
School D	575	590	+ 15
School E	568	571	+ 3
School F	558	573	+15
Statewide	558	583	+25

Table 5

*Percentage of Students Scoring Proficient or Above on the 2018 and 2019 administration of
Wisconsin Forward, 5th Grade Cohort*

Group	% 2018 (Before)	% 2019 (Year 1)
<i>Wonders</i>	6	36
School A	33	48
School B	40	40
School C	31	28
School D	53	38
School E	56	33
School F	41	38
Statewide	40	43

Additionally, the percentage of students performing within the *Proficient* category on the Wisconsin *Forward* increased following the first year of *Wonders* implementation and scores were similar to comparison schools (see Table 5). Here, the *Wonders* group includes the 5th grade cohort, who were 4th graders in 2019 and were 3rd graders in 2018. In 2018, about 6% of students scored within the *Proficient* category, the lowest percentage among the comparison schools. In 2019 about 36% of students scored within the *Proficient* category, on par with several schools.

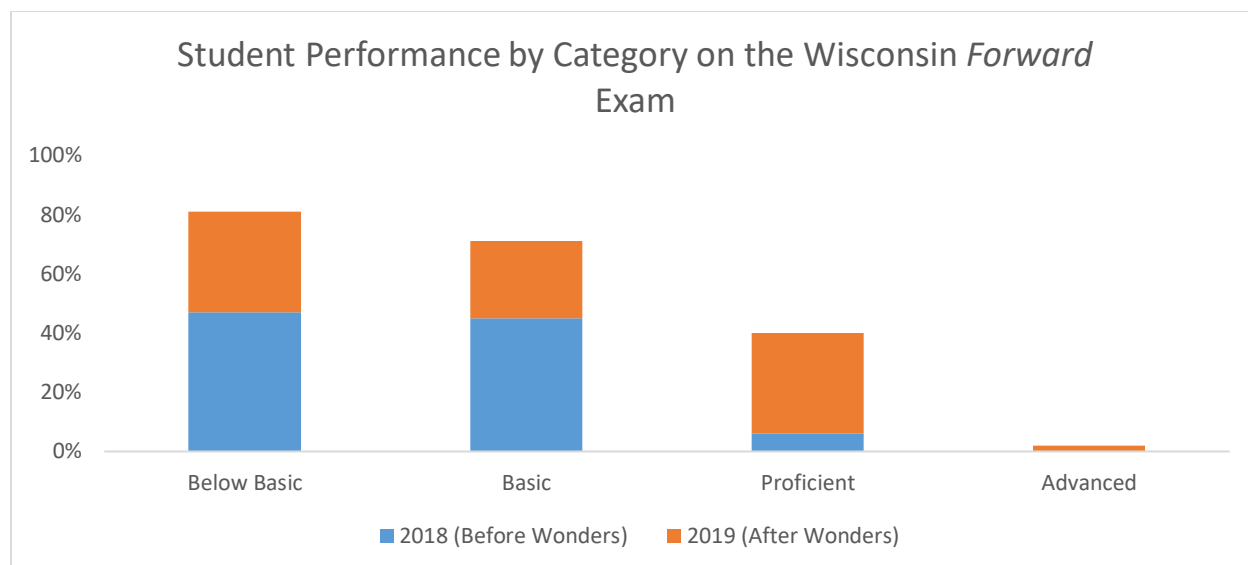


Figure 14. Student Performance by Category on the Wisconsin Forward Exam

Figure 14 highlights further student performance by category on the Wisconsin *Forward*. Before exposure to *Wonders*, the majority of 5th students — about 92% of those in the sample — performed in the *Below Basic* or *Basic* category on the Wisconsin *Forward* Exam in English Language Arts, and only about 6% were considered *Proficient*. No student scored within the *Advanced* category. After one year of exposure to *Wonders*, about 34% of students scored in the *Proficient* category and 2% of students scored in the *Advanced* category. Roughly 47% of students ($n = 21$) increased their performance category by at least one level from 2018 to 2019 and about 31% ($n=14$) increased their performance category from either *Below Basic* or *Basic* to *Proficient*.

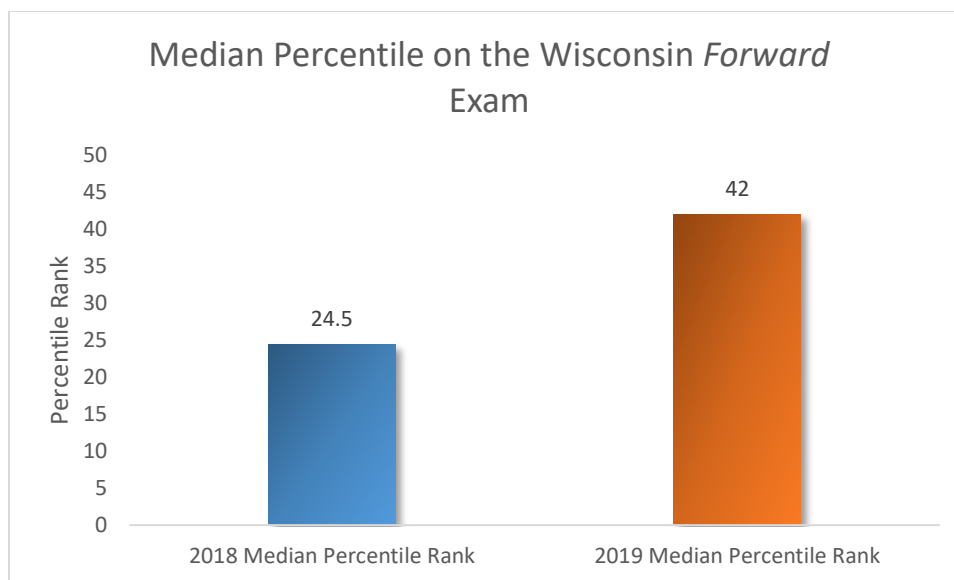


Figure 15. Median Percentile Rank on the Wisconsin Forward Exam

Figure 15 shows the median percentile rank on the Wisconsin *Forward* Exam in 2018 (before *Wonders*) and in 2019 (after one year of implementation). The percentile ranks show how well students performed compared with students across the state. In 2018 and before implementation of *Wonders*, the median percentile rank on the exam was about 25. This means that the typical student scored similarly or higher than about 25 percent of students, state-wide. After participating in the first year of *Wonders*, the median percentile rank on the exam was about 42. This means that the typical student scored similarly or higher than about 42 percent of students, state-wide.

Teacher Survey

In the spring of 2020 and to help capture teachers' perceptions about *Wonders*, teachers were asked to complete an online survey. The survey contained a combination of open-ended response and Likert items addressing teachers' general thoughts about *Wonders*, perceived effectiveness of *Wonders* for selected subgroups of students, strengths and weaknesses of provided professional development, and ease of use during the period of remote instruction. After one email follow up, 25 of 26 teachers responded to complete the survey.

What Were Teachers' Overall Impressions of Wonders?

The majority of teachers found *Wonders* effective for students, with 84% (n = 22) selecting *Extremely* or *Moderately* effective. Similarly, 76% (n = 19) of teachers responded that *Wonders Completely* or *Mostly* met instructional needs. When asked to elaborate on responses, several teachers mentioned the variety of resources and the structure as particular program strengths:

By using Wonders as a school, we are all on the same page. I feel like Wonders provides so many different stories for the kids to read to help with their love of reading, but also get them interested in other books. Students become more confident as they grow as a reader. Students learn the routine of Wonders quickly.

It meets the needs of all of my learners the high the on-level and the approaching Level. It also works for EL learners as well as students and with IEPs. It is also engaging and rigorous. The spiral of the curriculum is also powerful, and impressive as somebody who works with students across the grade levels.

It has rigorous, carefully selected texts that integrate with a weekly theme. It hones the student's essential skills, such as summarizing, but gives so many opportunities to allow students to go above and beyond the text. When this is seen as a resource instead of as a manual, I believe I have the tools I need to help my students succeed.

When asked what teachers most liked about *Wonders*, teachers provided a wealth of responses. The most frequent response included the quality and quantity of literature embedded within the program (40% of teachers), followed by the program's ease of use (24%):

The quality of literature and the good balance between fiction and nonfiction. Also, having the online planning component and presentation already completed made planning much more manageable.

I love how easy it is to use. I can still be me while I am teaching reading. I don't have to worry that I will miss a skill.

The interventions, the leveled readers, the spiral of the curriculum, the digital tools, ConnectED, lit anthology, EVERYTHING!

About 20% of teachers (n=5) each reported the online components, wealth of resources, and structure/consistency of *Wonders* were the program's most liked components:

I like all of the on-line components of Wonders. Very user friendly. The stories in the basal are great. There are plenty of resources to use. I also like having the leveled readers as options for small-group instruction.

I love that I have every resource in my toolbox that I can choose from. What I like MOST about Wonders is that every unit has carefully selected texts that are rich, engaging, and thematic. It would take me forever to have created something like that, and even then, my own bias wouldn't have selected some of the great stories/texts that are out there.

Less common responses included the rigor (8%), writing component (4%), and progress monitoring tools (4%).

When asked for recommendations for program improvement, responses were varied. About 24% (n = 6) of teachers each reported the time required to complete program activities was a challenge and the amount of available resources was sometimes overwhelming. Similarly, about 24% of teachers reported they would like to see more individualized writing and phonics activities. Nevertheless, the majority of teachers (84%) indicated they would use *Wonders* again:

I would choose Wonders because it is an enriching, engaging, and organized curriculum resource that benefits my students. Wonders provides captivating fiction and non-fiction stories that my students enjoy.

Rich and engaging texts-Tons of resources-Online Data platform-Best differentiation tool I have seen in reading curriculums-Online games - plus other online resources students can do at home. I guess that's more than one, but those are my reasons!

I think the overall curriculum is great and very easy to follow.

I'm seeing major gains from the previous students as they come to me. I believe that will just keep improving.

It has so many resources to use that are at various levels for all students.

The curriculum combines phonemic awareness, phonics, literature, high frequency words, and printing as well as the small group differentiation and WonderWorks intervention activities. Overall a solid curriculum with some weaker areas

How Well Does Wonders Support the Needs of Student Subgroups?

When asked to determine how well *Wonders* supports the instructional needs of selected subgroups of students, for example, students with special needs, gifted students, and English Language Learners, responses were varied. Roughly half (56%) of teachers reported that *Wonders Completely* or *Mostly* supported the instructional needs of students with special needs and about 36% of teachers felt the program *Somewhat* supported the needs of these students. Several teachers reported that the materials were too challenging for special education students, and several teachers reported they did not have students identified with special needs in their classrooms.

Special education student needs are individualized and sometimes a package program does not offer all these students need. The tier 2 interventions are great for my population. The Wonderworks materials are also great. The adaptive learning online is good too.

I really think the quality of Wonders has pushed students no matter if they have an IEP or not. Let me be VERY clear here. It is NOT Wonders job to meet the needs of a student who has an IEP. It is the job of a special education teacher to use Wonders, adapt, or modify it to break down the barriers preventing that particular student from succeeding. I believe Wonders has enough resources to do that.

Roughly half of teachers reported that *Wonders Completely* or *Mostly* supported the instructional needs of English Language Learners (52%) and about 40% felt the program *Somewhat* supported the needs of English Language Learners. It should be noted that many teachers reported they did not have English Language Learners in their classrooms and did not feel as though they could adequately provide feedback. Several teachers felt as though program components were

too challenging for ELL students. One teacher reported that it would be helpful to have materials in Spanish.

A higher percentage of teachers (72%) reported that *Wonders Completely* or *Mostly* supported the learning needs of gifted students, while 24% of teachers reported that *Wonders Somewhat* supported the needs of gifted students. Several teachers felt as though the leveled readers were not challenging enough for gifted students. Several teachers reported they did not have gifted students in their classrooms and therefore could not provide feedback.

What Are Students' Perceptions of Wonders?

When asked what students enjoyed most about *Wonders*, an overwhelming majority — 92%, or 23 of 25 teachers — reported that students most enjoyed the variety of stories and the read-alouds:

My students love the interactive read alouds. They love having the workshop small books and leveled readers at our small group tables (they get to have their own books)! They love Research & Inquiry with adjustments to meet our needs - partners fill out a bubble map about what they learned about theme for that week, then make a project that matches their bubble map. Sadly, some weeks we run out of time for this, so students don't get the full experience when this is so important for their learning.

There are many stories in the lit anthology that my students have enjoyed. They are often disappointed when we are unable to read the whole story.

Students additionally enjoyed the online/digital resources (24% of respondents) and the predictability of routines (16% of respondents):

My students very much enjoy the stories, how they all connect, and the online resources. Even stories that I myself may think they won't like, I'm usually surprised.

I think they liked how they knew what was coming next, the structure of the program. I also think they enjoyed most of the stories content.

When asked what students enjoyed least about *Wonders*, about 40% of teachers (n = 11) reported that students least enjoyed the assessments/tests, followed by 'Writing Workshop' (about 20%,

or 5 teachers). Less frequent responses included the rigor of the program, difficulty logging into the student dashboard, and length of stories, all reported by one teacher.

What Are Impressions of Professional Development Training?

Feedback regarding the quality of the professional development training was remarkably positive. Respondents perceived the consultants as very knowledgeable and supportive; one respondent referenced the trainers as a ‘dream team’ who was “constantly available and able to get me and my teachers out of the toughest jam”. When asked about particularly helpful components of the training, nearly all teachers reported that simply having a consultant available to answer questions was extremely helpful. Many teachers indicated the review of online components was especially appreciated, as was assistance in setting up the online calendar:

Teaching us about all of the components of Wonders. The weekly planner has been very helpful to print off to make sure to touch on all the "must dos" since there is so much.

Having the trainer come back after we were able to use Wonders. Had more questions once we were using the materials. I thought our trainer did a nice job and was knowledgeable of questions asked.

With distance learning for the final few months, the online training/connected information was most helpful

Several teachers expressed a desire for additional training, which would have provided opportunity for a more comprehensive review of the many resources included in the program. Suggestions for future trainings included a ‘live’ demonstration of a lesson (instead of video lessons) and grouping training sessions by grade levels. Aside from the training and support provided directly by trainers and consultants, teachers supported each other by engaging in weekly collaborative sessions to examine progress monitoring data, monitor use of curriculum to ensure consistency, share materials, examine resources, and answer questions:

We plan together each week. We look at data collected from progress monitoring and other assessments. We look at small groups work as well. We talk about how we grade assessments so that we are doing it the same way. We talk about how we do things for the daily routine of Wonders.

We share ideas on how we are using parts of the curriculum and how to make it fit in 31 with our day. We meet weekly to discuss the upcoming weeks and what challenges we might see coming.

We discuss the curriculum on a weekly basis - how we are teaching it, what works best for our students, and use our data to reflect on how we can better use Wonders to reach our students.

How did you Implement Wonders During Periods of Remote Learning?

Of particular interest was teachers' experience implementing *Wonders* during the spring of 2020 as school districts transitioned to remote learning in the midst of (or due to) following the COVID-19 pandemic. About 76% of respondents (n = 19) reported they found *Wonders* easy to use when teaching remotely. Program elements most frequently utilized by teachers included the weekly stories/read aloud (n = 15), games (n = 8), vocabulary cards (n = 7), and skills (n = 3):

When it came to assignments, Wonders was very easy to use. Everything online was easily assigned through Google Classroom. Teaching was a little more challenging, but that is true for all subjects right now.

Having the on-line component has been so helpful! I used the Building Background/Introduce Vocabulary slides, assigned some of the games/had students read the story by signing into Wonders.

Our students were able to log in to Wonders and read all of the stories online as well as listen to them if they were unable to read the content.

When asked what features of *Wonders* were particularly useful for remote teaching, 40% (n = 10) of teachers specifically mentioned that sharing and uploading materials through Google Classroom was extremely easy and facilitated access by students and teachers. Many teachers reported accessing the stories and interactive games:

The online activities were great for remote teaching. It was wonderful that there was a link to directly assign readings and games through Google Classroom.

It is SUPER easy to upload to Google Classroom which was really great!

The stories were what we used but moving forward I can see the interactive games being useful as well as immediate access to spelling lists and vocabulary.

We used many of the pieces in collaboration with Google classroom to create assignments for our students. The games and stories were easily accessible to teachers as well as students.

It was helpful that the students were already familiar with the online component of Wonders. I think it may be harder if I don't get to use in person time to teach the students how to use the site and where to access the activities I would like them to do.

While many teachers reported that *Wonders* was easy to implement during the period of remote learning, there were some challenges. Several teachers noted that sharing assignments through Google Classroom, while easier, prevented teachers from accessing student results unless assignments were shared directly through the *Wonders* platform—yet many teachers opted to share assignments this way to make things easier for parents:

We could not see the results of games and reading assignments when assignments were shared through Google Classroom. Instead, the assignments needed to be assigned through the Wonders platform. It was a bit more confusing for students.

When games or activities are assigned through Google Classroom we are unable to see the student's work.

It has been more difficult to watch the progress of my students when sharing Wonders activities through Google Classroom.

There are just certain routines that you can't recreate when you are not in person. That is nothing Wonders can fix. We just need to be with our kids.

Finally, when asked what four words teachers would use to describe *Wonders*, about 36% (n = 9) of teachers used terms to reflect the comprehensive nature of the program (“complete”, “resourceful”, “loaded”, “abundant”, “thorough”). Similarly, about 36% (n = 9) of teachers used terms that indicate the program is engaging (e.g., “fun”, “interesting”, “interactive”, “enjoyable”, “exciting”). About 32% of teachers (n = 8) used terms to reflect program rigor (e.g., “challenging”, “rigorous”, “advanced”, “enriching”). Finally, about 20% (n = 5) of teachers used terms to reflect the structured nature of the program (e.g., “predictable”, “structured”, “organized”).

Summary

This report provides an analysis of student performance on several measures of achievement (*STAR Reading*, *STAR Early Literacy*, and *Wisconsin Forward*) across one full year of *Wonders* participation and a portion of a second year. The report additionally captures teachers' perceptions about *Wonders* while implementing the program in a traditional and remote setting. Care should be used when interpreting these results, for several reasons. It should be noted that results presented reflect student performance *for only a portion of the second year*, as students were not administered a third benchmark assessment at the end of the school year and following the school's transition to remote learning. It also should be noted that some comparisons across years could not be made, given the differing skills highlighted on the *STAR Reading* and *STAR Early Literacy* assessments. Finally, the cyclical nature of performance on repeated measures of this type (e.g., the consistent pattern of lower scores on the beginning of the year administration following summer break) sometimes makes interpretation challenging.

Across grade levels, gains were noted for several groups of students. For example, first graders demonstrated gains in the median percentile score during the second year; the median percentile score increased from 49 at the beginning of the second year to 82 by the middle of the second year (immediately before the school transitioned to remote learning). Further, about 93% of first grade students were determined to meet benchmark on *STAR Reading*. Second grade students exhibited an increase in about 14 percentile units during the first year and about 15 percentile units during the second year, with significant gains both years. Third graders demonstrated marked gains during the first year of implementation, with the median percentile increasing by about 14 units. Fourth graders exhibited gains across both years of implementation; however gains were also noted during the year before implementation of *Wonders*. A higher percentage of fourth graders did however, meet benchmark during the second year of *Wonders* (85%) compared to the first year (63%) and the year prior to implementation (61%). Fifth graders demonstrated significant gains the second year of implementation and a higher percentage of students (85%) met benchmark standards on the *STAR*. Fifth graders additionally made significant gains on the *Wisconsin Forward* from 2018 to 2019 following the first year of *Wonders* implementation, and gains were comparable to students attending comparable schools. The percentage of students scoring within the *Proficient* category increased, from 6% before implementation to 34% after implementation.

Teachers' perception of *Wonders* was largely positive. For example, the majority of teachers felt that *Wonders* was effective (84%) and met instructional needs (76%). Teachers felt that *Wonders* contains a wealth of resources which sometimes seemed overwhelming; yet 84% of teachers indicated they would use *Wonders* again. Over half of teachers felt that *Wonders* supported the needs of student subgroups (e.g., special needs or gifted students); however, it should be noted that many teachers reported they could not adequately respond, since they did not have these students in their classrooms. Teachers reported that students enjoyed the variety of stories and the digital resources, but least enjoyed the assessments and tests. Teachers largely considered *Wonders* easy to use during the period of remote learning and many cited the ability to upload and share materials through Google Classroom an important feature. Finally, teachers described *Wonders* in terms that reflect the program's comprehensive, engaging, rigorous, and structured format.

References

- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. (2nd Ed.). Hillsdale, NJ: Lawrence Earlbaum Associates
- Renaissance Learning (2016). STAR early literacy technical manual. Retrieved from: <http://doc.renlearn.com/KMNet/R004327609GJB0BF.pdf>.
- Renaissance Learning, Inc (2018). Renaissance STAR Reading and STAR Early Literacy: Unified Benchmarks and Cut Scores. Retrieved from <https://help2.renaissance.com/US/PDF/SR/SRUnifiedBenchmarksCutScores.pdf>.
- Renaissance Learning, Inc (2020). Defining benchmarks in Star assessments. Retrieved from www.renaissance.com.

U.S. Department of Education: National Center on Intensive Intervention (2016). Review of progress monitoring tools [Review of STAR Early Literacy]. Washington, DC: Author. Retrieved from <http://www.intensiveintervention.org/progress-monitoring/star-early-literacy-3>.