Data Analysis and Evaluation of the McGraw Hill's

***Everyday Mathematics* Program, Its Implementation,**

and its Impact on Test Scores in an Arkansas District

Daniel Henry, Ph.D., Director, Auburn Center for Evaluation

David Marshall, Ph.D., Auburn University

AUBURN
CENTER FOR
EVALUATION

AUBURN
UNIVERSITY

Executive Summary

The participating dstrict is located in Central Arkansas and uses a McGraw Hill mathematics curricula in its public schools. *Everyday Mathematics* is used for grades K-6. The Auburn Center for Evaluation (ACE) was asked to analyze math test data for students using the program to answer four evaluation questions:

1. Do differences exist between fall and spring MAP assessment scores for schools using the *Everyday Mathematics* curriculum in the district?

2. How do fall and spring MAP assessment scores compare to normative data for schools using the *Everyday Mathematics* curriculum in the district?

3. What barriers and supports did the curriculum have for teachers implementing the program?

4. What was the experience of teachers who were using *Everyday Mathematics* in their classrooms?

To answer the first and second questions, paired-samples *t*-tests were employed. The second question was answered with descriptive statistics. To answer the third question and fourth questions, multiple focus groups (held over Zoom) and a two day site visit with classroom observations, teacher and administrator interviews were employed to understand the lived experience of participating teachers in implementing the curriculum. Although this (qualitative) feedback has already been provided to McGraw Hill, it is included here to provide a more complete understanding of the implementation.

Key Findings

1. Overall, spring assessment scores were significantly greater than fall assessment scores across grade levels and both years of data analyzed. In almost all cases, effect sizes were substantial.

2. Overall, students earned scores on the fall and spring assessments that were below the normative data. Exceptions to this can be found on the 2020-21 assessments for grades K-2.

3. Asian American students outperformed the normative data for almost all fall and spring assessment comparisons.

4. Across almost all comparisons for both years, students' growth scores from fall-to-spring fell short of the NWEA normative data. Exceptions were found in a few subgroups.

Limitations of the Evaluation

1. Data for the evaluation were completely dependent on the quality and quantity of information provided by the district.

2. Transience or mobility of the student population in participating schools is a concern. Only students with both fall and spring assessment scores were considered for these analyses, and it is possible that they differ systematically from students who failed to take one of the assessments.

3. Intermediate and long-term shifts in knowledge, attitudes, perceptions, and achievements in mathematics may not have resulted from the McGraw Hill curricula alone. Many schools in Arkansas have multiple federal, state, and local initiatives in effect at the same time. Since many of these initiatives are aimed at improving academic achievement, this may or may not have had an impact on the assessment scores – or may have had less of an impact compared to other factors not accounted for in these analyses.

4. Interpretations based on statistical significance alone should be made with caution.

5. The two years for which data were analyzed involved substantial disruption to the educational enterprise. The COVID-19 pandemic closed schools for in-person learning in the spring of 2020; what followed was an uneven transition to emergency remote instruction.

6. The 2020-21 school year took place during the ongoing pandemic and families were given the option of having their children attend school in one of three instructional options including: (1) in-person learning five days per week; (2) blended learning with two days per week of in-person learning; and (3) full-time remote learning. Differences in

learning associated with the different learning modalities likely had a substantial impact on assessment scores, and that is not factored into these analyses.

Data and Method

As part of a larger evaluation effort, two evaluation questions were answered:

1. Do differences exist between fall and spring MAP assessment scores for schools using the *Everyday Mathematics* curriculum in the district?

2. How do fall and spring MAP assessment scores compare to normative data for schools using the *Everyday Mathematics* curriculum in the district?

The study site is located in Central Arkansas and serves approximately 12,000 students. They employ McGraw Hill mathematics curricula in their schools. *Everyday Mathematics* is used for grades K-6. Of interest was to learn if statistically significant differences existed between the fall and spring assessment administrations, and to learn how they compare with the Northwest Evaluation Association's (NWEA) normative data.

Data Sources and Analyses

Data were provided by the district. The MAP assessments are administered three times annually in the school district. Fall and spring assessment scores were compared for this evaluation. All analyses were conducted using Stata 15. The first and third evaluation questions sought to learn if differences existed within subjects (i.e., from two points in time among the same group of individuals). As such, paired-samples *t*-tests were conducted to learn if within-subjects differences existed. Effect sizes were calculated when warranted. The 2019-20 data for grades 3-10 included race/ethnicity subgroups. The 2020-21 data included race/ethnicity and gender subgroups for grades K-10, and English language learner (ELL), Section 504, special education, and talented and gifted (TAG) subgroups for grades 3-10. All data only reflect those test takers who completed both fall and spring assessments each year. Subgroups with fewer than 10 individuals were excluded from this report as an ethical consideration. Subgroups for students

identified as *American Indian or Alaskan Native* or *Native Hawaiian or Other Pacific Islander* were not included in any of the analyses for this evaluation due to small sample sizes. These individuals are included in the overall grade-level statistics.

COVID-19 Pandemic

Data for the 2019-20 and 2020-21 school years were analyzed separately. Both school years took place during the ongoing COVID-19 pandemic and as such, the conditions under which the two school years occurred differed substantially – both from normal operation and from each other. The 2019-20 school year featured a significant disruption to learning when schools closed for in-person instruction in March 2020 as part of an effort to curb the spread of the virus (Marshall et al., 2020). Schools transitioned to emergency remote instruction for the remainder of the school year. The 2020-21 school year opened during the ongoing pandemic and school districts across the United States reopened with a range of instructional options for families (Marshall & Bradley-Dorsey, 2022). The district's (2020) *Roadmap to Reopening* plan allowed families to choose between three instructional options: (1) traditional learning five days per week in person with half-days on Wednesdays; (2) blended learning with two days per week of in-person instruction and three days of remote learning; and (3) full-time online learning. Students could transition between the three options at different times during the year. A student who attended remotely to begin the school year who transitioned to a blended learning option might have had a very different schooling experience than a student who was in-person for the duration of the school year. It is imperative to view this evaluation's findings with these contextual factors in mind.

Comparison of Fall and Spring Scores

The *Everyday Mathematics* curriculum was used in the district for grades K-6. The first two evaluation questions seek to learn: (1) whether the fall and spring scores statistically differed from one another and (2) how these scores and the associated growth score compared with normative data. The 2019-20 data were examined first. For each grade 3-6 overall, and across all subgroups, the mean spring MAP assessment scores were statistically greater than the fall MAP assessment score. The 2020-21 data were compared for grades K-6. With the lone exception of sixth grade students who received special education services, all analyses conducted yielded findings indicating that spring assessment scores were significantly better than fall scores. See Table 1 for means, standard deviations, *p*-values, and effect sizes.

Table 1.

Fall to Spring Paired-Samples t-test Findings

		Fall		Spring			
	N	**M**	**SD**	**M**	**SD**	***p*-value**	**Cohen's d**
2019-2020							
Grade 3	782	185.89	5.66	196.02	5.82	<.001	8.14
Grade 4	768	196.90	5.64	204.72	6.05	<.001	6.15
Grade 5	865	206.80	6.07	213.97	6.59	<.001	5.24
Grade 6	828	208.76	3.92	213.56	4.17	<.001	3.81
2020-2021							
Grade K	692	144.26	3.50	158.07	4.15	<.001	4.61
Grade 1	759	161.49	6.28	181.32	7.59	<.001	6.40
Grade 2	703	175.04	4.78	185.36	5.90	<.001	6.55
Grade 3	706	185.64	5.82	194.04	7.26	<.001	3.23
Grade 4	718	195.58	6.50	203.85	8.37	<.001	3.47
Grade 5	731	205.16	6.46	209.70	7.78	<.001	1.80
Grade 6	610	209.33	4.30	212.81	5.51	<.001	1.08

Comparison with Normative Data

The fall and spring assessment scores were also compared with MAP's normative data. Data associated with the 2019-20 assessment were compared with the 2015 normative data (Northwest Evaluation Association, 2015). The fall and spring normed scores represent the 50th percentile of scores; half of all scores fell below the mean and exceeded it. The same is true for the normed growth scores. The normed score represents the 50th percentile for growth among test takers. Half of all MAP test takers will have growth that exceeds this mean; half will fall short of it. Overall, only Asian American students earned scores on fall and spring MAP assessments that exceeded the norms. No subgroup exceeded the norms in terms of fall to spring growth. See Tables 2-5 for a comparison of normative data and district fall assessment, spring assessment, and growth scores.

Table 2.

2019-20 District MAP Scores Compared with Normative Data – Grade 3

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	782	190.4	185.89	203.4	196.02	13.00	10.13
Asian	34	190.4	193.21	203.4	202.61	13.00	9.40
Black	326	190.4	181.67	203.4	191.79	13.00	10.12
Latino/a	91	190.4	183.45	203.4	192.52	13.00	9.07
White	448	190.4	192.92	203.4	203.33	13.00	10.41
Two or More Races	30	190.4	182.68	203.4	193.73	13.00	11.05
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Table 3.

2019-20 District MAP Scores Compared with Normative Data – Grade 4

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	768	201.9	196.90	213.5	204.72	11.60	7.82
Asian	30	201.9	211.13	213.5	217.98	11.60	6.85
Black	296	201.9	191.75	213.5	199.13	11.60	7.38
Latino/a	84	201.9	194.88	213.5	204.31	11.60	9.43
White	315	201.9	200.42	213.5	208.48	11.60	8.06
Two or More Races	39	201.9	200.10	213.5	206.38	11.60	6.28
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD	Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Table 4.

2019-20 District MAP Scores Compared with Normative Data – Grade 5

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	865	211.4	206.80	221.4	213.97	9.90	7.17
Asian	27	211.4	220.27	221.4	231.19	9.90	10.92
Black	359	211.4	201.73	221.4	208.29	9.90	6.56
Latino/a	85	211.4	206.08	221.4	212.94	9.90	6.86
White	353	211.4	211.03	221.4	218.59	9.90	7.56
Two or More Races	38	211.4	207.81	221.4	214.63	9.90	6.82
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD	Above norm by more than 1 SD	

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Table 5.

2019-20 District MAP Scores Compared with Normative Data – Grade 6

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	828	217.60	208.76	225.30	213.56	7.70	4.81
Asian	13	217.60	222.38	225.30	227.77	7.70	5.38
Black	362	217.60	204.91	225.30	208.96	7.70	4.05
Latino/a	88	217.60	202.41	225.30	208.37	7.70	5.96
White	337	217.60	214.06	225.30	219.62	7.70	5.56
Two or More Races	26	217.60	208.79	225.30	211.15	7.70	2.36
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD	Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Data associated with the 2020-21 assessment were compared with the 2020 normative data (Northwest Evaluation Association, 2020). Subgroup data for ELL students, students with 504 plans, students receiving special education services, and TAG students for grades 3-6. Overall, *Everyday Mathematics* students earned scores exceeding the normative data on the fall assessment in grades K-2; this was true for most subgroups for these grade levels as well. Only Asian Americans earned the scores exceeding the normed scores in every grade level. White students exceeded the normed scores in grades K, 1, 2, and 4. Consistent with what one might expect, students identified as talented and gifted (TAG) earned scores exceeding the norm and

students receiving special education services earned scores below the norm for each grade level. Only Asian American fifth grade students (M=13.20) had a growth score that exceeded the 2020 MAP normed data. See Tables 6-12 for a comparison of normative data and district fall assessment, spring assessment, and growth scores.

Table 6.

2020-21 District MAP Scores Compared with Normative Data – Grade K

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	692	139.56	144.26	157.11	158.07	17.54	13.81
Asian	34	139.56	155.64	157.11	169.86	17.54	14.22
Black	273	139.56	141.22	157.11	153.26	17.54	12.03
Latino/a	78	139.56	141.01	157.11	154.04	17.54	13.03
White	372	139.56	146.90	157.11	162.68	17.54	15.77
Two or More Races	31	139.56	143.44	157.11	158.02	17.54	14.58
Female	353	139.56	144.70	157.11	158.80	17.54	14.10
Male	339	139.56	143.80	157.11	157.30	17.54	13.50
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD	Above norm by more than 1 SD	

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Table 7.

2020-21 District MAP Scores Compared with Normative Data – Grade 1

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	759	160.05	161.49	176.40	172.70	16.35	11.21
Asian	22	160.05	173.95	176.40	181.32	16.35	7.36
Black	296	160.05	158.50	176.40	168.40	16.35	9.90
Latino/a	89	160.05	159.26	176.40	170.66	16.35	11.40
White	311	160.05	164.00	176.40	176.95	16.35	12.95
Two or More Races	94	160.05	167.20	176.40	170.64	16.35	2.13
Female	350	160.05	161.80	176.40	172.10	16.35	10.30
Male	409	160.05	143.20	176.40	173.20	16.35	12.00
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Table 8.

2020-21 District MAP Scores Compared with Normative Data – Grade 2

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	703	175.04	175.04	189.42	185.36	14.38	10.31
Asian	33	175.04	184.49	189.42	198.28	14.38	13.79
Black	275	175.04	169.78	189.42	178.57	14.38	8.79
Latino/a	75	175.04	173.78	189.42	183.29	14.38	9.51
White	286	175.04	179.22	189.42	190.51	14.38	11.59
Two or More Races	31	175.04	176.55	189.42	188.82	14.38	12.27
Female	320	175.04	174.80	189.42	185.20	14.38	10.40
Male	383	175.04	175.20	189.42	185.50	14.38	10.30
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup.

Table 9.

2020-21 District MAP Scores Compared with Normative Data – Grade 3

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	706	188.48	185.64	201.08	194.04	12.60	8.40
Asian	30	188.48	202.30	201.08	210.40	12.60	8.10
Black	274	188.48	181.43	201.08	188.01	12.60	6.58
Latino/a	76	188.48	182.12	201.08	191.26	12.60	9.14
White	288	188.48	189.21	201.08	199.04	12.60	9.83
Two or More Races	35	188.48	186.85	201.08	192.13	12.60	5.29
Female	318	188.48	174.80	201.08	193.40	12.60	8.40
Male	388	188.48	175.20	201.08	194.60	12.60	8.40
ELL	41	188.48	176.38	201.08	186.53	12.60	10.15
504	49	188.48	181.05	201.08	191.42	12.60	10.37
SPED	100	188.48	173.26	201.08	180.24	12.60	6.98
TAG	116	188.48	199.71	201.08	208.88	12.60	9.17
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup; ELL=English language learners; 504=students with a 504 plan; SPED=students receiving special education services; TAG=talented and gifted students

Table 10.

2020-21 District MAP Scores Compared with Normative Data – Grade 4

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	706	199.55	195.58	210.51	203.85	10.96	8.27
Asian	30	199.55	203.18	210.51	213.87	10.96	10.68
Black	274	199.55	190.56	210.51	197.55	10.96	6.98
Latino/a	76	199.55	192.02	210.51	199.38	10.96	7.36
White	288	199.55	201.59	210.51	211.35	10.96	9.76
Two or More Races	35	199.55	193.90	210.51	201.76	10.96	7.86
Female	318	199.55	195.30	210.51	204.10	10.96	8.80
Male	388	199.55	195.90	210.51	203.60	10.96	7.70
ELL	41	199.55	185.62	210.51	191.04	10.96	5.42
504	49	199.55	193.83	210.51	200.39	10.96	6.57
SPED	100	199.55	177.46	210.51	181.52	10.96	4.07
TAG	116	199.55	209.67	210.51	219.69	10.96	10.02
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup; ELL=English language learners; 504=students with a 504 plan; SPED=students receiving special education services; TAG=talented and gifted students

Table 11.

2020-21 District MAP Scores Compared with Normative Data – Grade 5

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	610	214.75	209.33	222.88	212.81	8.13	3.48
Asian	11	214.75	224.45	222.88	237.65	8.13	13.20
Black	247	214.75	204.82	222.88	206.79	8.13	1.97
Latino/a	66	214.75	207.41	222.88	211.32	8.13	3.92
White	253	214.75	213.39	222.88	217.81	8.13	4.42
Two or More Races	29	214.75	212.40	222.88	217.49	8.13	5.09
Female	302	214.75	210.40	222.88	216.20	8.13	5.80
Male	308	214.75	208.30	222.88	213.30	8.13	5.00
ELL	26	214.75	197.20	222.88	201.69	8.13	4.49
504	59	214.75	205.49	222.88	209.44	8.13	3.95
SPED	71	214.75	191.00	222.88	191.41	8.13	0.42
TAG	152	214.75	221.31	222.88	229.32	8.13	8.01
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup; ELL=English language learners; 504=students with a 504 plan; SPED=students receiving special education services; TAG=talented and gifted students

Table 12.

2020-21 District MAP Scores Compared with Normative Data – Grade 6

	N	Fall Normed	Fall Mean	Spring Normed	Spring Mean	Growth Normed	Growth Mean
All Students	731	209.13	205.16	218.75	209.70	10.96	4.53
Asian	26	209.13	224.75	218.75	230.07	10.96	5.32
Black	299	209.13	200.16	218.75	203.11	10.96	2.95
Latino/a	87	209.13	202.85	218.75	207.80	10.96	4.95
White	283	209.13	209.50	218.75	215.53	10.96	6.03
Two or More Races	34	209.13	203.73	218.75	208.34	10.96	4.61
Female	352	209.13	205.30	218.75	209.90	10.96	4.60
Male	379	209.13	205.10	218.75	209.60	10.96	4.50
ELL	34	209.13	192.62	218.75	199.24	10.96	6.62
504	79	209.13	204.73	218.75	208.71	10.96	3.98
SPED	96	209.13	187.10	218.75	190.53	10.96	3.44
TAG	173	209.13	218.47	218.75	224.73	10.96	6.26
KEY	More than 1 SD below Norm		Less than 1 SD below Norm		Above norm by less than 1 SD		Above norm by more than 1 SD

Note: Fall and spring scores reflect mean scores overall and for each subgroup; ELL=English language learners; 504=students with a 504 plan; SPED=students receiving special education services; TAG=talented and gifted students.

Qualitative Data Analysis for *Everyday Mathematics*

As part of understanding the implementation process, focus groups (both in person and over Zoom), interviews, and observations were conducted with K-6 teachers in the participating district. Teachers talked about aspects of the program they liked, including the support materials, the flow of the lessons, and other aspects of the program they liked. In addition, they gave constructive and useful feedback about the technical (software) aspects of the program and the Redbird curriculum tool that fell short of meeting their teaching needs.

Note: The following responses were gathered during 3/14/21 Zoom meetings with participants. The evaluator met with teachers at six elementary schools and asked them: "What do you like about Everyday Math?" and "What issues with Everyday Math are a barrier for you?" as well as "Can you tell me about your experience with training provided by McGraw Hill?"

1. Supports

Many teachers were very positive about Everyday Math- Some had used it in other schools and had positive experience there, "I had no idea what I was walking into- by the time I had gotten them in fifth grade- they were teaching me stuff. You really have to give it (the program) three or four years."

1. Teachers also commented that students enjoyed completing small accomplishments along the way: "They're saying 'we did this and we did this' and are very excited about moving through the lessons."

Teachers had many positive things to say about the program:

1. "I really like Everyday math this one of my top two math programs that I've taught."

2. "I really like the program because it lines up the skills the five domains in math every day math is more geared towards each individual standard."

3. I had to go in and learn all the pieces- if in the beginning here's the PD- I got the same

4. "I love that they have the books online- it goes to the exact same lesson where they're actually using the information from the books."

5. "I like that math boxes that are included ."

6. "I like the math boxes also- I like them as a review- I like the ability to score."

7. "I do love the teacher's guide."

8. "I do love how they use the manipulatives."

9. "It's set up nicely it has a lot of manipulatives"

10. "The games are good- with COVID it's hard to do the games I was familiar they're not going to realize how important- the kids are having fun- they have no idea- we learned it and then when were done with the unit- they're figuring out how to work- "we had to do fractions" kids were able to name the skills they had used."

11. "Lessons flow well- takes 40 minutes- starting with the math message- I think it works well- they love the Everyday math games."

12. "The warmup and the focus then the practice. With those three steps I was more able to make it my own so I know the things that we like to cover- math boxes support what we're doing it together."

2. Challenges

Teachers had wide-ranging and insightful comments about issues that they had encountered using Everyday Math. Some teachers spoke about the difficulty of starting a new curriculum in such a challenging (global pandemic) time, others said there were structural issues in how the parts of the lessons interacted with each other, while still others said they were

just beginning their teaching careers and didn't really feel they were getting the most out or the

program.

1. "All the information is overwhelming- this is the first time I've done second grade for a long time- but there's so much information I feel like I can't keep up."

2. "It's just change- there's so much changes and it's so stressful to have this computer based system."

Some teachers reported there was a problem because while differing parts of the materials used the same title for lessons, they were different lessons: "There is one lesson that I did last week, we did another assignment and then I noticed there was a page that was titled the same in Math Masters- the kids said, 'I don't understand it this doesn't make sense' and the lesson did not work out because it was not the same content."

Teachers felt their lessons were limited by the global pandemic:

1. "We watched a video of a teacher doing a lesson kids were in tables kids could share things

2. "They're all spread apart- we don't share anything- we break out rooms with distancing. So I went to the Dollar Tree and bought cards for them they were $7.- I wish we could get decks for all my classes."

3. "I really like Homelink and I'm glad that it comes in Spanish too- lots of kids don't have parent input that the program often requires. I make sure I don't take grades for it (extra work at home)- I think it's a lot to ask for some of my families."

4. "For all of this (the games and extra exercises) to work, we'll actually to be face to face- I have to spread things around on the table between groups and have enough materials for everyone."

5. "I'm having trouble with the games- finding time to teach how to do it first- how do we do that- I don't know how to do it better."

6. "A lot of the program have to do with games- they can't interact with each other like that- they can't play the same board games-it makes it hard to do game based lessons when so much time is spent just setting it up."

3. Redbird

The single biggest negative issue in the use of Everyday math was the Redbird component. Teachers complained that it wasn't well-tied to the lessons of the day and also meant that students were "playing games instead of doing what I want them to do":

1. "I think it's great except for Redbird. When it was sold to us it was something that I was really excited about that's not what it turned out to be."

2. "I have the ability to show exactly what they're doing in Red Bird- that part's great- but it would be great to have a setting to keep them from playing the games."

3. "Redbird is not good- nowhere near as user-friendly (as the rest of the curriculum)."

4. "My main frustration is Redbird. It won't automatically talk to the kids. Sometimes it works on touchscreen and sometimes it doesn't and the questions ramp up too fast. It needs to be like Lexia."

5. "I need it (Redbird) to be independent practice but I need to have something that shows them (through examples) rather than just instructing them on what to do."

6. "I don't necessarily have access playing with Redbird- I don't know if it's just not user friendly or what."

5. Pedagogical Issues

One of the largest adjustments for teachers in PSSD initial use of Everyday Math has been adjusting to a "spiral" or cascading form of concept delivery rather than a more traditional approach of "mastery before moving on".

1. "The content with the math message- it's like a bell ringer- is supposed to be connected to the previous lessons. Lots of times it's something we haven't even go over yet. Sometimes I will just skip that part or substitute something else."

2. "The biggest complaint that I have and a lot of teachers I have was jumping around. We're finding it's not working. If I touched on something like time, the next day would touch a completely different subject. Sometimes it will be ten lessons until they see it again- we're wanting something that we're able to teach it and keep practicing and reinforcing what they're doing."

3. "Kids are forgetting- they're not getting reinforced quickly- mastery vs. spiral."

4. "Traditionally units are set- they're about specific things. We do them until I know kids understand and them I move on. This doesn't work like that."

5. "With Everyday Math I'm worried to get the skills in- it's fun I just worry that they're getting everything that they need they didn't know how to write what they're learning."

6. "The spiral is just weird for me- we do a lesson one time and then they do it on the progress checks not completing then they just complete the skill once - no way to do a progress check with games."

6. Software Issues

The single biggest concern from teachers using Everyday Math related to software was the lack of integration with the district's record-keeping software, Schoology. Teachers also commented

that the use of touchscreen responses was inconsistent within the program, and that the learning

curve for accessing what they needed to do was sometimes high for their students:

1. "When I'm in the teacher's edition sometimes i can't tell beforehand what their screens will look like- I had to project they couldn't see this and they couldn't that- sometimes I'm thinking the toolbox should be there but it's not."

2. "I am a first year teacher- within the school I got introduced to Everyday math- I have had a lot of issues with the software. Everything is a new window- If I say, "let's play with pennies- and they I can't find where the lesson - I can't teach them how to do it- I copy and paste into Cammy (?) I basically have stopped try to use the actual program during class."

3. "Some of my class can do it on their own-from it needs more to be kid friendly all touchscreens need to work the same way. I would like something to tell me more clearly if students miss a problem a certain amount of times- it's real disjointed- it requires a lot of teacher attention."

Note: The following responses were gathered during a site visit to Little Rock on May 10 and 11, 2021. The evaluator met with the staff of ten elementary schools (approximately 60 teachers in total) and asked them (among other questions): "What are the most important aspects of Everyday Mathematics?" and "What issues with Everyday Mathematics are a barrier for you?" as well as "Can you tell me about your experience with support provided by McGraw Hill?" Because this was the last of several interviews with the teachers, I was able to follow up with questions about past issues, particularly pacing, Red Bird, and Everyday Mathematics/Schoology integration.

1. Supports

Many teachers were very positive about Everyday Mathematics. They were pleased with the

materials provided and reported that their students liked many aspects of Everyday Mathematics

as well, including the games:

1. "In the kit there a lot of flash cards and other useful materials- I really appreciated that."

2. "I really like the quality of materials in the kit- a lot of good stuff for teachers and students."

3. "I like that I can teach math and they can work on it as I do it in real time."

4. "I do like the Math Boxes review part."

5. "I like that I can check the Home Links page and that it actually explains to the parents what we are doing on this sheet so that the parents can help us keep their children on pace."

6. "My kids grew from their winter (NWEA) scores- winter to spring was a big improvement."

7. "My students really love the games- we've made some games ourselves."

8. "I like that the components are online and so we could use them whether a student is virtual or not."

9. "This is actually more on the child development side- we're going to draw pictures as we understand things, especially in kindergarten. I like that the children are given the space to do that in the workbooks."

10. "A lot of these lessons had amazing games to play, and the kids really loved them."

11. "Their math scores have shot up- as the NWEA gets closer we do a lot of math talk- we talk about how we use math all the time. They have learned to draw pictures and write a sentence as other ways of showing they know the answer."

12. "After the first NWEA test they grew- they're motivated to do well."

13. "My kids are actually learning this stuff- there are some lessons that they had really gotten into. It's a much different experience than what we have had in the past."

14. "I like the games for sure."

15. "The warmup portion is excellent- the five to six minutes we spend doing that makes the lessons more meaningful for the students."

16. "I like the digital component of it- it can be shown on the computer the same way I would show it on the whiteboard."

17. "Using manipulatives are a strength. It takes twice as long to teach using them, but that will be better next year and as we go along."

18. "Trainings are good- they made it easier to use the curriculum and lets us know what is possible."

19. "I like the layout of the books- it's logical and easy for the children to use."

20. "The kindergartners bridged that digital gap and these kids have learned to create a community."

2. Challenges

During the interviews, a number of aspects of the Everyday Mathematics that were less than optimal were discussed. Teachers had wide-ranging and insightful comments about issues that they had encountered using Everyday Mathematics.

1. "The first entire semester we just have journaling (blank) pages (in our workbooks), but now we finally have something to do. I was having to supplement the first semester doing it all virtually. They didn't use it (the workbook) during either of the first grading periods."

2. "I found that what the district was providing- packs of cards, counters, and math cubes- wasn't adequate to meet my needs. I've spent a lot of money at the Dollar Store this year."

3. "One of the biggest challenges has been learning the program itself. I know there's a lot available, but I don't really know how to create assignments or lots of other things. I need more in-depth training."

4. "I did not use their manipulatives while I was teaching online."

5. "My kindergartners had to learn all the technology to even use the program. Keyboarding was very difficult for some of them."

6. "Mine found the games they'd rather do than their assignments."

7. "My first graders had a hard time learning how to click and drag more than anything."

3. Pandemic/Distance Learning Challenges

Because of the global pandemic, this was an unprecedented year in American education. Participating teachers found the pandemic impacted their experience of teaching and delivering the curriculum:

1. "We have kids coming that lost at least a half a year of math- it's going to be a challenge for a while."

2. "One of the great things about the program are all games. This not a great year to try to do curriculum- because we couldn't do the partner games-pull out the cards we couldn't do that because of social distancing.

3. "Even if we put them in breakout rooms they couldn't really do learning groups the way they do them in person- it really hurt us this year."

4. "I can't wait for them to actually have a normal year and all be back here so we can do this the right way."

5. "This year a lot of the setup for the lesson- partner lesson, partner games, and so on were not possible because of the number of distance learning kids and the fact that we were using new curriculum."

6. "My experience is that we're not going to finish the curriculum- they missed a chunk of the third grade."

4. Redbird

1. *Although interviews done earlier this year indicated that teachers were unhappy with the Redbird aspect of the curriculum, the last focus groups indicated teachers were more mixed in their use and acceptance of the software.*

2. "I have used Redbird more because it's prescribed in the texts. I have found that some of my students who are low-performing may not get the lesson or they may do poorly on something else, but they use and like Red Bird. It seems to help some students that have been extremely low."

3. "I haven't had any issues with it. It works very well, especially with very low kids."

4. "I use it with fidelity- so I don't have a problem with Redbird- I could use an online lesson but the kids like to have that other interesting thing to do (Redbird)."

Some teachers found that Redbird was very difficult to use because the adaptive/AI nature of it made it difficult to understand what different students in the classroom were actually doing in terms of the curriculum and the daily lesson:

1. "My point is it's hard to pinpoint where they are (in Redbird). It's not user friendly when they're trying to look at data. It's too hard to move them forward. You don't know how the AI is going to work. When there is a mismatch with the assignment (not directly

linked to the lesson) the kids got really bored. Just learning how to use it was difficult for some kids."

2. "Redbird does not come across as a reward for the students who do it-even the games kids are presented with aren't entertaining and the kids don't like it."

5. Pedagogical Issues

Teachers were divided when asked about various curricular aspects of the program. Teachers worried that the cascading nature (exposure and then coming back for more enrichment later) confused some of their children and they were frustrated by it. Two examples of the "one lesson" exposure that frustrated teachers the most were long division and telling time. Others wished that assessments in the curriculum were given more frequently (especially in the case of summative unit assessments) and it was also noted that the end of the year summative exams were too long, especially for younger grades.

1. "When they get exposed more closely to the subjects, they really learn it. I'm frustrated by the brief exposures."

2. "I would really like to have more time given to the first exposure to a concept."

3. "The first exposure ever to a subject should be given more than one lesson. I can't believe long division popped up for a day and then disappeared."

4. "What I have seen is they're getting exposed briefly over and over to something means it was like I was reteaching same concepts more than once."

5. "Please- more than one day introducing concepts."

6. "I taught counting time- this lesson was a one day lesson- and then the kids didn't see it again for four or five weeks."

However, other teachers saw the cascading as a benefit and said that their students were starting to see the relationship between different curricular concepts through exposure.

1. "I think in the upper grades the cascade is very useful to help students learn."

2. "I feel like I was a little bit pleasantly surprised. I think that the showing it to them, the cascade method, is working. In fractions it took me a little while to get onboard."

3. "I do like the cascading."

4. "I do like how we introduce things and then come back to them."

5. "The more the students use *Everyday Mathematics* the more they will see the spiral has happened and how the concepts are related."

6. "The summative exams are too long- we need more unit tests more often- the kids thought there would be a very difficult test based on the long number of pages."

7. "The (summative) assessments take too long at 45 minutes per student. Can't they break up the summative assessments?"

6. Pacing of Material

Although teachers had been concerned about the fast pace of the lessons in earlier interviews, some seemed less concerned during these focus groups:

1. "I don't think the kids were struggling with the pace of the lessons at the beginning of the year implementation. I actually think the pacing is very slow, and thankfully very early childhood education focused."

2. "They missed a good part of last year- hard to know about the pacing. It did seem fast to them, but is that fast compared to a normal year? I don't know."

3. "The pacing was okay- they get exposed to the routine- it's organized differently than what people are used to in previous years."

4. "Pacing was extremely fast. These kids missed half of third grade last year so we've been playing catch-up all year."

5. "The pace for us was way too quick- we're not going to finish."

7. State Testing

Teachers were asked how well they thought Everyday Mathematics aligns with district and state testing and benchmark goals as well as how well they thought it prepared students for state testing. Teachers were mixed in their opinions:

1. "Especially at the beginning of the year the early stuff is not well aligned with district goals."

2. "I think that it does because there are some assignments that they really talk about the state testing. For example, Mental Math- a lot of those questions you need to use your mind and it (the curriculum) really did help them. A lot of the lessons you could tell were directly aligned with state standards."

3. "We're using all the manipulatives here, which is great, but the manipulatives do not really prepare them for what they need to do on the state tests."

4. "There's a responsibility to meet state testing requirements and I don't think there's enough there (in the curriculum) for my kids to do well on the testing."

5. "Actually, I think it is pretty well aligned. I just used what I knew could be helpful to my students and didn't worry about the rest."

8. Suggestions for Changes

Finally, teachers were asked "If there is one thing that you could change in the Everyday Mathematics curricula, what would it be?" Their answers reflected a number of different aspects of the curriculum as well as suggestions about assessments and materials:

1. "I would say give us something at the beginning of the year other than blank pages in their journals."

2. "I like the journals that are provided for the students in their workbooks, but at the beginning of the year it needs to be something that the kids need to do, work, in the notebook format- there's a huge gap between early childhood and elementary belief systems. Not everything needs to be measurable. We're working (on the kindergarten level) with the kids more at their own pace, but here (*Everyday Mathematics*) they're expected to sit down and crank it out."

3. "Maybe somehow slow the curriculum down and revamp Red Bird. For me it's hard to know exactly what to change because of the (pandemic) year."

4. "The curriculum to be more linear for earlier grades and for kids who are low."

5. "I would like there to be paperwork for the kids to do for the routine everyday assignments."

6. "The toolkit needs to be beefed up for more teacher flexibility."

7. "Sometimes it's hard to think of a unit having sustainability for four weeks- maybe they should break it up a little bit."

8. "I would like to have two volumes of the math journal- breaking it up would make it easier for the students."

9. "End of the year assessment is something that seems very long. And it's worse because even though there are seven problems, they're put on separate sheets of paper so it looks very scary to my students."

10. "I would like the material to be more consistent with how the online assessments are graded."

11. "I wish that we had automatic notifications of missed assignments and where the students

    are. I wish teachers have more control- the option to go in there and reset them to a

    proper grade- If I go back and reset them the next time I look they haven't saved. It's

    very frustrating."

Summary

1. The implementation of *Everyday Mathematics* during one of the most difficult teaching years in American history produced a wide range of results, both for teachers and students. While teachers liked many aspects of the program, and were grateful for the support and training provided by McGraw Hill, as can be expected in the first year of any curricular implantation they found that some parts of the program (structure of lessons, supporting materials, online games) worked better than others (software interfaces, glitches with students working remotely, a new way of presenting concepts to students-cascade rather than linear, and the Redbird program). Teachers were forthcoming and constructive in their responses and were interested in improving the program and their use of it.

2. Overall, spring assessment scores were significantly greater than fall assessment scores across grade levels and both years of data analyzed. In almost all cases, effect sizes were substantial. The one exception was for grades 7-10 in 2020-21 where all findings were statistically, but not practically significant.

3. Overall, students earned scores on the fall and spring assessments that were below the normative data. Exceptions to this can be found on the 2020-21 assessments for grades K-2.

4. Asian American students outperformed the normative data for almost all fall and spring assessment comparisons.

5. Across almost all comparisons for both years, *Everyday Mathematics* students' growth scores from fall-to-spring fell short of the NWEA normative data. Exceptions were found in a few subgroups.

References

[District Name]. (2020, July 9). *[District Name] roadmap to reopening 2020-2021.* Updated

December 4, 2020. https://core-

docs.s3.amazonaws.com/documents/asset/uploaded_file/796210/Roadmap_to_Reopening

_Updated_12.8.20.pdf

Marshall, D. T., Shannon, D. M., & Love, S. M. (2020). How teachers experienced the COVID-

19 transition to remote instruction. *Phi Delta Kappan, 102*(3), 46-50.

https://doi.org/10.1177/0031721720970702

Marshall, D. T., & Bradley-Dorsey, M. (2022). Reopening schools in the United States. In D. T.

Marshall (Ed.), *COVID-19 and the classroom: How schools navigated the great*

disruption (pp. 147-164). Lexington Books.

Northwest Evaluation Association. (2015). *2015 NWEA Measures of Academic Progress*

normative data. Portland, OR: Author.

Northwest Evaluation Association. (2020). *2020 NWEA MAP growth normative data overview.*

Portland, OR: Author.