Evaluation of Beast Academy in Mankato Area Public Schools: Summary Memo

Introduction

WestEd was contracted by Art of Problem Solving (AoPS) to conduct an independent evaluation of its Beast Academy program in Mankato Area Public Schools (MAPS), a school district serving approximately 8,500 K-12 students in Mankato, Minnesota. In MAPS, Beast Academy (BA) is traditionally used in the elementary and middle grades to support two groups of students: gifted and talented students ("Cluster" students), and high ability learners from populations that have been historically marginalized in programming for academically talented students ("Rising Scholars"). Rising Scholars includes high ability students who are culturally diverse, economically diverse, linguistically diverse, and twice exceptional students.

WestEd acknowledges that this study occurred during the COVID-19 pandemic. The impact of the pandemic on students' learning environments and on the district's use of Beast Academy cannot be overstated. Although WestEd maintained a rigorous design and supplemented study-collected data about Beast Academy exposure with data from the district and AoPS to ensure fidelity of treatment/comparison designation, we nonetheless advise caution in interpreting the results given these data were collected during an unusual time in students' education.

The purpose of this memo is to provide an easily accessible summary of the evaluation findings. The technical details of the study, including methodology, treatment determination, analysis, and results, can be found in the technical report (Beast Academy Evaluation Technical Report). The Technical Report contains critical information pertaining to the study, methods, data analysis, and interpretation of the results and is meant to be read as a supplement to this Summary Memo. This Memo is intended to provide readers with streamlined access to study highlights and findings.

Evaluation Approach

WestEd's evaluation estimated the impact of exposure to two years of Beast Academy programming on student achievement in mathematics in the Mankato Area Public Schools district. All students enrolled in grades 3, 4, and 5 in school year (SY) 2019–20 who remained enrolled in the district through grades 4, 5, and 6 in SY 2020–21 were eligible to participate in the study. The evaluation also examined the math-based attitudinal outcomes associated with Beast Academy participation. To estimate the impact of Beast Academy on achievement, WestEd analyzed baseline and outcome data from the NWEA Measures of Academic Progress (MAP) interim assessments in

mathematics administered by the district.[1] The evaluation team supplemented this analysis with student surveys that captured students' attitudes and perceptions related to math in the 2020–21 SY, such as their perceived sense of self-efficacy to learn and do difficult math work.

In each analysis, WestEd sought to determine the extent to which math performance or attitudinal patterns differed between students who used the Beast Academy program and their peers who were not exposed to the program. Usage was defined in several ways. First, WestEd evaluators received anonymized programmatic data from the district that identified students as either Cluster or Rising Scholar Beast Academy participants. To obtain a more nuanced definition of treatment, the evaluation team also received detailed usage data from the district, which contained student-level information on the number of lessons attempted and time spent doing math-related work on the Beast Academy Online platform. Together, these data allowed WestEd to estimate the impact of Beast Academy for students who used the program in greater or lesser "dosages" over the 2019–20 and 2020–21 school years. Students were identified as Beast Academy users in the survey analysis based on a self-reported measure of participation in Beast Academy over the one-year period of SY 2020–21.

WestEd's analytic methods included statistical modeling approaches that isolated the effect of Beast Academy usage on the outcomes while controlling for other factors that could be related to the results. Where appropriate, the evaluation team disaggregated results by grade level and treatment group (i.e., Cluster or Rising Scholar). For a complete discussion of the evaluation design and methodology used, including the matching procedures referenced below, see the Technical Report that supplements this memo.

Evaluation Findings

Student Achievement Results

From a sample of 1,027 students enrolled in grades 2 through 4 in SY 2018–19 (the SY prior to the study intervention period), WestEd used a rigorous matched comparison group design to pair Beast Academy students with non-Beast Academy students who shared similar characteristics at the start of the study, including their baseline mathematics achievement levels. The analysis compared the math outcomes of both groups two years later to determine the extent to which any observed differences between the two could be attributed to exposure to the Beast Academy program.

---

[1] WestEd was originally contracted to evaluate the impact of Beast Academy over a one-year period using state assessment data. Due to the disruption in statewide testing as a result of the COVID-19 pandemic, the evaluation plan was adjusted to accommodate a two-year study period using the NWEA assessment data.

WestEd | Improving learning, healthy development, and equity in schools and communities

The evaluation team found that Beast Academy had statistically significant, positive effects on mathematics achievement. Specifically, students who had a record of Beast Academy use in both years of the intervention period scored on average 8.8 scale points higher on the spring 2021 MAP assessment in mathematics than their matched comparison-group peers who had no exposure to Beast Academy during that period.

The impact of Beast Academy on mathematics achievement varied by treatment group, with the largest effects seen for Rising Scholar students. Students identified in district administrative records as Rising Scholars scored on average 11.2 scale points higher than the comparison group. Students identified as Cluster students scored 10.1 scale points higher than the comparison group. A third treatment group, students who were not identified as Rising Scholar or Cluster students but who had a record of Beast Academy use in the two intervention years, scored on average 6.7 scale points higher than the comparison group.

The impact also varied by the level of engagement on the Beast Academy Online platform, as measured by the number of lessons attempted over the two-year intervention period. The higher the number of lessons attempted, the larger the gap in math achievement compared to students who had no exposure to Beast Academy Online. For the average Beast Academy user in Mankato Area Public Schools—one who attempted 120 lessons on the platform over two years—this effect was equivalent to a 4.9-point increase in math scale scores over the average score for the comparison group. Attempting 120 lessons on the platform was associated with a math score of 224.4, about 5 points higher than the comparison group average of 219.5. Figure 1, below, illustrates the upward trend in math scores associated with discrete levels of Beast Academy Online use.

WestEd | Improving learning, healthy development, and equity in schools and communities

![WestEd logo](WestEd.org)

Figure 1. Higher levels of Beast Academy Online use were associated with larger differences in mathematics performance between Beast Academy students and their matched comparison group peers who did not use the program
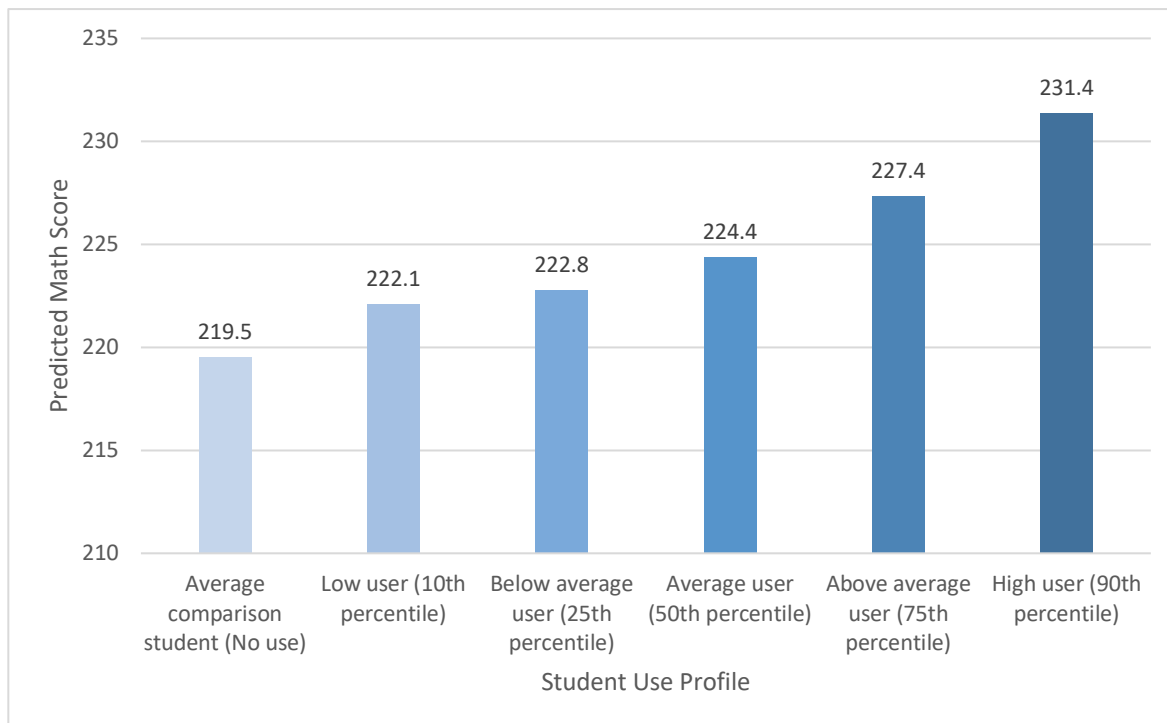
Note: Results based on a weighted sample of 1,094 MAPS students enrolled in grades 4, 5, and 6 in SY 2020–21. "Predicted math score" refers to the expected MAP scale score associated with each level of student use on the online platform. Student use profiles calculated based on the cumulative number of lessons attempted on the platform over SYs 2019–20 and 2020–21.

Student Survey Results

In addition to analyzing student achievement patterns, in the spring of 2021 WestEd surveyed all MAPS students who were enrolled in grades 4 through 6 during the 2020–21 school year. The survey assessed students' attitudes and perceptions toward math on eight different domains, using established measures validated through prior research: (1) motivation to study; (2) effort and perseverance; (3) interest in mathematics; (4) cooperative learning; (5) self-concept in mathematics; (6) growth mindset; (7) self-efficacy in math; and (8) beliefs about mathematics as a learnable subject.[2]

---

[2] See the Technical Report for details on the sources and properties of these measures.

WestEd | Improving learning, healthy development, and equity in schools and communities

WestEd received 1,061 usable responses to the survey. This number represented the sample of students whose parents consented to their students' participation, who themselves assented to participate in the study, who themselves completed the survey, and whose treatment status could be confirmed based on a self-reported measure of Beast Academy use in the 2020–21 school year. The evaluation team compared the average responses of students who reported using Beast Academy with the average responses of students who did not use the program. WestEd's estimation method compared Beast Academy students to their non-Beast Academy peers within the same school, to control for unobserved differences in school settings that may have influenced their attitudes and perceptions toward math.

The evaluation team found that Beast Academy students had more positive attitudes and perceptions toward math on six of the eight domains measured. The findings listed below were statistically significant. WestEd's analysis revealed the following patterns:

- *Motivation to study.* Beast Academy students scored .24 scale points higher, on average, than their non-Beast Academy peers on the 4-point scale for this measure. The average Beast Academy student score was 3.18, compared to an average of 2.94 among non-Beast Academy students. In practical terms, this result means that Beast Academy students expressed higher levels of agreement with statements related to their motivation and persistence to do difficult math work, such as, *"When I don't understand a problem, I keep working until I find an answer."*

- *Effort and perseverance.* Beast Academy students had an average score of 3.57 compared to the non-Beast Academy average of 3.28, a difference of .29 points on the 4-point scale for this measure, meaning Beast Academy students reported engaging in effortful math work at greater frequencies than non-Beast Academy students.

- *Interest in mathematics.* Beast Academy students had an average score of 3.06 compared to the non-Beast Academy average of 2.76, a difference of .30 points on the 4-point scale for this measure, meaning Beast Academy students expressed higher levels of agreement with statements related to their interest in math.

- *Self-concept in mathematics.* Beast Academy students had an average score of 3.50 compared to the non-Beast Academy average of 2.79, a difference of .71 on the 4-point scale for this measure, meaning Beast Academy students had higher appraisals of their own abilities in math.

- *Self-efficacy in math.* Beast Academy students had an average score of 2.94 compared to the non-Beast Academy average of 2.47, a difference of .47 points on the 5-point scale for this measure, meaning Beast Academy students expressed higher levels of confidence in their ability to learn and do difficult math work.

WestEd | Improving learning, healthy development, and equity in schools and communities

- *Beliefs about mathematics as a learnable subject.* Beast Academy students had an average score of 5.33 compared to the non-Beast Academy average of 5.07, a difference of .26 points on the 6-point scale for this measure, meaning Beast Academy students were more likely to believe that math could be learned as opposed to the belief that math is only accessible to high-ability students.

- No significant differences were found between Beast Academy and non-Beast Academy students on the scale scores for *cooperative learning* and *growth mindset*.[3]

- These patterns were consistent when disaggregating the data by grade level.

Beast Academy students were also more likely than non-Beast Academy students to agree with positive statements related to their interest in and enjoyment of math in the current school year. Beast Academy students had higher average responses to the items, *"Math is interesting this year"* (3.97 versus 3.62, respectively), and *"I enjoy math this year"* (4.01 versus 3.45, respectively), compared to their non-Beast Academy peers who were enrolled in their same schools. These items were measured on a 5-point scale, where 1 indicated *Totally untrue* and 5 indicated *Totally true*.

Additionally, WestEd examined open-ended responses to the questions, *"What is your favorite thing about Beast Academy?"* and *"What is your least favorite thing about Beast Academy?"* **The predominant theme from these responses was that Beast Academy was more challenging, and often more engaging, than students' typical mathematics work.** Among the key themes observed by WestEd evaluators were the following:

- Students noted that they enjoyed being challenged on the platform, and that these challenges often made math work feel fun, fresh, or new. As one student put it, *"I like Beast Academy because it is difficult and it teaches math in different ways."*

- Students liked the collaborative nature of the program, noting that they enjoyed working with their classmates to solve difficult problems. One student summed up this sentiment by saying, *"My favorite thing about Beast Academy is working as a team and figuring out problems together."*

- Students appreciated the opportunity to try problems again if they did not get the answer right the first time; some further noted how the program encouraged continued effort. As one student put it, *"[Beast Academy] makes me want to keep trying and not give up in math even when it's hard."* Some students wished they could have more opportunities to attempt a problem before seeing the answer or having to start over.

---

[3] Responses to the four items that comprised the growth mindset scale featured high levels of missing data. As a result, the evaluation team considered estimates related to growth mindset to be unreliable. See the Technical Report for details related to this measure.

WestEd | Improving learning, healthy development, and equity in schools and communities

- Most answers to the "least favorite thing about Beast Academy" prompt had to do with how challenging the program was in a negative sense. While many students appreciated that the program pushed them to work through difficult problems, some students expressed frustration at the level of difficulty they encountered, and, less frequently, the lack of clarity in the explanations to problems that they got wrong.

Reflection on the Results

Overall, the evaluation results provide evidence to suggest that Beast Academy positively influenced mathematics achievement and math-related attitudes among the cohort of MAPS students who were enrolled in grades 4, 5, and 6 in the 2020–21 school year. Due to the unprecedented disruptions to learning conditions that occurred during the study period as a result of the COVID-19 pandemic, WestEd urges caution in extrapolating the results beyond this specific population of students. While the evaluation team made concerted efforts to preserve the rigor of the analyses, the study was conducted during an extraordinary time in the educational trajectories of students. Despite the evaluation team's confidence in the findings, there are many potential unmeasured factors that could have affected students' performance on and engagement in math-related activities during this time. For example, the way the district used Beast Academy changed over the course of the study period, and some students may not have received a full year of mathematics instruction from the same teacher. We therefore recommend that the results be viewed as a snapshot of the relationship between Beast Academy use, on the one hand, and math achievement and attitudes, on the other, within the unique context of Mankato Area Public Schools at this time.

Notwithstanding the need to interpret these results cautiously, the findings reflect a positive association between Beast Academy and student outcomes. The strength of the evidence found by the evaluation team is greatest for the impacts on math achievement. Because the student achievement analyses used a rigorous matched comparison group design, the significant differences in math scores between Beast Academy students and their matched peers can be reasonably attributed to participation in the Beast Academy program, while acknowledging the potential confounding factors that may also have been at play. The evaluation team's use of student-level usage data allowed for a more refined measure of student exposure to the program—and, consequently, a more nuanced measure of treatment—which could have improved the precision of the estimates. On the other hand, since the usage data only captured students' engagement with the online portion of the Beast Academy program, it may have missed other important ways in which students engaged with the program, for example, through direct teacher instruction and support.

WestEd | Improving learning, healthy development, and equity in schools and communities

When comparing the results based on the district-provided indicators of Beast Academy exposure with those based on the usage data, similar trends can be seen with respect to Beast Academy students' higher mathematics performance relative to their peers who did not receive the treatment. With the district indicator data, the estimated average effect of Beast Academy exposure was 8.8 scale points higher than the average comparison student; with the usage data, the estimated average effect was a more conservative 4.9 scale points. It is beyond the scope of this evaluation study to unpack the differences between the district indicator of Beast Academy exposure and the usage data gathered through the online platform. Nonetheless, both the district indicator data and the usage data reveal a positive association between Beast Academy and student performance. Given that the usage data provide a more nuanced glimpse into students' actual levels of engagement on the online platform, there is an opportunity to further explore these data to understand how Cluster and Rising Scholar students differ from each other in terms of their usage and achievement patterns. From an implementation perspective, more can also be learned about how teachers use Beast Academy to support the unique learning needs of Cluster and Rising Scholar students.

The student survey results can be interpreted as measures of associations between students' average attitudes and perceptions toward math at the end of the 2020–21 school year and their self-reported usage of Beast Academy at any time during the 2020–21 school year. The largest positive association between Beast Academy exposure and math attitudes was found for the *self-concept in mathematics* measure (.71), followed by *self-efficacy in math* (.47), *interest in mathematics* (.30), *effort and perseverance* (.29), *beliefs about mathematics as a learnable subject* (.26), and *motivation to study* (.24) in order of descending magnitude of differences between the two groups. As with the student achievement analyses, all results were highly statistically significant, meaning it is unlikely that they could have been obtained due to chance. Results from the open-ended survey items largely corroborated the trends from the scale analyses.

WestEd's survey estimation approach controlled for unobserved school-level characteristics, ensuring that no two students were compared across different school contexts. However, the absence of linked pre-test measures of the outcomes meant the evaluation team could not control for preexisting differences in students' attitudes and perceptions toward math. Consequently, results of the survey analysis cannot be exclusively attributed to participation in Beast Academy. The treatment status indicator for the student survey analyses also lacked the nuance of its usage counterpart in the achievement analyses, being as it was a self-reported dichotomous indicator of Beast Academy use. Unlike the usage data, this measure could only indicate whether students had used Beast Academy at any point in the prior school year, even if the extent of that use was limited to a single session or lesson; it did not capture richer patterns related to dosage or length of exposure. Nevertheless, the data show clear patterns of differences between Beast Academy and non-Beast Academy students. Future research could probe the potential causes for these differences, and the mechanisms by which Beast Academy may help promote feelings of self-

concept, self-efficacy, interest, effort, beliefs, and motivation in math. An in-depth qualitative examination of students' experiences using the program, for example, might ask students to share, in their own words, what specific features of the Beast Academy program contribute to such feelings. By the same token, a more rigorous quantitative study, such as one that randomly assigned students to different levels of exposure to the program, could help bring to light the extent to which the differences in outcomes observed in WestEd's survey are driven by specific components of the program or by other unmeasured characteristics of the students themselves.

An equally important question, especially in the context of the COVID-19 disruptions to students' learning environments, concerns which learning conditions are most conducive to maximizing the potential benefits of Beast Academy. The results from the student achievement analyses suggest that higher levels of use on the online platform—at least as measured by the number of lessons attempted—predicted higher math scores among the population of MAPS students examined for this study. It may be reasonable to ask what a realistic number of lessons attempted could be in a given school year, for a given student. To aid in these kinds of discussions, the study identified several student use profiles with details about the predicted mathematics score associated with discrete levels of engagement on the platform (i.e., number of lessons attempted). These and more details are described in the Technical Report.