*Susan Davis-Becker, Ph.D.*

*Kyle Kilworth, M.Ed.*

# Evaluation of Teach to One Roadmaps

March 2024

## Evaluation Overview

New Classrooms contracted with ACS Ventures, LLC (ACS) to evaluate the utility of the assessments within *Teach to One Roadmaps* (*Roadmaps*). The product uses both a diagnostic assessment and formative skill-level assessments of both pre- and on-grade skills in order to provide each student with a personalized learning pathway to proficiency. In this evaluation, ACS explored the relationship between skill-level mastery within *Roadmaps* and state assessments (state) administered to the same students at the end of the school year. ACS framed the evaluation around three questions:

1. What is the relationship between student performance within *Roadmaps* and their performance on the state assessments?

2. Do these relationships vary across subgroups of students within the data (e.g., across states and grades)?

3. Are there types of skills that are more or less predictive of performance on the state assessment?

This report details the data used in this evaluation, the approach to addressing each of the questions listed above, and the interpretation of the results for the *Roadmaps* tool.


## Data Sources

New Classrooms coordinated with three states to collect student performance information on the *Roadmaps* assessment as well as the state assessment. A summary of the data included in this evaluation is in Table 1. Across the three states, data was collected from between 1 and 3 grade levels (spanning grade 5 through grade 9).

Table 1. Student Samples by State and Grade

| State | Grade | # Students |
|---|---|---|
| State 1 | 5 | 100 |
| State 1 | 6 | 84 |
| State 1 | 7 | 121 |
| State 2 | 8 | 44 |
| State 3 | 9 | 201 |

Table 2 provides a high-level summary of the data collected for the students in this study. This included:

- Students' *Roadmaps* performance summarized as the percent of prerequisite and on-grade skills that students have mastered at the end of the school year (i.e., % of skills)
- Students' state assessment results in the form of total exam score (either raw or scaled score) and the performance level associated with the student's total test score (5=Exceeds Expectations, 4=Meets Expectations, 3=Approaches Expectations, 2=Partially Meets Expectations, 1=Did Not Meet Expectations).

Because the state assessment exam scores are reported on different scales, the scores were converted to z scores (position of a score relative to the mean, measured in standard deviation) to facilitate analysis across different grades and states. Additional details on student state assessment performance and distribution of scores can be found in Appendix A.

Table 2. Student Performance Results by State and Grade

| | Roadmaps | | | State Assessment Results | | | | |
|---|---|---|---|---|---|---|---|---|
| Assessment | Min | Max | Average | Min Exam Score | Max Exam Score | Average Exam Score | Min Z score | Max Z score |
| State 1 – Grade 5 | 0% | 93% | 30% | 663 | 778 | 713.91 | -2.27 | 2.02 |
| State 1 – Grade 6 | 15% | 100% | 57% | 671 | 794 | 727.04 | -1.97 | 2.61 |
| State 1 – Grade 7 | 1% | 100% | 56% | 665 | 817 | 729.86 | -2.19 | 3.47 |
| State 2 – Grade 8 | 8% | 100% | 51% | 650 | 795 | 737.09 | -2.88 | 1.92 |
| State 3 – Grade 9 | 1% | 100% | 47% | 12 | 85 | 53.93 | -2.10 | 1.55 |

The strength of this dataset is that it represents a sampling of students from different school districts and thus provides a comparison of the *Roadmaps* results to those from three state assessments. However, some samples (e.g., State 2 grade 8, State 1 grade 6) were smaller and therefore, interpretation of results based solely on the students in those samples should be limited.

## Analysis and Results

### Relationship between performance on Roadmaps and the state assessments

To begin this investigation, ACS combined the data across states / grades to evaluate the relationship between performance on the *Roadmaps* and the state assessments. To start, *Roadmaps* scores were correlated with the state assessment Z scores. The overall correlation was 0.795, indicating a strong relationship between performance on these two measures. The full relationship is shown in Figure 1.

The relationship between performance on these two measures was further explored by comparing the *Roadmaps* performance to the performance level identified for each student based on their state assessment score. The results of this comparison are shown in Figure 2. For this visualization, performance on the *Roadmaps* assessment was categorized into one of five groups (0-12.5%, 12.5-25%, 25-50%, 50-75%, 75-100% of skills mastered). As shown in these results, the lowest state performance levels (e.g., Level 1 and Level 2) had the most students who mastered only a small proportion of *Roadmaps* skills whereas students in the higher state assessment performance levels (Levels 4 and 5) had students with more skills mastered.

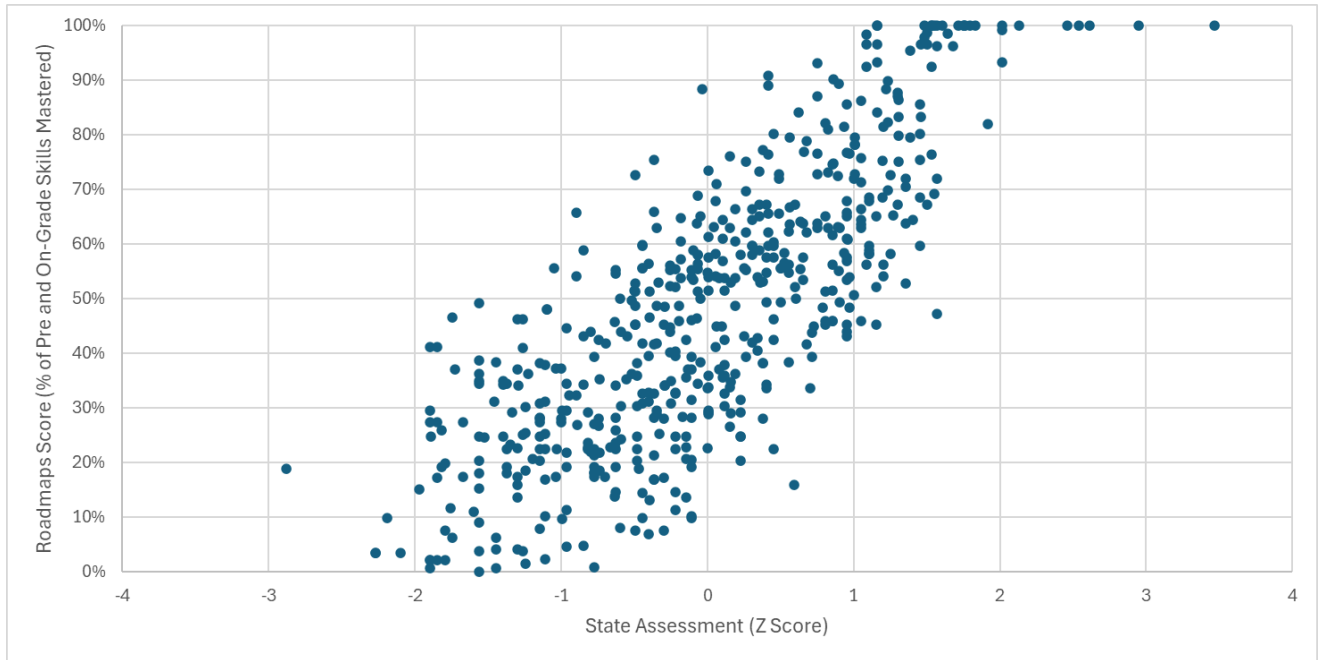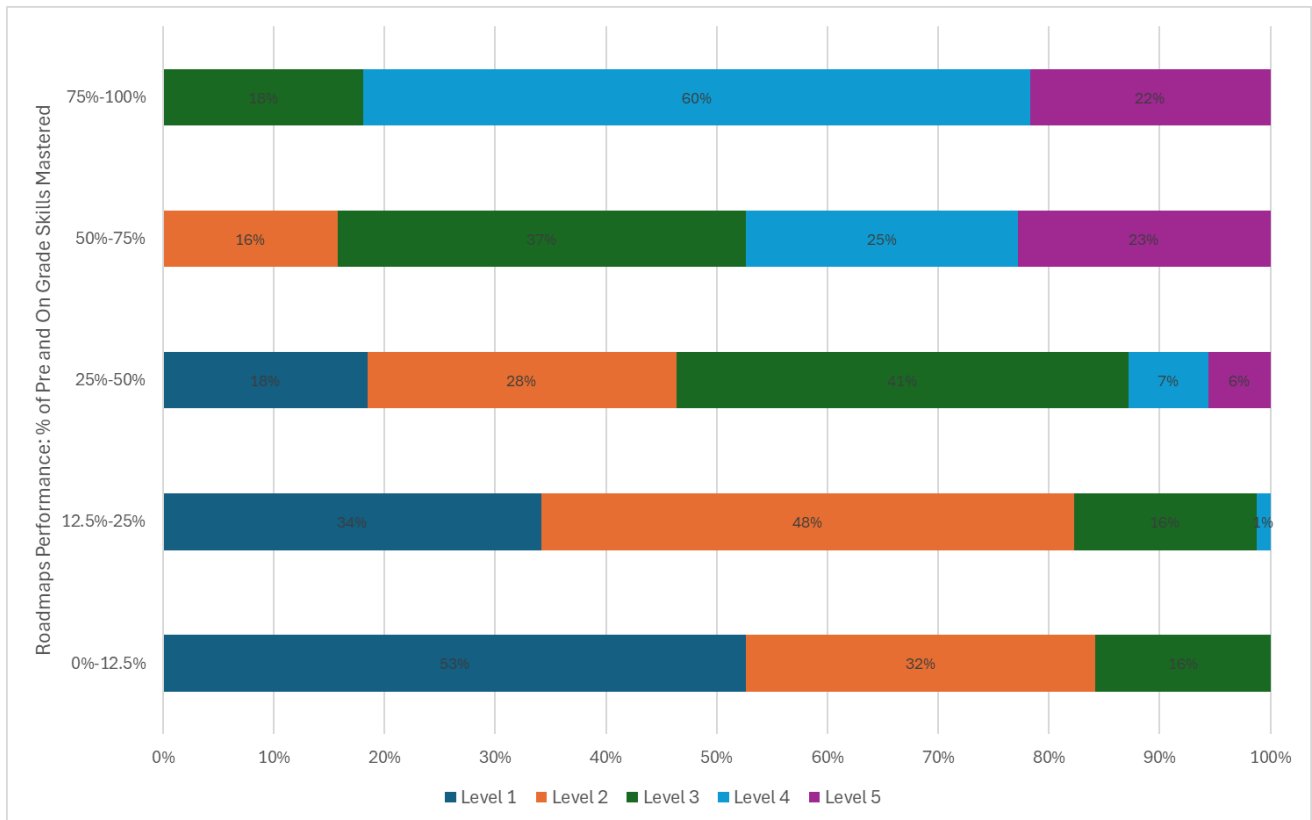Figure 1. Relationship Between Roadmaps and State Assessment Results - Scores



Figure 2. Relationship Between Roadmaps and State Assessment Results – Performance Levels

The predictive utility of the *Roadmaps* results to the state assessment exam scores was evaluated using simple linear regression. The overall regression was statistically significant ($R^2$ = 0.5769, $R^2$ (adjusted) = 0.5738, F=185.8, p < .0001) meaning that the *Roadmaps* results were a significant predictor of performance on the state assessments.

## Relationship Across States and Grades

The strength of the relationship between the *Roadmaps* scores and the state assessment scores was further evaluated for the different states and grades represented in this dataset. Table 3 shows the correlations between these two sets of scores by state / grade and Figures 3 and 4 show the trend of this relationship across scores for each grade and state. Overall, these results confirm that the strength of the relationship is consistent across grades and states. The lowest correlation is for the 8th grade sample from State 2 which may be due to the small sample size.

Table 3. Correlation Between Roadmaps and State Assessment Results by State / Grade

| Grade | State | Students | Correlation |
|-------|-------|----------|-------------|
| 5 | State 1 | 100 | 0.780 |
| 6 | State 1 | 84* | 0.867 |
| 7 | State 1 | 121 | 0.852 |
| 8 | State 2 | 44* | 0.715 |
| 9 | State 3 | 201 | 0.760 |

**\*** Results based on N<100 should be interpreted with caution

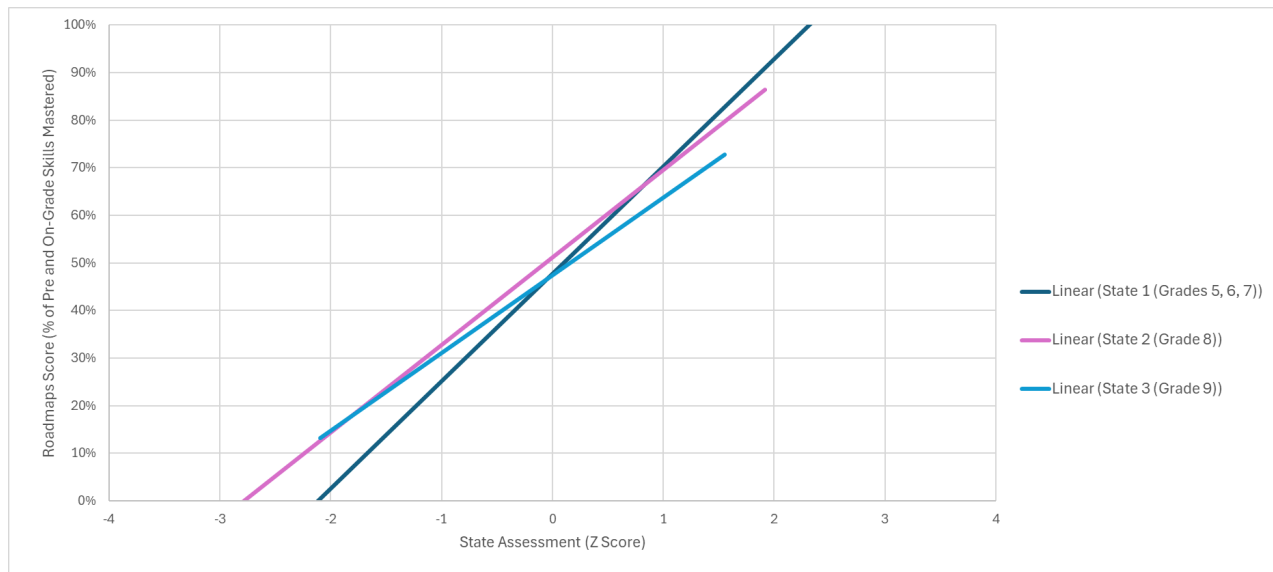Figure 3. Relationship Between Roadmaps and State Assessment Results by State - Scores
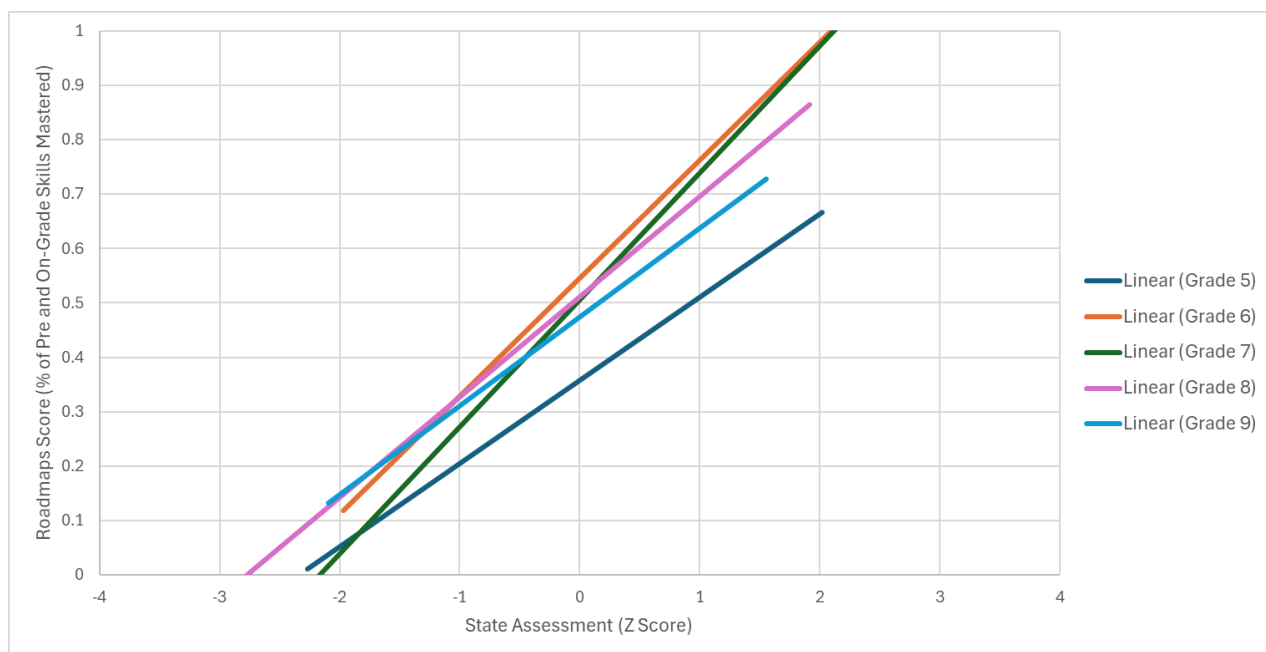
Figure 4. Relationship Between Roadmaps and State Assessment Results by Grade - Scores



Figures 5 and 6 show the relationship between *Roadmaps* scores (categorized into 1 of 5 levels) and performance level based on state assessment scores across grades and states. As with the comparison of scores shown in Figures 3 and 4, these results confirm that the strength of the relationship is consistent across grades and states. To note, for 6th grade, no students were in the lowest performance category for their *Roadmaps* scores.

Further, proficiency in state 3 is defined as achieving a level 3 or higher while students in states 1 and 2 must achieve a level 4 or 5. We did not see any student who mastered 100% of their pre and on-grade skills who did not also score at a level to be marked as proficient on their state assessment.

Figure 5. Relationship Between Roadmaps and State Assessment Results by State – Performance Levels
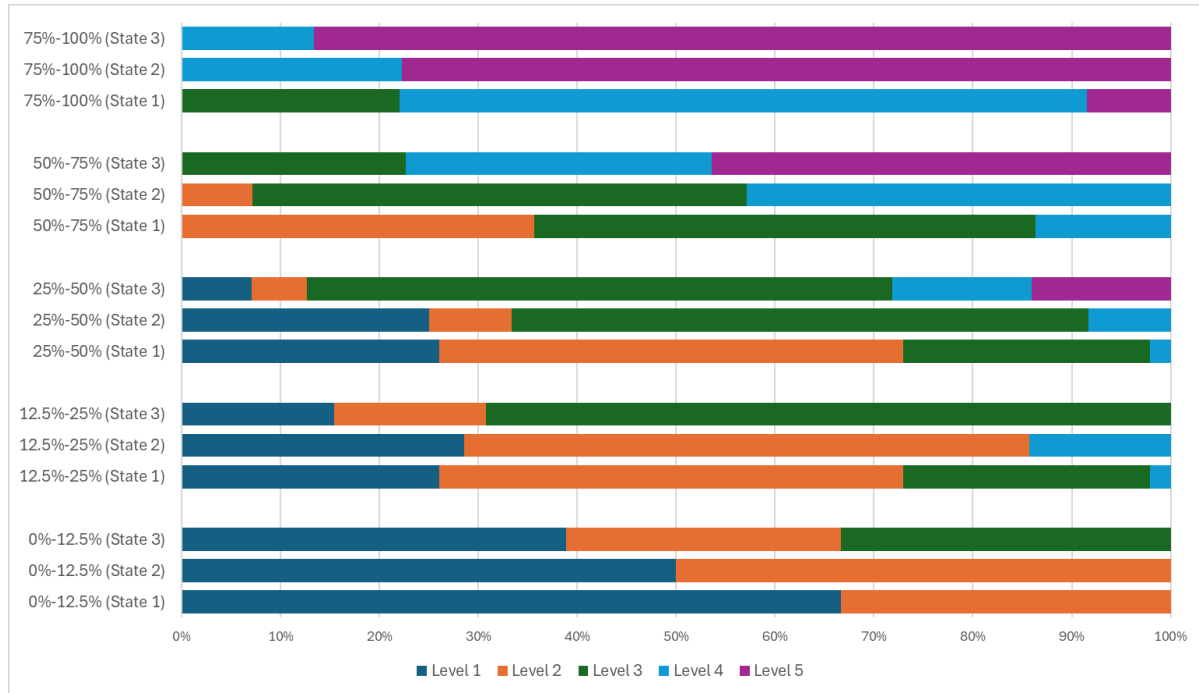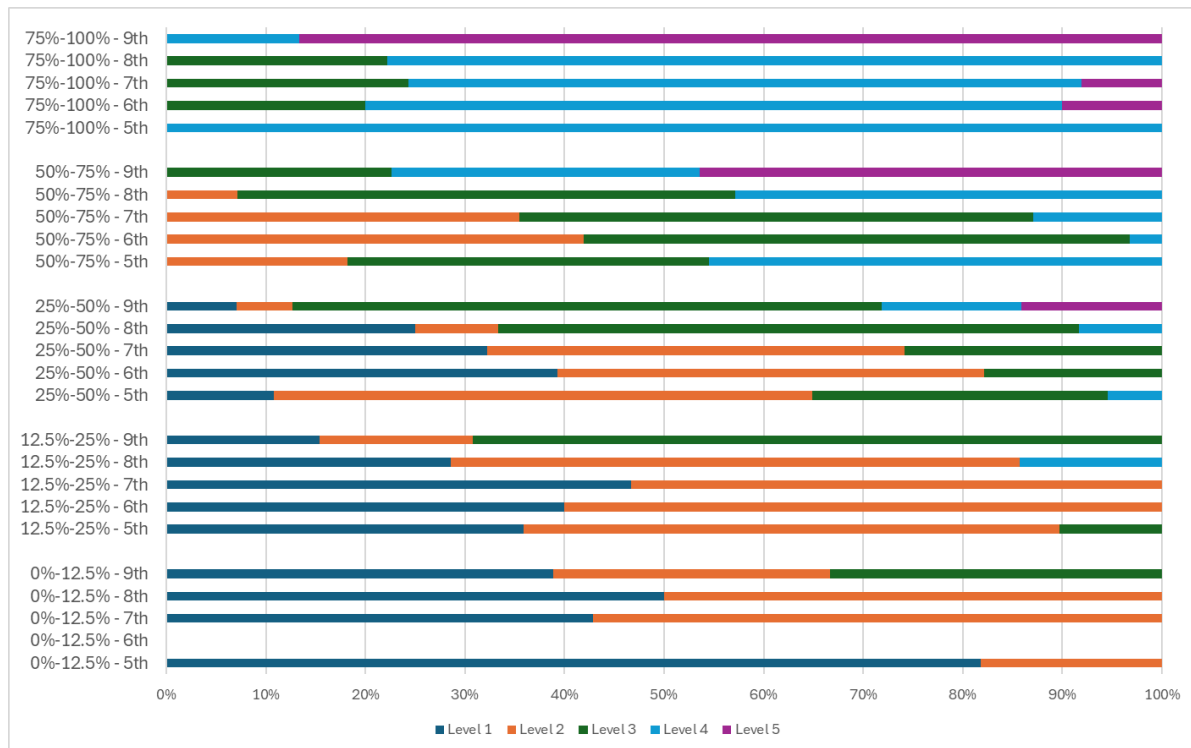


Figure 6. Relationship Between Roadmaps and State Assessment Results by Grade – Performance Levels

The predictive utility of the *Roadmaps* results was explored across states and grades through regression analyses. The results, shown in Table 4, indicate that the *Roadmaps* performance was a significant predictor of state assessment results across all five grades and three states. The 8th grade / State 2 sample has the lowest predictive utility ($R^2$ and $R^2$ adjusted) which may be due to the smaller sample size.

Table 4. Regression Results by Grade and State

| Group | State | Students | $R^2$ | $R^2$ (adj) | p-value |
|---|---|---|---|---|---|
| 5 | State 1 | 100 | 0.5031 | 0.4822 | < .0001 |
| 6 | State 1 | 84* | 0.6849 | 0.6731 | < .0001 |
| 7 | State 1 | 121 | 0.6934 | 0.6828 | < .0001 |
| 8 | State 2 | 44* | 0.4781 | 0.4246 | < .0001 |
| 9 | State 3 | 201 | 0.5239 | 0.5142 | < .0001 |

**\*** Results based on N<100 should be interpreted with caution

**Evaluation by Domain**

Each skill within the *Roadmaps* tool is linked to one of five domains. Table 4 shows the number of skills, at each grade, mapped to each domain.

Table 4. Number of Skills by Domain

| Domain | Grade 5 | Grade 6 | Grade 7 | Grade 8 | Grade 9 |
|---|---|---|---|---|---|
| Number Sense | 55 | 63 | 65 | 46 | 36 |
| Algebraic Thinking | 14 | 24 | 26 | 44 | 72 |
| Functions | 0 | 0 | 0 | 9 | 25 |
| Geometry | 18 | 23 | 28 | 36 | 6 |
| Statistics and Probability | 1 | 9 | 13 | 3 | 8 |

The predictive utility of *Roadmaps* scores within each domain was further analyzed by state / grade. The results of this analysis are shown in Table 5. Across all grades, all *Roadmaps* domain scores were significant predictors (p<.01) of the state assessment scores. The exception to this are three of the domains at grade 8 (Geometry, Functions, and Number Sense) which may be due to the limited sample size (N=44) combined with the large number of skills included as predictors.

Table 5. Regression Results of Roadmaps Domain Scores to State Assessment Results by Grade

|  | $R^2$ | $R^2$(adj) | p-value |
|---|---|---|---|
| **Grade 5 (State 1)** | | | |
| Geometry | 0.6167 | 0.5429 | < .0001 |
| Algebraic Thinking | 0.5719 | 0.5719 | < .0001 |
| Number Sense | 0.7999 | 0.6263 | < .0001 |
| Statistics and Probability | 0.1021 | 0.09299 | 0.0012 |
| **Grade 6 (State 1)*** | | | |
| Geometry | 0.7999 | 0.7231 | < .0001 |
| Algebraic Thinking | 0.7813 | 0.7025 | < .0001 |
| Number Sense | 0.9157 | 0.7085 | < .0001 |
| Statistics and Probability | 0.5771 | 0.5257 | < .0001 |
| **Grade 7 (State 1)** | | | |
| Geometry | 0.7757 | 0.7075 | < .0001 |
| Algebraic Thinking | 0.7557 | 0.6881 | < .0001 |
| Number Sense | 0.8860 | 0.7513 | < .0001 |
| Statistics and Probability | 0.7556 | 0.7260 | < .0001 |
| **Grade 8 (State 2)*** | | | |
| Geometry | 0.8222 | 0.4540 | 0.0574 |
| Algebraic Thinking | 0.9989 | 0.9837 | 0.0026 |
| Functions | 0.3782 | 0.2136 | 0.0387 |
| Number Sense | 0.8471 | 0.3427 | 0.1941 |
| Statistics and Probability | 0.2900 | 0.2553 | < .0001 |
| **Grade 9 (State 3)** | | | |
| Geometry | 0.3046 | 0.2831 | < .0001 |
| Algebraic Thinking | 0.7657 | 0.6339 | < .0001 |
| Functions | 0.5107 | 0.4408 | < .0001 |
| Number Sense | 0.6767 | 0.6057 | < .0001 |
| Statistics and Probability | 0.3682 | 0.3418 | < .0001 |

***** Results based on N<100 should be interpreted with caution

## Summary

This report summarizes the evaluation of the personalized pathways embedded into Teach to One *Roadmaps* as a predictor of state assessment results. For this evaluation, data was collected from three states including students' *Roadmaps* scores (percent of pre and on-grade skills mastered) and state assessment results (reported scores and performance levels). Each of the research questions are presented below along with the related findings.

1. What is the relationship between student performance within *Roadmaps* and their performance on the state assessments?

    o *Roadmaps* scores were highly correlated with state assessment scores and were a significant predictor of performance on the state assessment.

    o This strong relationship was also evident through comparison of *Roadmaps* scores across students at different performance levels.

    o We did not identify any students who demonstrated mastery of all of the skill-level assessments on their personalized academic roadmap and who did not score proficient

2. Do these relationships vary across subgroups of students within the data? (e.g., across states and grades)

    o The strong relationship between *Roadmaps* scores and state assessment results was found across state samples and grade levels.

    o This was evident through correlations and regression analyses which indicated *Roadmaps* was a significant predictor of state assessment scores across all state samples and grade levels.

3. Are there types of skills that are more or less predictive of performance on the state assessment?

    o Mastery of skills within each domain were significant predictors of performance on the state assessment. However, the skill domain with the most predictive utility varied across grade levels.

    o The five domains within the *Roadmaps* system were also predictive of state assessment results and this relationship held across grades with just one exception.

## Appendix A: State Assessment Information

State assessment scores were converted to Z scores which reflect the distance a score is from the group mean in standard deviation units. This allowed the scores to be analyzed together across states and grades. The distribution of scores on the Z score metric is shown in the graph below.

Table A.1 Exam Z score distribution
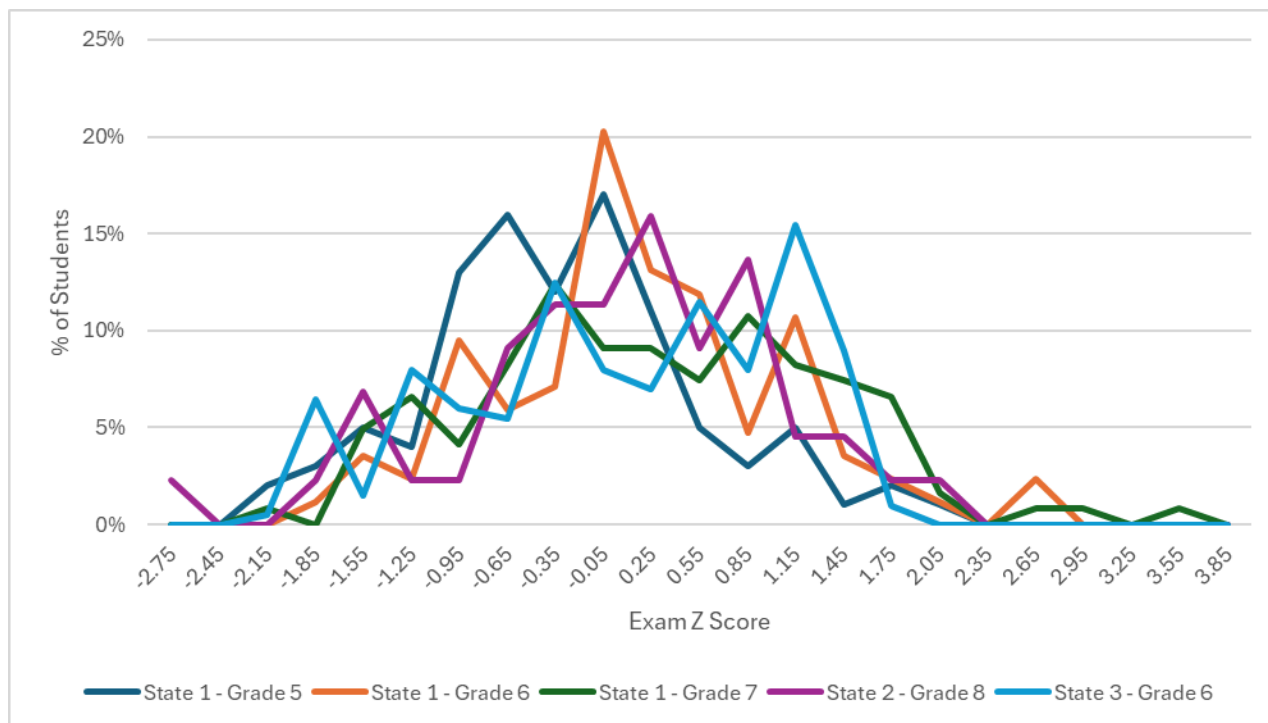


The three state assessments used in this study report out student performance across five achievement levels. Each level is briefly described below.

- **Level 5**: **Exceeds Expectations**
  - o Students at this level demonstrate exceptional mastery of grade-level academic expectations. They are well-prepared for success in various post-graduation options.
- **Level 4**: **Meets Expectations**
  - o Students meeting this level have met the District's rigorous grade-level academic expectations. They are building essential skills for future success.
- **Level 3**: **Approaches Expectations**
  - o Students at this level are approaching grade-level expectations. While they haven't fully met them, they are making progress.
- **Level 2**: **Partially Meets Expectations**
  - o Students partially meet grade-level expectations. There's room for improvement, but they are working toward proficiency.
- **Level 1**: **Did Not Meet Expectations**
  - o Students at this level did not meet grade-level expectations. Additional support and intervention may be needed to help them progress.