Using MAP for Strategic Framework Milestones and SIP Metrics

Feedback from various stakeholders has led us to examine the use of MAP to measure Strategic Framework Goal #1: *Every student is on track to graduate as measured by student growth and achievement at key milestones.* In particular, we have received three specific questions regarding our use of MAP data for Strategic Framework Milestones and SIP Metrics for 2016-17:

1. What is the best way to measure growth on MAP?
2. How should the district and schools set MAP goals for growth?
3. How should the district and schools set MAP goals for proficiency?
4. Should we track progress based on Proficient-Advanced or Basic-Proficient-Advanced?

In this document, we summarize the key issues for each of these questions and provide our recommendations.

1. What is the best way to measure growth on MAP?

Currently, MMSD uses the percent of students meeting or exceeding fall to spring growth targets on the MAP assessment as both a Strategic Framework Milestone and School Improvement Plan (SIP) metric. In addition, this metric receives significant attention in our public reporting on MAP in other venues and teachers have been trained over the past several years to use it to measure progress at the classroom and student level. We have included growth as a complement to MAP proficiency; it allows us to look not just at how students are performing, but also improvement during the year.

MAP growth targets are calculated by NWEA based on average gain for students with the same fall score and in the same grade. NWEA uses a national sample to create these targets, so the growth target represents the average performance of American students. For example, a grade 3 student scoring a 196 on the MAP reading assessment in the fall would have a growth target of 9 points. This means that students in the United States in grade 3 who scored a 196 in the fall, on average, scored a 205 in the spring – a gain of 9 points, and the growth target for every student scoring a 196 in the fall. Different resources from the vendor lead to slightly different interpretations of how growth within MMSD fits into a national context; however, in general, we believe that growth in MMSD at the elementary level has consistently exceeded national averages while growth at the middle school level has been close to average.

We have received questions about whether MAP growth is an appropriate measure to use for measuring Goal #1.

Focus on the Current Approach:
The table below outlines the pros and cons of maintaining our current approach to MAP growth:

Area	Advantages	Disadvantages
How useful and effective is the measure itself?	• Used regularly across MAP districts • Simple to use • Targets already provided to staff via data from the vendor, so goals and needed progress to meet them are not a mystery • Clear connection from district to school to class to student level data • Can set ambitious goals using these measures, even if they are based on national averages. For example, we as a district strive for percent meeting growth targets well above 50%, which means more than half of our students are meeting or exceeding average gains • Do not require new calculations from central office staff	• Growth target reflects individual student performance relative to national average and achievement gaps persist nationwide • Students meeting growth targets over time will not necessarily reach proficiency
What technical tools and resources do we already have in place?	• All existing Dashboard programming, including progress monitoring tools for schools, SIP Goals Reports, and Strategic Framework milestones, relies on our current approach; a new approach would require substantial reprogramming • NWEA (vendor) has existing resources around MAP growth • Data Use Guides produced by MMSD already are available explaining MAP growth	• N/A
How does this work impact teachers, principals, and other stakeholders?	• Years of effort already put into place building user understanding of MAP growth • Use of tools reflecting current approach to MAP growth is routine and regular in schools • Use of MAP growth is routine throughout district practices like quarterly progress reviews, intervention selection, planning rosters, and program evaluations • Families have knowledge of measure as currently used	• Understanding of MAP growth is not universal • High performing schools may not be able to maintain increases over time

Alternatives:

Through our review of our data and MAP materials posted by other districts, we identified several alternative approaches to measuring MAP gain/growth. It is important to note that district-level accountability measures are rare and district goals on these measures even rarer, so most of our review focused on public reporting as opposed to accountability metrics by necessity. The table below outlines those approaches as well as the advantages and disadvantages we see to each approach. The last five columns evaluate each option on five criteria we see as important for practical and effective implementation, with green denoting a yes answer:

        C1: Do tools already exist?
        C2: Can we track progress at the student level?
        C3: Is every student included?
        C4: Is the measure readily accessible to staff?
        C5: Can we compare to other districts or the nation?

Strategic Framework Milestone Option *Example of Student Success*	Advantages	Disadvantages	C1	C2	C3	C4	C5
1. Percent of students meeting fall-spring growth target (status quo) *Fall score of 240, fall to spring growth target of 6 → spring score of 246 or higher*	• See table above	• See table above	✅	✅	✅	✅	✅
2. Percent of students improving RIT score from fall to spring *Fall score of 240 → spring score of 241 or higher*	• Easy to understand and calculate	• Not ambitious • Students improving may not be keeping pace with national averages		✅	✅	✅	
3. Percent of students exceeding fall-spring growth target *Fall score of 240, fall to spring growth target of 6 → spring score of 247 or higher*	• Easy to understand and calculate • Students exceeding growth target are outpacing national averages so will be closing gaps	• Requires construction of new tools and developing new understanding in schools		✅	✅	✅	
4. Percent of students meeting multiple of fall-spring growth target (e.g. 1.5 times growth target) *Fall score of 240, fall to spring growth target of 6 → spring score of 249 or higher*	• Easy to understand and calculate • Students exceeding growth target are outpacing national averages so will be closing gaps	• Requires construction of new tools and developing new understanding in schools		✅	✅	✅	
5. Average RIT score gain percentile *N/A – not a student level measure – goal would be something like "school is above 60th percentile for average RIT score gain"*	• Tables available from NWEA (MAP vendor) to facilitate calculation • Potential for most recent comparison to national averages	• Goal tracking cannot happen at student level • Averages susceptible to outliers • Does not have sustained history of use by NWEA			✅		✅
6. Percent of students improving proficiency band *Fall score basic → spring score proficient*	• Measures progress toward proficiency • Easy to understand	• Does not recognize large movement within proficiency bands • Less effective for students starting with high scores	✅	✅			
7. Cohort-based proficiency changes over time *Fall grade 3 score basic → spring grade 5 score proficient or higher two school years later*	• Measures progress toward proficiency • Focuses accountability on students with stable enrollment across time	• Mobility in MMSD means many students will be excluded from models	✅				
8. Value added *N/A – not a student level measure – goal would be something like "school has above average value added"*	• Highly sophisticated method that accounts for many demographic characteristics	• Data not received until far after school year • Not a student level measure • Politically sensitive			✅		✅

Recommendation: Although we acknowledge the limitations of the fall-spring growth measure, we recommend continuing its use based on the balance of considerations outlined above. This measure is simple, accessible, and already in place. Alternative approaches that are more statistically sophisticated are technically possible for us to produce; however, with every Strategic Framework milestone, we have sought a balance between information, transparency, and accessibility. We believe the minor gains in information that could be achieved by a more sophisticated model would be more than offset by issues in transparency, with staff no longer able to quickly and easily calculate which of their students are meeting the goal, and issues in accessibility, as the relatively simple explanation for MAP growth is replaced by a more complicated explanation of a different model. We also want to emphasize the large amount of work that would accompany a new approach in terms of developing new documentation and understanding, as well as reprogramming many technical tools; this process would take months. In addition, other districts using MAP, including some of the largest school districts in the country, tend to use the MAP growth measure in the same way we do in MMSD. Finally, any new approach would require us to explain why our results changed from one year to the next; explaining how changes in methodology impact results can be difficult, and individuals who do not understand the changed methodology might erroneously credit or criticize the district.

For MAP Growth, we recommend moving away from the practice of recommending increases for every school in every year. Based on what we know about MAP growth, our district's results already are strong (NWEA states that 50% of students meeting growth targets would be typical and we consistently exceed that number), and some schools are well beyond national averages. For results that are far above average, continued improvement may not be realistic, and even some decline may not represent worse performance but instead regression toward a mean.

Therefore, we propose developing thresholds above which schools would be expected to maintain their percent meeting growth targets, rather than continuing to increase. The thresholds we propose are based on approximately the 70th percentile nationwide and are as follows:

	Grades 3-5	Grades 6-8
Reading	65%	60%
Math	70%	65%

As an example of what these means in practice, under our previous model, an elementary school with 71% growth would have received a goal recommendation of 73%. Under our new recommendation, this same school would have received a recommendation to stay above 65%. We believe this new approach is a fairer treatment for schools and groups that are already very high performers.

2. How should the district and schools set MAP goals for growth?

The SIP Goals Report that schools use to set their SIP achievement goals provides a recommended goal for each measure and student group. As a reminder, recommended goals for **MAP Proficiency** were **2%-5%**, and recommended goals for **MAP Growth** were **2%-5%-10%** for 2015-16, based on performance relative to district averages. We have been asked to examine whether these recommended goals are sustainable, realistic, and the best way to set goals.

First, we should consider the goal trajectories used in prior years and what they mean for future data. Baseline data for this process started with the 2012-13 school year. We use grade 5 reading data as an example.

Grade 5 Reading Growth

Legend: Actual — Original Goal Trajectory — Recommended Goal
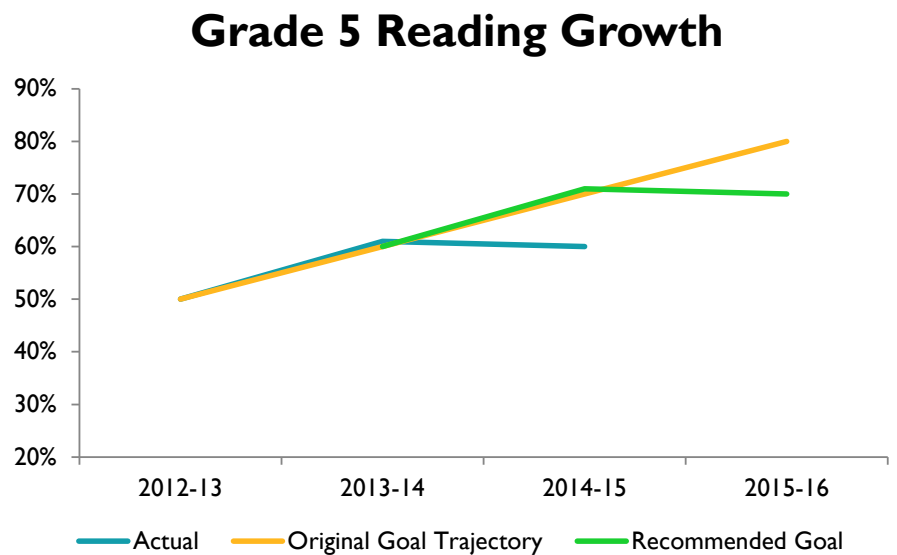
For MAP growth, our initial growth trajectory involved a 10 percentage point improvement each year for the district. This goal has extended to SIPs for the past three years, as schools near district averages have received the goal recommendation of 10% improvement; that recommendation changed to 5% starting in 2015-16. The graph to the right illustrates our **original trajectory** of 10 percentage points a year, our **recommended goals** for each year (the previous year's actual result plus an improvement of 10%), and our **actual results** from each year.

This graph shows us that the original plan of 10% improvement in growth per year would have placed us around 80% in the current school year. Although we believe in setting ambitious goals, the idea that we would continue to improve 10 percentage points every year likely was not realistic, and now that we are around 60% of students meeting growth targets, we may want to consider a lower target than 10 percentage points each year, as even 5 percentage points is relatively large.

MAP Growth Goals and Results 2014-15
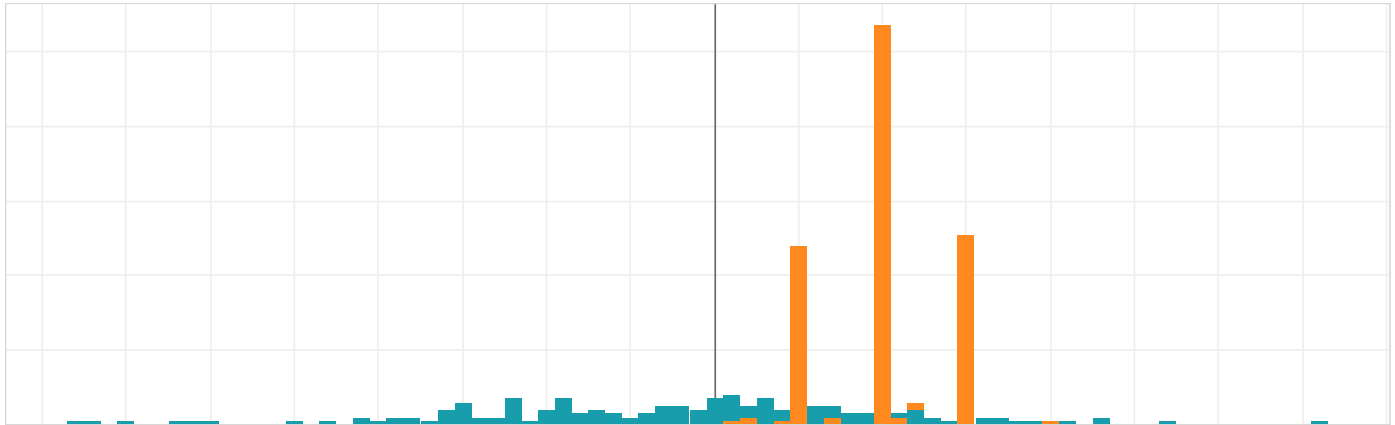
Almost all schools set goals for MAP growth that aligned with a district recommendation: 5%, 10%, or 15%. In addition, we see that very few schools actually achieved growth improvements of 5% or more, with changes in growth generally clustering around 0%.

Recommendation: Schools/groups within 10 percentage points of the MAP growth threshold would receive a recommendation for 2% improvement and schools/groups more than 10 percentage points from the threshold would receive a recommendation for 5% improvement.

3.   How should the district and schools set MAP goals for proficiency?

For MAP proficiency, our initial growth trajectory involved a two percentage point improvement each year for the district. This goal also has extended to SIPs for the past three years, as schools near district averages have received the goal recommendation of 2% improvement. The graph to the right illustrates our **original trajectory** of two percentage points a year, our **recommended goals** for each year (the previous year's actual result plus an improvement of 2%), and our **actual results** from each year.

Grade 5 Reading Proficiency
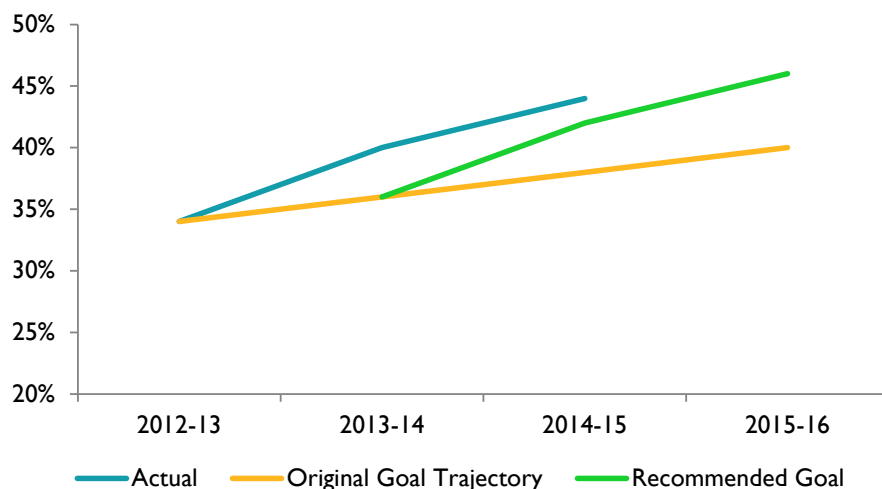
Our proficiency improvements have consistently outpaced the original intended trajectory of two percentage points a year. Our recommended goals based on the prior year's results have been higher than the original trajectory for the past two years. Based on this evidence, there is no reason to believe our original trajectory was unrealistic, and continued recommended improvement of 2% per year for the district appears reasonable.

Now that we have examined the sustainability of these goals, we will examine how realistic they are based on recent results at the school and focus group level. The graphs below show the **goals set by schools (graphed in orange)** and the **actual changes they achieved (graphed in teal)** at the school and student group level for 2014-15. Note that these graphs reflect only focus groups for which goals were set.

MAP Proficiency Goals and Results 2014-15
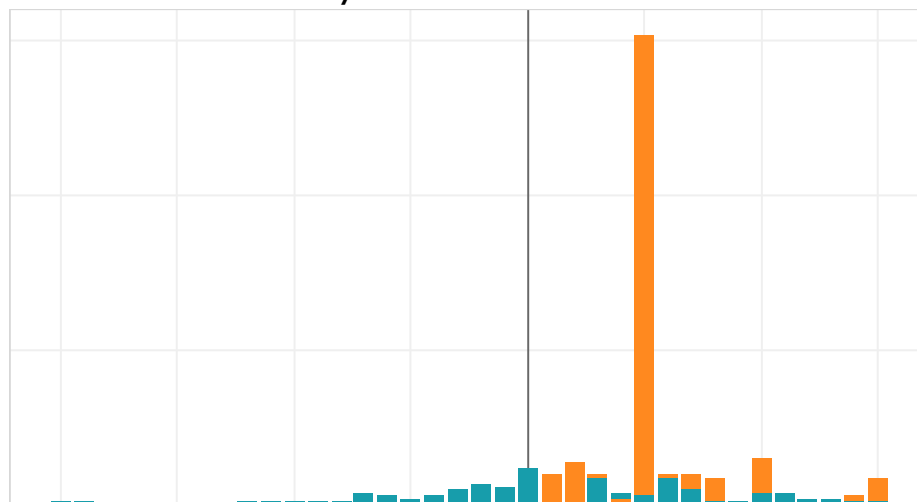
From this graph, we see that the most frequent goal for MAP proficiency improvement was 5% by a large margin. The next most common goals were 10%, 2%, 1%. A fair number of schools had proficiency improvements within groups of 5% or more, and even more exceeded 2%.

Recommendation: For MAP Proficiency, we believe that 2%-5% remains an attainable model for goal recommendations. Many schools were able to attain progress within and even above these recommendations, as evidenced by the blue bars appearing to the right of the 2% column. However, we believe setting goals of 10% or 15% may not be realistic in practice and suggest that when schools do so, they should provide compelling rationale that such increases are reasonable.

4.   Should we track progress based on Proficient-Advanced or Basic-Proficient-Advanced?

Our third outstanding question is whether we should continue to use our current approach to proficiency, which includes students scoring "Proficient" or "Advanced," or whether we should include the "Basic" result band in our rollup of reported proficiency.

We used data from the [American Institutes for Research College & Career Readiness & Success Center](#) to examine other states' practices for reporting proficiency. AIR includes information on state reporting standards for 49 states and the District of Columbia.

All 49 states, as well as DC, use Proficient or higher as their measure of progress (although specific state-level tests vary). In addition, based on an online search for other districts' materials, we can find only isolated cases of districts using Basic or higher as opposed to Proficient or higher, and even then often for specific cases (e.g. Students with Disabilities or promotion standards).Districts often cited as peers or role models for MMSD do not report on a rollup of basic or higher scores, instead using proficient (or the equivalent) or higher; this includes Ann Arbor (MI), Austin (TX), San Francisco (CA), Chicago (IL), and Cambridge (MA).

Choosing to include Basic in our proficiency rates would mean losing the ability to compare MMSD's results to other districts, state, and national level results.

Based on our review of this data, we cannot find any evidence of Basic or higher being used as a reporting metric. In addition, we have significant concerns about the perceptions of using this standard, which is lower than those used by every state in the nation. Prior to the re-norming of the WKCE, Wisconsin had one of the lowest standards for proficiency in the nation. Once the WKCE was re-normed to align with the NAEP (and our MAP cut scores were aligned with those assessments as well), Wisconsin's state proficiency standards (and, by extension, our MAP proficiency standards) came more into line with the national context.

We are concerned that including Basic represents lowering the bar for what we expect in terms of student success. The Basic result band expresses that "students are able to show some of the knowledge and skills, but not consistently or to the level that is expected at that grade level." We believe our standard should be higher. We also believe using basic or higher would be internally inconsistent with our approach to the high school profiles in Goal #2, where we require students to attain a C or better. Data showing the percent of students scoring Minimal or Basic will be included in the district's annual MAP report and is always available on demand via the Data Dashboard.

Recommendation: We recommend that MMSD maintains the use of Proficient-Advanced as our proficiency standard for Strategic Framework Milestones and SIP Metrics. This practice is aligned with national trends and makes our results comparable to state assessments and other common district practices.