Continuous Improvement, Evaluation Cycles, and Scaling What Works in MMSD

for Board of Education

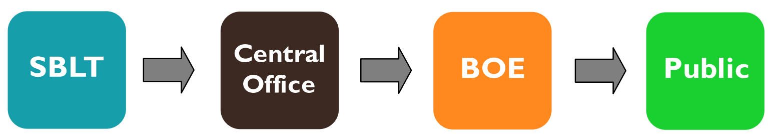
To ensure effective execution of our district strategy, we will utilize several crucial ways of working that build on our past practice but are in better alignment with our core values. In this document, we focus on our processes in three areas: Continuous Improvement, Evaluation Cycles, and Scaling What Works.

Continuous Improvement

In MMSD, our work around the Strategic Framework follows an Annual Planning Cycle, which involves periodic reviews of progress at all levels, annual data reporting and evaluations of major plans and high-intensity partnerships.

Reviews of Progress

The central resource for planning within the cycle for each school is the School Improvement Plan (SIP), while the central resource for each Central Office department is the Central Office Measures of Performance (COMP) and the central resource for the Board of Education is the Strategic Framework. The Annual Planning Cycle includes three to four reviews where schools, Central Office, and the Board review progress, reflect on recent actions, and plan for the future on a regular cycle. Each review follows a similar cycle:



For Schools

For schools, quarterly reviews happen during quarterly School-Based Leadership Team (SBLT) Institutes. At these SBLT Institutes, all school principals and select members of their SBLT meet to discuss SIP and Strategic Framework progress, review data, and plan for the work ahead. At each meeting, we review a consistent set of recently available outcome data (such as MAP scores or climate survey results) at the district level, with school-specific resources available for the teams attending. After the SBLT Institutes, attendees can take these materials back to their schools to engage in a building-based review.

For Central Office

For Central Office, regular reviews happen around Central Office Institutes (COI). At the Central Office Institutes, all Central Office staff spend the first half of the day reviewing progress on Strategic Framework Priority Areas and other data. Each department produces a Central Office Measures of Performance (COMP) plan each year, which is the Central Office department equivalent of a SIP. The second half of a COI day may be spent working on the COMP, engaging in shared problem-solving, debriefing on shared learning from the morning, and/or team-building. In addition, at the end of the year, Central Office departments engage in a review of Climate Survey and Central Office Satisfaction Survey results specific to their department.

For Board of Education

For the Board, reviews happen during board retreats with the Senior Leadership Team (SLT). The first retreat of the school year lays out the scope of Priority Projects that will be undertaken in support of the Strategic Framework. All retreats will include updates on the Strategic Framework. The last retreat of the year is our end of year review.

Annual Data Reporting

The table below reflects major data projects conducted by RPEO's research team and provided to the Board on an annual basis. **Please note that these projects do not reflect ad-hoc requests, one-time research projects and regular consulting functions. These dates are subject to change.**

Title	Typical Month
ACT/Aspire	October
Hmong Student Data	October
American Indian Student Data	October
MAP	October
PALS	October
Attendance	November
Behavior	November
Summer School	November
Arrest and Citation Report	November
Enrollment	November
Human Resources Report	November
Climate Survey	June
High School Completion	June
Annual Report	July

Review cycles for materials sent to the Board of Education vary, but a typical review process for a report is as follows:

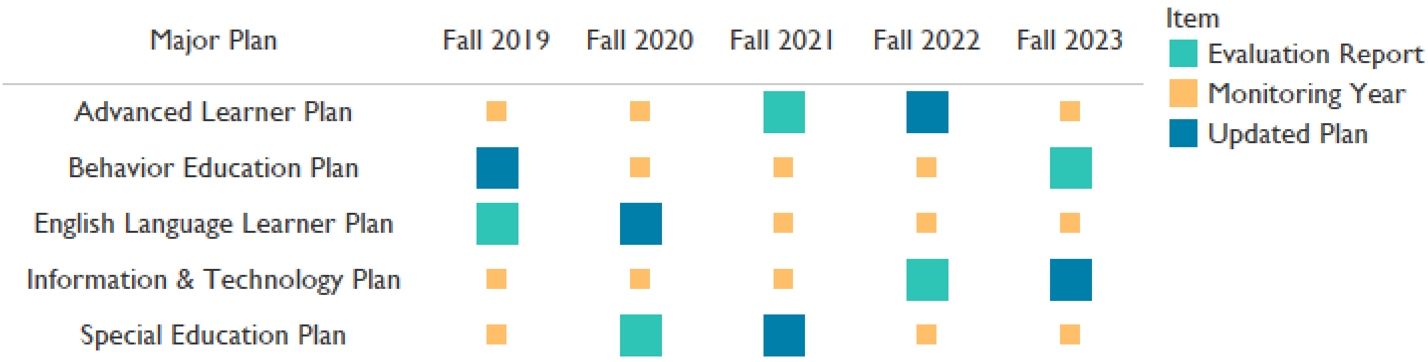
1. Draft report completed by RPEO staff
2. Report shared with program leadership; feedback incorporated
3. Report shared with Superintendent; feedback incorporated
4. Report shared with SLT/COLT; feedback incorporated
5. Report shared with Principals via Bulletin or meeting
6. Report shared with Board

Evaluation Cycles

Evaluation for Major Plans

Starting in 2015, the district operated on a cycle of evaluating major plans every three years. Through this work and feedback from leaders throughout the district, we have learned more about what worked well and where improvements can be made. We found that these evaluations can be informative, but they also require major investments of both money and staff time and do not always line up well with the timing of plan renewals. Based on this information, we are recommending a couple changes to the evaluation cycle for major plans. These changes include modifying the timing of evaluations so that the initial evaluation happens after three years and all subsequent evaluations occur on a five-year cycle; having all evaluations conducted by an external consultant; and having the Research & Program Evaluation Office focus on creating more robust annual monitoring updates. We describe the revised cycle below, calling out the specific changes and rationale for them within each section.

Currently, MMSD has five major plans that fit within the major plan evaluation cycle: English Language Learner, Advanced Learner, Special Education, Behavior Education Plan, and Ignite (formerly the Information & Technology Plan). We recommend evaluating these plans initially after three years, and then every five years from that point forward. The graphic below shows the recommended timeline of reports through the Fall of 2023. We suggest this timeline for multiple reasons. First, summative evaluations are expensive, so we believe the district should invest in only one each year if possible. Second, we believe a cycle that focuses on only one major summative evaluation a year allows for the greatest focus and highest quality decision-making. Finally, given that evaluations should inform plan updates, and plans often take a year or more to revise, a three-year evaluation cycle means we would regularly be evaluating plans that had been in place for less than two years, barely enough time to overcome implementation dips and begin to see results. The graphic below shows the timing for evaluations through 2023.



Monitoring Year: Program shares materials with the Board of Education including progress and highlights from the year; these updates typically are shared in February and reflect data through the prior school year. We intend annual monitoring updates to be the main vehicle for continuous improvement of our major plans. These updates will include narrative from plan leaders about what they have done over the past year, what they have learned, and what they will do differently moving forward, including any planned adjustments to the plan. They also will include data on a set of key indicators that allow for formative reflection on success and challenges. Example indicators include suspension and behavior event totals for the Behavior Education Plan and advanced learner identification trends for the Advanced Learner Plan. The Research & Program Evaluation Office will produce more robust versions of these annual monitoring updates than in previous years. We believe our best investment is in increasing the prominence of annual monitoring updates for each plan to allow both ourselves and the Board to be nimble, engage in robust continuous improvement, and act if concerns arise.

Evaluation Report: Produced five years after the last Evaluation Report. Focuses on implementation and outcomes since the inception of the plan or last report and draws evaluative conclusions about progress and effectiveness. These reports typically will be shared near the beginning of the school year and will be produced by external consultants. By contracting this work out, MMSD ensures a third party review while freeing up capacity for RPEO to work on the annual monitoring to drive continuous improvement.

Updated Plan: Using the results of the summative evaluation completed the prior year, as well as lessons learned through regular monitoring and continuous improvement over the past five years, an updated five-year plan will go into effect. Because updated plans would be informed by the findings of each evaluation, these plans would take effect the year after the summative evaluation is presented. For example, a summative evaluation presented in the fall of 2025 would reflect data from the 2020-21 through 2024-25 school years and be used to inform an updated plan drafted during 2025-26 and taking effect for 2026-27.

Each of these resources focuses on the success of the plan, not of the student group referenced. For example, the ELL Plan evaluation focuses on whether the plan has been successful, not whether ELLs overall have succeeded in the district. We review the progress of specific student groups continually, with a special focus on end-of-year data, so we do not believe these evaluations should attempt to assess the success of both the plan and the student group. To keep the evaluations focused on the plans, we believe every major plan needs a clear set of goals and metrics included as part of the plan.

Evaluation for High-Intensity Partnerships

The content below reflects our understanding under the previous Evaluation & Review Cycle. We recommend modifying this content once the new Director of Strategic Partnerships is hired and has a chance to meet with our community partners.

Currently, MMSD has nine high-intensity partnerships formalized per Board Policy 7544 that are expected to be both monitored and evaluated on a regular basis. High-intensity partnerships will be formalized through Board approval of a three-year Memorandum of Agreement (MOA), unless otherwise specified with the MOA. Any substantive changes to program models during the three-year MOA cycle will be brought to the Board for review and approval.

All high-intensity partnerships are expected to engage in an annual monitoring process which involves partnership teams reviewing designated monitoring year outcomes and data (per MOA) twice a year. Partnership teams review first and second quarter monitoring year outcomes and data during a mid-year review meeting in quarter three. Partnership teams review end of year monitoring year outcomes and data and prepare an Annual Report during quarter one. Annual Reports will be shared with the Board of Education via the Weekly Update and will be used by partnership teams to inform revision of the MOA on the three-year cycle.

Evaluations of high-intensity partnerships are driven by Board Policy 7544 – School and Community Partnerships which establishes a procedure for formalizing and monitoring any partnership between MMSD and community entities. Given that all high-intensity partnerships involve a higher investment of resources to carry out, we expect them to engage in an external evaluation on a regular cycle. We will implement a three-year cycle for external evaluation that coincides and supports MOA revision and renewal. Specific evaluation years are subject to change, pending resources available to fund evaluation efforts and alignment with grant and/or other evaluation expectations.

	18/19	**19/20**	**20/21**	**21/22**	**22/23**
AVID/TOPS/College Club	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year
CBITS	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal
Forward Madison	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year
Behavioral Health in Schools	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year
Juventud and Escalera	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year
PEOPLE Program	Monitoring Year	Monitoring Year	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal
Schools of Hope AmeriCorps - Elementary	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year
Schools of Hope Middle School Program	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year
Youth Apprenticeship	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year	Monitoring Year
Early College STEM Academy - Phase 2	Approval	Monitoring Year	Monitoring Year	Evaluation, MOA Renewal	Monitoring Year

* All high-intensity partnerships were be renewed during 2015-2016 to transition the MOA to the three-year evaluation and renewal cycles listed above. Thereafter, all high-intensity partnerships will reflect three-year terms.

Scaling What Works

As a district, we want to sharpen our processes for piloting and determining what needs to be brought to scale and at what pace. By doing so, we ensure that new ideas address clear gaps or needs, determine metrics for success at the outset and discontinue what does not work. Most important, we ensure that ideas brought to scale are sustainable and supported by the community we serve. Coburn (2003) posits four dimensions of scaling: depth, sustainability, spread, and shifting ownership. Research like this, as well as our own experience, tells us that scaling is complex and goes beyond increasing participation; it involves the ways we spread ideas, as well as the ways we assess and sustain change as it reaches new levels.

In this document, we focus on district-level scaling decisions, outlining a process for how we will structure district-level conversations about what practices to pilot and how and when to bring them to scale. We anticipate scaling conversations to apply to district-level decisions in areas such as:

- Instructional resources (e.g. Ignite)
- Instructional designs (e.g., DLI)
- Programs (e.g., AVID/TOPS)

We want to emphasize that not all work done in the district has to fall into this process. Many of the things we do are projects or work streams that are not subject to this kind of review. We also recognize that many ideas originate in schools and in the community, and that many small-scale practices are in place that could be regarded as local pilots. These ideas are important and valuable, and serve to inform and create ideas that are then considered at the district level within step one of the process. They would not, however, be considered part of this scaling process at this point. Future iterations of this process will likely include discussions of how we generate ideas and determine whether to consider scaling them. For this first iteration, however, we have chosen to focus on district-level practices that are sponsored and brought forward by Senior Leadership Team members.

The Process for Piloting, Validating and Scaling

To aid this type of decision-making, we have outlined our process for piloting, validating and scaling. The process includes decisions relating to the ways we want to try or scale an idea, as well as those that assess whether that idea is working to help us decide whether to continue scaling up the idea. As we progress through the process, the standard for evidence we would want to see to move forward should increase.

The process includes the following steps:

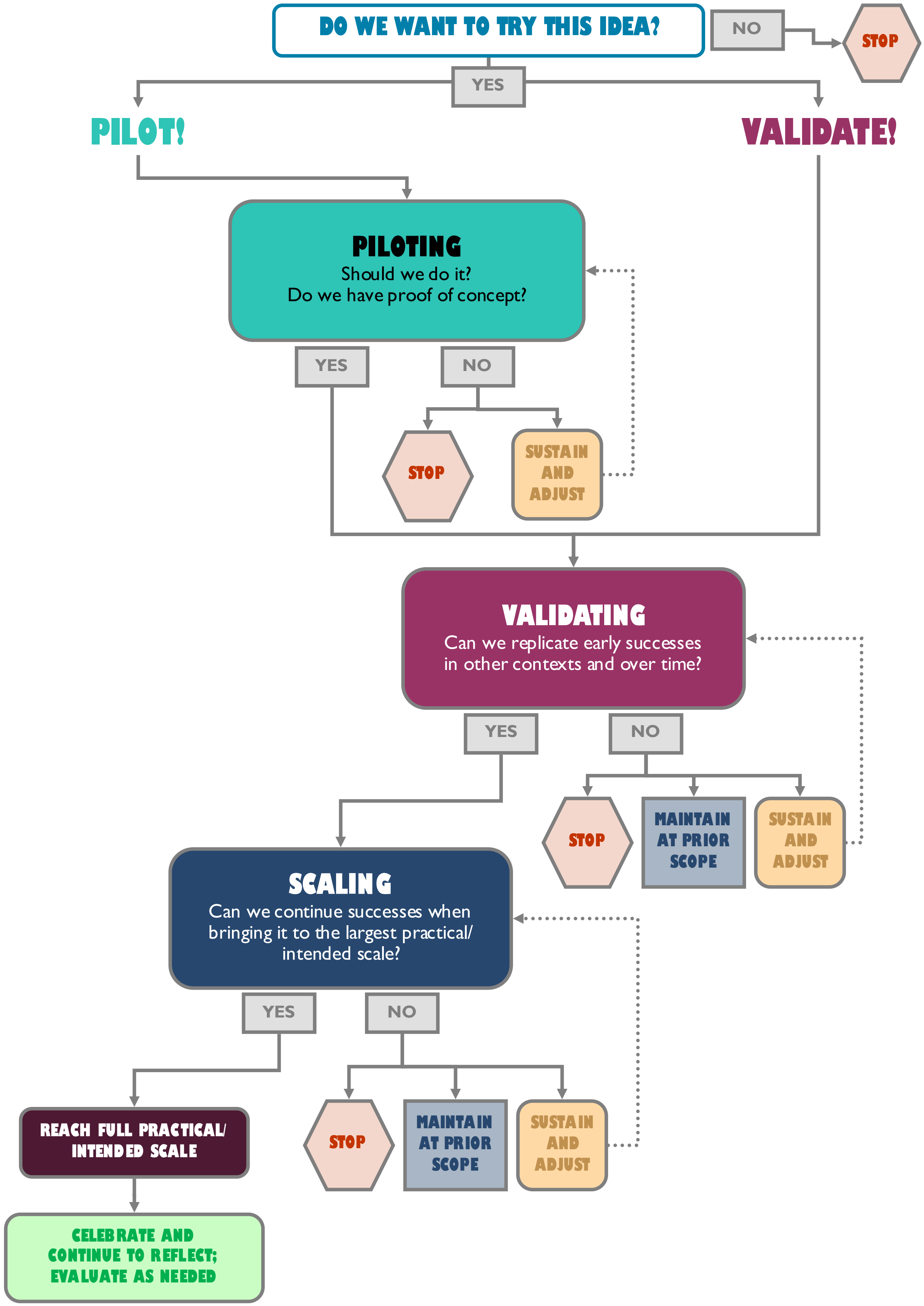
- **Do we want to try this idea?** *Decision: Pilot, Validate, or Stop*
- **Piloting** *Decision: Should we continue? Yes or no*
- **Validating** *Decision: Should we continue? Yes or no*
- **Scaling** *Decision: Should we continue? Yes or no*

As a follow-up to the decision not to continue, there are three possible options: adjusting the idea and restarting the current step with a different approach, maintaining the idea at its current level of implementation, or stopping the idea entirely.

We acknowledge that some district-level practices are underway that may already be at different steps in this process. For existing ideas and practices, we will do our best to identify where within this process they currently lie, but we know the fit may be imperfect. We also recognize that the process outlined here focuses on actions taken by Central Office leaders, the Senior Leadership Team, and the Board of Education. The voices of school staff, students, families, and community members should be integrated by those parties into the decisions we make at each step. That work is not outlined explicitly here, but is critical to making this process meaningful.

On the next page, we present an overview of the scaling process that shows how we intend to move across the steps. Subsequent pages provide more detail about what we believe happens within each step.

Scaling Process Overview



Do we want to try this idea?

The purpose of this step is to determine whether/how to try an idea. In this step, Central Office program leaders and Senior Leadership Team should answer the following questions:

- What will we do?
- Why will we do it? Does it fill a clear gap or need? Will it enhance equity?
- Why do we think it will work? Does local experience or research suggest it will work?
- Do we have stakeholder support and excitement?
- What will it cost to try it? Think about money, resources, and time, both initially and at broader scale.

If the answers to most questions above are positive, leaders can recommend moving forward with this idea.

Program leadership and Senior Leadership Team then recommend how to try this idea. Options include:

- We **pilot** ideas when they are ***promising practices*** (e.g., have strong theory or promise, but we feel we would need to learn more locally to be confident it would work for us)
- We **validate** ideas when they are a ***best practice*** (e.g., we are confident it will work based on evidence and the practice is best to put in place at a larger scale right away)

If we determine it is an ***emerging practice*** (e.g., some evidence exists, but we are not completely sure), we choose pilot or validate based on our confidence in the idea.

If we will **PILOT**, then move to the piloting step. If we will **VALIDATE**, then skip the piloting step and go straight to validating.

Piloting

At the start of this step, we focus on a couple of key questions:

- Where should this pilot take place?
- How will we know if it works? What are our metrics for success?
- When will we check in on results?

Then, we try the idea at a small scale.

Once we reach the time when we agreed we would check in on results, we reflect on the pilot to determine whether we should continue to scale this idea. In this step, we are asking whether we have achieved proof of concept (e.g., evidence that demonstrates a design is feasible and has practical potential). We ask questions like:

- How did we implement it?
- Did it show results that indicate it might work? Did we reach our metrics for success?
- Do we believe it can work if we continue to expand?
- Can we afford to do this? Think about financial resources, time, and opportunity costs.

Program leadership work with the Research & Program Evaluation Office as needed to determine the evidence and findings for each question. This evidence is then presented to the Senior Leadership Team. At the piloting stage, although we want to see positive results to continue, we might be satisfied with stories or perceptions of why the idea shows promise, as opposed to quantifiable outcomes that approach causality.

If there is sufficient evidence to determine proof of concept and potential for this idea to work at a greater scale, leadership and the Board would determine yes, this idea can be scaled further. If not, we would either choose to sustain and adjust the idea, which means we would try the pilot again using a different approach, or discontinue the idea.

Validating

The purpose of this step is to determine how we will spread this idea. In this step, Central Office program leaders and Senior Leadership Team answer questions such as:

- What conditions helped the pilot succeed?
- Where else do those conditions exist already? Where do we want to create them?
- Do we have stakeholder support and excitement?
- What evidence will we need to show success in validation? What are our metrics for success?
- When will we check in on results?

The answers to these questions, as well as the type of idea we are implementing, help program leadership decide key elements of expansion, including whether it will be immediate or phased, voluntary or required, across some students and schools or across all, and whether implementation plans are centralized or localized.

Then, we try the idea at an expanded scale.

Once we reach the time when we agreed we would check in on results, we review what we've learned to see if we can replicate success in additional contexts and over time. We ask questions like:

- How did we do it?
- Is it working? Have we achieved our metrics for success?
- Who is it working for? Does it enhance equity? Has it filled a clear gap or need?
- Have the benefits justified the costs?
- Based on the evidence we see, how broad should this idea's eventual reach be?

Program leadership work with the Research & Program Evaluation Office and/or external evaluators as needed to determine the evidence and findings for each question. At this point, the evidence required to validate spread is more rigorous than at the pilot phase, and may include data collected and analyzed in more rigorous ways. This evidence is then presented to the Senior Leadership Team and Board of Education.

If there is sufficient evidence to determine validate expansion and potential for this idea to work at a greater scale, leadership and the Board would determine yes, this idea can be scaled further. If not, program leadership would determine whether to continue implementing at the prior scale, to adjust the plan for expansion and reevaluate, or to discontinue.

Scaling

The purpose of this step is to determine how we will take this idea to full scale. In this step, Central Office program leaders and Senior Leadership Team should answer questions such as:

- Who is it working for?
- What conditions helped validation succeed?
- Where else do those conditions exist already? Where do we want to create them?
- Can we afford to do this? Think about financial resources, time, and opportunity costs.
- Do we have stakeholder support and excitement?
- What does full scale look like? How far can and/or should this idea spread?
- What evidence will we need to show success at or near full scale?
- When will we check in on results?

The answers to these questions help program leadership decide key elements of expansion, including whether implementation is consistent across schools or school-specific, and whether the idea will extend to all students or some students.

Then, we try the idea at or near its fullest scale.

Once we reach the time when we agreed we would check in on results, we review what we have learned about implementing at scale. We ask questions like:

- How did we do it?
- Is it working? Have we achieved our metrics for success?
- Who is it working for? Does it enhance equity? Has it filled a clear gap or need?
- Have the benefits justified the costs?
- Can positive results be sustained?
- Does our community support it at full scale?
- Are we ready to transfer ownership away from central office leaders and into schools?

Program leadership work with the Research & Program Evaluation Office and/or external evaluators to determine the evidence and findings for each question. At this point, the evidence required to determine success at full scale should be rigorous; this step should have the highest bar of proof among all steps in the process. This evidence is then presented to the Senior Leadership Team and Board of Education.

If there is sufficient evidence to determine success at full scale, leadership and the Board would determine yes, this idea can become standard practice at full scale. If not, the idea would be paused and program leadership would determine whether to continue implementing at the prior scale, to adjust the plan for scaling and reevaluate, or to discontinue.

After an idea reaches full scale, it becomes part of our district's regular routines and practices. We reflect, monitor, and adjust our ways of working on a regular basis as we go through the continuous improvement processes described earlier in this document. Data updates, research reports, and program evaluations are among the types of review we may use as part of continuous improvement.