



Inputs, Outputs, & Outcomes Deep Dive for Staff

Inputs, Outputs, and Outcomes

Inputs are the resources and activities invested in a program or strategy. These are typically known at the beginning of a cycle and measure the effort applied. In a school system, inputs include staff, books, programs, facilities, buses, and everything else for which resources are invested prior to the start of school.

Outputs are the results of a particular set of inputs, usually knowable in the midst of a cycle. They assess or measure the implementation of the program or strategy. In a school system, outputs include formative and interim assessment scores, benchmark results, grades, quarterly financial reports, and everything else that measures implementation of inputs throughout the school cycle.

Outcomes reflect the impact of the program or strategy, typically knowable at the end of a cycle. They are a measure of the effect on the intended beneficiaries. In a school system, outcomes include graduation rates, summative assessment scores, end-of-year audit results, end-of-year retention rates, and everything else that measures the final results of all the inputs and outputs at the end of the school cycle.

Inputs, outputs, and outcomes in schools fall into two categories: **adult** or **student**. For example, adult outcomes measure what adults know or are able to do at the end of the school cycle, while student outcomes measure what students know or are able to do at the end of the school cycle. School systems exist to improve **student outcomes**. All other activities should support or be in service to this.

Example:

- A desired student outcome might be increasing the percentage of students scoring 3 or higher on AP exams. This measures what students know or are able to do by the end of the cycle. Outcomes don't occur by chance; they arise from inputs by staff that generate outputs, which in turn lead to outcomes, and if effective, to improved student outcomes.
- An input that may help lead to this desired outcome is increasing the number of seats available in AP courses. This is an investment of resources at the beginning of the school year.
- Outputs that may lead to this desired outcome might be an increase in the number or percentage of students enrolled in or on track in AP courses. These outputs measure fidelity of the implementation of the input and are knowable throughout the school year.

Details

	Inputs	Outputs	Outcomes
When do they occur?	Beginning of cycle	Middle of cycle	End of cycle
What is their purpose?	Resources and strategies used to accomplish the outcomes	Simultaneously measures input implementation and measures likelihood of outcome achievement	Measure of results that the school system has created
Who selects them?	Superintendent; requires technical expertise	Superintendent; requires technical expertise	Board; clarifies community vision
Which order are they selected?	Select these last; select outputs then select inputs that are likely to create the outputs	Select these second; select outcomes then select outputs that would be predictive of the outcomes	Select outcomes first; always begin with the end in mind; select student outcomes first, then adult outcomes

Understanding Assessment Types, Goal Setting, and Progress Monitoring

Diagnostic, Formative, Interim, & Summative Assessments

Closely related to the concepts of inputs, outputs, and outcomes are the four key types of assessments: **diagnostic, formative, interim, and summative**. While inputs, outputs, and outcomes can refer to any organizational processes, these assessment types specifically focus on evaluating learning and measuring performance. These terms are most commonly used in academic contexts.

- **Diagnostic Assessments** are administered before or at the start of an instructional cycle to measure students' existing knowledge and skills. Typically given at the beginning of the school year, they help educators determine which topics require more or less instructional time. Diagnostic data generally serves as an **input**, guiding instructional decisions for the first semester.
- **Formative Assessments** occur in the midst of an instructional cycle to measure student learning relative to specific expectations. These assessments, often designed by and/or administered on an ongoing basis by teachers or campuses, help refine instructional practice. Formative metrics—data from formative assessments or other input-evaluation tools—generally serve to evaluate and refine the fidelity of implementation. They are considered **outputs** but are not likely to be aligned to summative assessments or metrics.
- **Interim Assessments** take place periodically within an instructional cycle to measure a cross-section of knowledge or skill development. Typically administered two to three times per year by campuses or school systems, they or other outcome-predictive tools serve as predictors of summative performance. Interim metrics track progress at key intervals during a cycle and help forecast final outcomes. Like formative assessments, they are considered **outputs** but, unlike formative metrics, they are explicitly aligned to summative assessments or metrics.
- **Summative Assessments** occur at the end of an instructional cycle to measure a cross-section of cumulative knowledge and skills attained over the course of an instructional cycle or school year. These assessments, administered at the end of a curricular unit, school year, or transition to a new schooling experience by school systems or states and serve the purpose of evaluating school or school system effectiveness. Summative metrics represent final performance data, tracking outcomes and to evaluate whether the school system / chief executive was successful. Summative metrics are generally a type of **outcome** -- and **when measuring what students know or are able to do, are a type of student outcome**.

Norm-referenced & Criterion-referenced

Another set of assessment-related concepts is the distinction between norm-referenced and criterion-referenced. These are two different ways of interpreting what the assigned rating will be after the raw scores are collected.

- **Criterion-Referenced Assessments** measure student performance against a fixed set of standards. Every student who meets or exceeds a predetermined threshold receives the same rating. For example, if a score of 9 out of 10 (90%) qualifies as an "A," all students scoring 9/10 earn an "A." These assessments are ideal for determining whether students have mastered specific knowledge or skills.
- **Norm-Referenced Assessments** compare student performance against that of their peers. A student's rating depends on how their score ranks relative to others. For example, if a student improves by five points but their peers improve by ten, their relative standing—and possibly their assigned rating—may decline. These assessments are useful for identifying where students stand in relation to their peers rather than measuring absolute mastery.

Proficiency, Growth, & Comparison

Another critical distinction in assessment is how expectations are constructed. Expectations can be based on **proficiency, growth, or comparison**, each serving a unique purpose in evaluating student progress.

- **Proficiency** defines expectations relative to a predefined competency level. It asks whether a student meets a specific standard, often expressed as being "on grade level" or "passing."
 - *Example:* "The percentage of 3rd grade students demonstrating grade-level proficiency in reading will increase from 40% in June 2021 to 70% by June 2026."
- **Growth** defines expectations relative to a student's starting point. It measures improvement over time, often expressed as "grew by more than a grade level" or "demonstrated expected growth."
 - *Example:* "The percentage of 3rd grade students who began the year below grade level in reading and grew by at least 1.25 grade levels will increase from 40% in June 2021 to 70% by June 2026."

- **Comparison** defines expectations relative to a control group. It assesses whether a student has made progress in relation to peers, often framed as “closing the gap.”
 - *Example:* “The percentage of 3rd grade students who were two grade levels behind in reading and closed the gap between themselves and the overall 3rd grade average will increase from 40% in June 2021 to 70% by June 2026.”

Static Goals & Cohort Goals

When setting goals, a key decision is whether to track a specific group of students over time (**cohort goals**) or to measure a fixed indicator over time with different students (**static goals**). By default, most people unknowingly use static goals—just as they often default to proficiency-based goals or assume assessments are criterion-referenced unless prompted to consider alternatives.

- **Static Goals** track overall system improvement by measuring different students using the same assessment over time.
 - *Benefit:* Provides insights into how a specific area of the organization is improving year over year.
- **Cohort Goals** track a defined group of students as they progress through the system, using different assessments at each stage.
 - *Benefit:* Offers a clearer picture of how individual students grow over time.

Milestones, Interim Goals / Interim Guardrails, & Goals / Guardrails

Closely related to the concepts of formative, interim, and summative are the concepts of **milestones, interim goals / interim guardrails, and goals / guardrails**. Where formative, interim, and summative focus specifically on methods of assessment and associated metrics, milestones, interim goals / interim guardrails, and goals / guardrails focus specifically on **project management** and **progress monitoring**. These terms are most commonly used in managerial or governance contexts.

- **Milestones** are dates by when specific inputs are implemented -- by when project deliverables are due. They describe and track staff actions and serve as a **management tool** for monitoring implementation progress. Milestones are **set by managerial teams** and are not governance instruments.
- **Interim goals and interim guardrails** are measures of progress toward a defined goal or guardrail. They are **output metrics** that are **predictive** of the goals / guardrails and that are **influenceable** by the chief executive. Interim goals and interim guardrails are **defined by the chief executive, in consultation with the board**, and are the data the chief executive uses to routinely report on performance to the board.
- **Goals and guardrails** are **policy statements** created **by the board, in consultation with the chief executive**, that are intended to capture the vision and values, respectively, of the community that the board represents.
 - **Goals** are **SMART** (Specific, Measurable, Attainable, Results-focused, Time-bound)
 - **Guardrails** set boundaries but are **not SMART**

Selecting Goal / Interim Goal Targets

When determining **targets for goals and interim goals**, the management team should evaluate at least **three key things**:

- **5-Year Average Growth:**
 - Identify the **previous five years'** summative percentage of year-over-year growth for the targeted student outcome.
 - Calculate the **average** those five numbers to identify the **minimum threshold** of average growth (the number to beat).
 - This analysis assumes a static trendline. If external circumstances have significantly disrupted past trends, adjustments should be made accordingly.
 - **Key Principle:** Targets should **exceed** historical performance, not merely match it.
- **Degree of Difficulty:**
 - Different tasks require different levels of effort, and this should inform target-setting.
 - **General assumptions:**

- Learning English is generally harder than learning math.
 - Younger students are easier to catch up than older students.
 - **Early gains** (e.g., moving from 0% to 10% proficiency) are easier than **closing final gaps** (e.g., moving from 90% to 100%).
 - Maintaining grade-level proficiency is easier than helping students who are far behind catch up.
- **Guiding Principle:** The more challenging the task, the **smaller** the target; the easier the task, the **larger** the target.
- **Resource Commitment:**
 - Key Assumptions:
 - The **average growth over the previous five years** is for a given student outcome is the **growth baseline**.
 - The **amount invested in strategies explicitly focused** on improving that student outcome as the **investment baseline**.
 - **Scaling Rule:** The level of **desired improvement** above the baseline growth will determine the **required level of new investment** -- the multiple times investment baseline.
 - Example: If current average growth is 1% per year with a \$100 investment, achieving growth of 4% per year would likely require a \$400 investment.
 - If there is a **lack of willingness or capacity** to invest at least that amount to improve performance, then the desired target is too high and should be adjusted accordingly.

Questions to Ask When Evaluating Interim Metrics

Knowing whether or not interim metrics are sufficiently meaningful is always challenging. Here is a set of questions that staff, board members, and coaches should review prior to their being accepted:

- Interim Metric Requirements
 - Are all of the interim metrics -- for both goals and guardrails -- SMART? This means they each have:
 - A starting point and an ending point
 - A starting date and an ending date
 - A specific measurement
 - In the case of interim goals, a specific student population
 - Are the interim metrics leading indicators relative to the goal/guardrail rather than lagging?
 - Are there exactly three interim metrics for each goal and for each guardrail?
 - If all three of the interim metrics are accomplished, will that likely ensure accomplishment of the goal?
 - Are all of the interim metrics influenceable by the superintendent? Does the superintendent have at least 80% authority over the implementation of the items being measured?
 - Are each of the interim metrics updateable multiple times per year?
 - Are each of the interim metrics outputs (mid-cycle results, knowable in the midst of the cycle) rather than inputs (resources/strategies, knowable at the beginning of the cycle)?
- Interim Metric Considerations
 - What is the degree of correlation between the interim metric and the goal metric?
 - Can the data be monitored on the school board's monitoring calendar within 30 to 60 days of when it is collected?
 - Do any of the interim metrics utilize data sourced from external to the school system (like CGCS' academic KPIs or operational KPIs)?
 - Do any of the interim metrics rely on data that is historically unreliable or highly variable between individual schools?
 - Do the interim metrics fully address the content of the goal/guardrail?
 - Are each of the interim metrics strong outputs ("late outputs" - measures of implementation quality) rather than weak outputs ("early outputs" - measures of participation)?
 - Are there significant unintended consequences that need to be considered?
 - Does each interim metric have only one data set rather than multiple?
 - Has the management team evaluated each interim metric for unintended consequences?
 - Is the metric predictive both downward and upward?

- Is the interim metric the data that management actually uses for ongoing decision making?
- Is there curriculum that explicitly teaches the material covered by the interim goals and does evidence exist that the curriculum is being implemented in schools?
- Do all of the interim metrics appear as part of the cabinet members' annual evaluations? (see [Cascading Alignment - Board to Cabinet.docx](#))

Upward vs. Downward Predictability

Many metrics—especially those related to **adult inputs and outputs**—may appear **predictive** because a decline in these measures often correlates with a decline in student outcomes. However, this does **not** mean that increasing them will necessarily lead to improved performance at the same rate.

Example:

- If **3rd graders are deprived of food**, their caloric intake will decline, and their academic performance will likely suffer.
- This suggests two potential **interim metrics: calories provided and calories consumed**—both of which would correlate with a decline in student outcomes if reduced.
- However, **doubling** a student's caloric intake **will not** double their academic performance.

Key Insight:

- A metric that is **predictive downward** (i.e., a decline correlates with failure) is **not necessarily predictive upward** (i.e., an increase leads to success).
- This is why **interim goals must focus on student outputs** rather than adult inputs, adult outputs, or even student inputs.
- While these other factors are often **necessary**, they are **not sufficient** for improving student outcomes.
- **Downward predictability is common, but upward predictability is rare.**

Making Monitoring Meaningful

Effective monitoring should provide insights that **management actively uses** to drive decision-making. If monitoring reports:

- **Do not align** with data management routinely uses, they become meaningless.
- Use data that is not actively used **only to satisfy school board requests or for the sake of compliance with this process**, they are **worthless**.
- Require management to **create new reports from scratch** each month over a long period, this reflects either:
 - A **fundamental misunderstanding** of continuous improvement at best, or
 - **Gross incompetence** in managerial leadership at worst.

Key Principle:

Monitoring should **reveal and reinforce** the decision-making process—not be an isolated bureaucratic exercise. *(See more about Effective Monitoring beginning on page 13 of this document.)*

Is Assessment Harmful to Students?

Categorically, no. Assessment, like any tool—whether a screwdriver, level, or pliers—is essential for effective construction (in this case, constructing knowledge). While it can certainly be abused, when used appropriately, it is a critical part of the learning process. To claim that all assessment is inherently harmful is not only misinformed, but it also fosters beliefs that can be detrimental to students.

Assessments serve several vital roles in education, but two of the most crucial **are long-term memory creation and instructional practice improvement**. When used responsibly, assessments help solidify learning and drive continuous improvement in teaching, benefiting students overall.

Assessment: Long Term Memory Creation

Part of the neurological science of learning involves students migrating information out of short-term working memory and into long term memory. One strategy for supporting this migration is assessment.

- When students are taught something on Monday of week 1 and then quizzed on it on Friday of week 1 (this is usually formative assessment), the process of migration begins.
- When the same material is assessed again at the end of the semester period (this is often interim assessment), the process continues and strengthens.
- When students are assessed on the same material again at the end of the year, requiring them to recall information learned months earlier (this is usually summative assessment), the associations with that learning are strengthened and reinforced in the brain's physical structures.

By incorporating **appropriately designed and spaced assessments**, educators can actively support the **strengthening and retention of knowledge**, helping students move from **initial exposure to long-term mastery**.

Assessment: Instructional Practice Improvement

The heart of effective education often lies in the dynamic between the learner and the educator. Research consistently shows that, among all the factors within a school system's control, instructional quality has the greatest impact on student performance. As the quality of instruction improves, student outcomes are more likely to follow suit.

One of the most reliable methods for teachers to enhance their instructional quality is through a cycle of teaching, assessing, reteaching, and reassessing. This process of continuous improvement forces teachers to regularly evaluate what worked, what didn't, and adjust accordingly. However, this cycle hinges on having assessment data—usually in the form of formative assessments—to inform decision-making.

Improving instructional practices is crucial for improving student outcomes, but without assessments, particularly formative ones, the process of improvement is nearly impossible.

Standardized Assessment: Performance Transparency

A commonly asked question is whether or not parents have the right to know how their child's school is performing compared to other schools.

- If the answer is "no", parents don't have the right to understand the relative performance of their child's school, then standardized assessment is less necessary.
- If the answer is "yes", parents should have the right, then standardized assessment is essential.

Interrater reliability is another issue. If the math teacher at one school gives students A's for the same work that the math teacher at another school gives her students B's, letter grades lose their reliability as indicators of student performance. One way to ensure fairness is through a common assessment across schools, scored using a consistent rubric, with teachers receiving regular training to apply these standards accurately. In other words, achieving fairness requires standardization.

Performance transparency isn't just important in K-12 education—it is fundamental in many professions. Fields such as medicine, engineering, construction, and technology use standardized assessments to ensure fairness and uphold professional standards. These measures help protect public safety and maintain quality—whether by ensuring doctors administer the correct treatments, engineers design structurally sound bridges, or IT professionals secure organizational data.

When performance transparency across multiple sites matters, well-designed standardized assessment is an essential and appropriate tool for ensuring fairness and accountability.

Standardized Assessment: System Practice Improvement

Transparency is just one benefit of standardized data. It also plays a critical role in helping complex systems identify outliers and make necessary adjustments. However, if the data isn't consistent—apples to apples—it becomes difficult to separate noise from signal and make informed decisions.

It's important to recognize that standardized assessments can be formative, interim, or summative, and the type of assessment significantly affects how the data should be used. For example:

- It doesn't make sense to use formative data to predict performance on a summative assessment.

- Likewise, it doesn't make sense to use summative data to inform weekly adjustments in teaching.

In general, system-level improvements are more effectively driven by aggregated interim and summative data than by formative data, as they provide a more comprehensive picture of long-term trends and performance.

Is Standardized Assessment Biased?

All assessments, whether standardized or not, carry inherent biases. Every assessment is shaped by the creator's worldview, language, and understanding of the material. Rather than asking, "Are standardized assessments biased?" the more important questions are:

- Is the bias harmful to the accuracy of the data collected?
- What measures have been taken to identify and mitigate harmful bias?

For example, consider this question from a 3rd-grade math assessment:

If 300 crayons are added to 100 crayons, how many crayons are there?

When asked, "Is this question biased?" many people might say no, but that is the wrong question. A better question is, "For whom is this question biased, and does that bias harm the collection of accurate data?"

If the students being assessed have no familiarity with crayons—such as asylees or children experiencing trauma—this question could certainly be biased. Does that bias harm the accuracy of the data? The answer depends on the context. For the general population, probably not, as the question only requires counting, not familiarity with crayons. However, for asylees or children unfamiliar with crayons, the bias could interfere with their performance, making it a valid concern for exclusion in certain contexts.

It is also essential to ensure that assessments are thoroughly vetted for bias. Each item should be reviewed by multiple teams to identify inappropriate elements, whether they concern the grade level of the material, language complexity, developmental appropriateness, or bias.

While all assessments have the potential for harmful bias, standardized assessments—when well-designed—are better equipped to minimize bias, thanks to their structured administration process, which is more challenging to achieve in non-standardized assessments.

Misuse of Assessment Data

The simple answer to the misuse of assessment data is using it for purposes it wasn't designed for. Academic assessments are not meant to define a child's worth or value. Anyone who asserts otherwise is misusing standardized assessment data. Another common misuse occurs when standardized summative assessment data is treated as if it were formative assessment data, or vice versa.

An important consideration when using data is the type of decisions being made. Standardized assessment data can be useful for resource allocation and assessing how well the school system is serving specific student groups. However, since standardized assessment data can be influenced by socioeconomic factors, it becomes problematic if it is used to limit students' access to opportunities—such as access to courses, grade levels, or institutions.

Item Types and Their Impact

Another area of concern involves how knowledge and skills are assessed. There are generally three types of assessment items: **selected response, constructed response, and performance tasks**. Most people think of standardized assessments in terms of the traditional Scantron sheets, which are selected response items. While efficient, these assessments are limited in the insights they provide.

Constructed response items ask students to generate their own answers, while performance tasks require students to demonstrate their knowledge or skills through activities. Both types often rely on a rubric for evaluation. Portfolio assessments are similar to performance tasks, as they involve evaluating a collection of the student's previous work. These assessments can be part of both standardized and non-standardized assessments.

The way assessments are designed is meaningful because it influences how educators design and deliver instruction. If a standardized summative assessment relies solely on selected response items, there is a tendency for educators to adopt similar formats for formative and interim assessments. Education leaders are increasingly recognizing this issue and are

shifting toward assessments that include more constructed response items, performance tasks, and portfolio assessments to better support comprehensive, student-centered teaching and learning.

Does Constant Test Prepping Improve Performance on Standardized Assessments

No. This is a myth perpetuated by, not surprisingly, the test prep industry. What does support improved performance is a small amount of preparation regarding the item types that will be used in the assessment and the general design of the assessment. This familiarity helps a test taker be better positioned to demonstrate what they know on the assessment. That's a very good thing. But it only takes a few hours to do this. Endless test prepping that begins weeks and months before an assessment is time wasted and is a practice that is actively harmful to children. What improves performance on standardized assessments? When students are taught the assessed material to a depth of knowledge that allows them to use the material in new ways. There are no shortcuts.

Project Management vs Progress Monitoring

Project management is about tracking the inputs -- the day-to-day actions -- of staff to determine whether they are likely to deliver the desired outputs. Project management is asking if the tactics being used to accomplish the strategy are being implemented effectively such that the milestones will be met. A common result of effective project management practices is revision of implementation plans and practices. Project management is a management duty, not a governance duty.

Progress Monitoring is about tracking the outputs -- the interim metrics -- of staff to determine whether they are likely to deliver the desired outcomes/results/goals/guardrails. Progress Monitoring is asking if the strategy being used to accomplish the results is actually likely to do so. A common result of effective progress monitoring is revision of organizational strategy. Progress monitoring happens at both the managerial and governance levels (though with significant differences in scope).

Arguments Against Measurement

Often, people will make arguments against measuring anything. While this position, taken to the extreme, is harmful and undermines our ability to serve our most educationally vulnerable children, the underlying pain behind it is not without merit. Two similar ideas summarize the strongest argument against measurement -- and simultaneously animate several of our most ardent recommendations regarding the composition of effective metrics.

- Campbell's Law states, "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."
- Goodhart's Law states, "Any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes."

In short: the moment you start measuring something -- and particularly, once meaning or consequence is attached to the measure -- the utility of the data collected by the measure begins to decline. THIS ISN'T ABOUT CHEATING — that's a separate phenomenon entirely and is a function of poor leadership, not assessment. Nor is it about teaching to the test. It's about the collapse of meaning after a while. These are sociologically and statistically observable and naturally occurring functions of measurement, not arguments for destroying the unbelievable good for children that standardized assessments can provide. While these collapsing effects are not instantaneous, they are real and warrant project management and performance management systems that are calibrated to fight against these tendencies.

Key strategies built into effective systems to combat these realities include changing interim metrics every one to three years, relying on strong outputs (for example, "how many students are on-track") for interim metrics which are much harder to game than inputs (for example, "how many of x did we do") or weak outputs (for example, "how many people attended/participated"), and relying on metrics that are already explicitly used for decision making rather than novel data that no one actually uses or tracks.

Goals / Summative Metrics

- **Definition:** Quantifiable outcome measurement of achieving/honoring your school system's vision or values
- **Function:** Used to track outcomes; whether the school system/chief executive was successful
- **Guiding Question:** What would honoring the vision and values of our community look like for this school system?
- **Formula:** The [measure] for [population/area] will [increase/decrease] from [starting point] on [starting month/year] to [ending point] by [ending month/year]
- **Examples:**
 - Number of high performing campuses as measured by the School Performance Framework will increase from W% on X [month/year] to Y% by Z [month/year]
 - Percentage of graduates persisting in their second year post-secondary without needing remedial courses will increase from W% on X [month/year] to Y% by Z [month/year]
 - Percentage of graduates having completed an associate's degree and/or been awarded an industry certification by graduation will grow from W% on X [month/year] to Y% by Z [month/year]
 - Percentage of students reading on grade level according to ABC instrument by the end of 3rd grade will increase from W% during X [month/year] to Y% during Z [month/year]

Effective Summative Metrics...	Moderate Summative Metrics...	Ineffective Summative Metrics...
✓ Are quantifiable (SMART), and the data needed to measure can be reasonably collected. ✓ Clearly align with the vision and values of the community. ✓ Are developed with baseline data and indicate an important improvement in student outcomes. ✓ Include the starting date (month and year) and the ending date (month and year) by which the metric will be achieved. ✓ Include a starting point (baseline) and ending point (target) ✓ Will challenge and disrupt the organization and require adult behavior change. ✓ Were developed through analysis of student needs. ✓ Community members were included in the development process. ✓ Where feasible, rely on external data rather than data produced by the school system's staff. ✓ The data is collected using a robust system for ensuring data integrity.	□ Includes vague markers for success (e.g. "more," "better," "large amounts") and data cannot be collected within a reasonable time □ Unclear alignment with the vision and values of the community □ Not developed with baseline data and/or, if achieved, will not have a significant impact on students □ Can be accomplished by only slightly tweaking what the organization is currently doing □ Include a date, but it is incorrect, too broad, or not aligned with the vision and values of the community	○ Cannot be measured. ○ Do not indicate overall success of the initiative, or are actually formative metrics or interim metrics ○ Are not ambitious or inspirational ○ Can be accomplished without changing anything the organization did last cycle ○ No date included

Interim Goal / Interim Guardrail / Interim Metrics

- **Definition:** Quantifiable output measurements of achievement that indicate progress towards the Summative Metric.
- **Function:** Used to track outputs; quantifiable progress at critical and ongoing junctures during a cycle
- **Guiding Question:** How will I measure whether or not I am on track to reach my Summative Metric?
- **Formula:** The [measure] for [population/area] will [increase/decrease] from [starting point] on [starting month/year] to [ending point] by [ending month/year]
- **Examples:**
 - The percentage of students on track in reading will increase from W% on X [month/year] date to Y% on Z date [month/year], as measured by district benchmark assessments
 - The percentage of 9th graders on track to graduate will grow from W% in X [month/year] to Y% [month/year] by Z as measured by the district's ABC on-track indicator

Effective Interim Metrics...	Moderate Interim Metrics...	Ineffective Interim Metrics...
✓ Are quantifiable (SMART), and the data needed to measure can be reasonably collected. ✓ Lead directly to the achievement of the Summative Metric. ✓ Are ambitious, but can be completed within the scope of the goal, and will need to be tracked on an ongoing basis in order to ensure overall success. ✓ Refreshed data is available multiple times per year ✓ Include the starting date (month and year) and the ending date (month and year) by which the metric will be achieved. ✓ Include a starting point (baseline) and ending point (target) ✓ Will challenge and disrupt the organization and require adult behavior change. ✓ Are predictive of the Summative Metric ✓ Are influenceable by the chief executive ✓ Where feasible, rely on external data rather than data produced by the school system's staff. ✓ The data is collected using a robust system for ensuring data integrity.	□ Include vague markers for success (e.g. "more," "better," "large amounts") □ Do not account for all aspects necessary to goal/guardrail success OR are not directly tied to overall goal/guardrail success. □ Can be accomplished by only slightly tweaking what the organization is currently doing □ Indicate trivial achievements or do not necessarily relate to the overall goal/guardrail success.	○ Cannot be measured. ○ Do not align with nor lead directly to the Summative Metric. ○ Can be accomplished without changing anything the organization did last cycle ○ Do not indicate ambitious achievements or require little to no effort. ○ Community opinions were a significant influence in the development process.

Goal Milestones / Guardrail Milestones / Formative Metrics

- **Definition:** Quantifiable output measurements that indicate fidelity of implementation of deliverables -- and ideally, though not necessarily, progress toward an Interim Metric.
- **Function:** Used to track fidelity of implementation; quantifiable progress at critical and ongoing junctures during a cycle
- **Guiding Question:** How will I measure whether or not I am faithfully implementing the chosen tactics?
- **Formula:** The [measure] for [population/area] will [increase/decrease] from [starting point] on [starting date] to [ending point] by [ending date]
- **Examples:**
 - The number of AP seats available will increase from W% on X date to Y% on Z date
 - The percentage of 9th graders with weekly counselor check-ins will grow from W% in X to Y% by Z

Effective Formative Metrics...	Moderate Formative Metrics...	Ineffective Formative Metrics...
✓ Are quantifiable (SMART), and the data needed to measure can be reasonably collected. ✓ Lead directly to the completion of the milestones/deliverables. ✓ Are ambitious, but can be completed within the scope of the initiative, and will need to be tracked on an ongoing basis in order to ensure overall success. ✓ Refreshed data is available multiple times per month. ✓ Include the starting date and ending date by which the milestone will be achieved. ✓ Include a starting point (baseline) and ending point (target)	□ Include vague markers for success (e.g. "more," "better," "large amounts") □ Do not account for all aspects necessary to initiative success OR are not directly tied to overall deliverable quality. □ Can be accomplished by only slightly tweaking what the organization is currently doing □ Indicate trivial achievements or do not necessarily relate to the overall timeliness of the milestones or quality of the deliverables.	○ Cannot be measured. ○ Do not align with nor lead directly to the timeliness of the milestones or quality of the deliverables. ○ Can be accomplished without changing anything the organization did last cycle ○ Do not indicate ambitious achievements or require little to no effort.

Project Management

- **Definition:** Process of tracking the extent to which the agreed upon tactics are being implemented with fidelity and milestones are being met
- **Practitioners:** This should only be conducted by the level of staff responsible for managerial implementation of a tactic and their supervisors; this is a managerial function, not a governance function
- **Frequency:** Project management should be practiced no less than once per month (though it is most commonly a weekly practice) and no more than the frequency with which implementation fidelity data is refreshed.
- **Function:** Used to track formative metrics (inputs).
- **Guiding Question:** How will I know whether or not I am on track to reach my Formative Metric?

Blue	All project deliverables have been completed and are of quality
Green	All milestones are on track to be completed on time and project deliverables are of quality
Yellow	Some project deliverables are not on track to be completed either on time or are not of quality
Orange	Without changes, some project deliverables are not on track to be completed
Red	Without significant changes, most project deliverables are off track and/or will not be completed
Gray	Project has not been started yet

Progress Monitoring

- **Definition:** Process of tracking progress regarding the extent to which the agreed upon results are likely to occur
- **Practitioners:** This should be conducted by staff responsible for managerial creation of results and their supervisors; this is both a managerial function (when it occurs between staff and their supervisors), and a governance function (when it occurs between the board and the chief executive)
- **Frequency:** Progress monitoring should be practiced no less than once per quarter (though it is most commonly a monthly practice) and no more than the frequency with which lead measure/interim measure data is refreshed.
- **Function:** Used to track interim metrics (outputs)
- **Guiding Question:** How will I know whether or not I am on track to reach my Summative Metric?

Blue	Results were completely delivered
Green	All interim metrics indicate that the results will be delivered
Yellow	Interim metrics indicate conflicting evidence regarding whether results will be delivered
Orange	Interim metrics indicate results are unlikely to be delivered without significant changes
Red	Interim metrics indicate results will not be delivered
Gray	Interim metrics are not available yet

Effective Goal Monitoring

Measuring Progress

Once SMART goals about student outcomes have been adopted, effective goal monitoring requires four main ingredients: monitoring calendar, monitoring report, superintendent participation, and board member participation.

Effective Monitoring Calendars

Before boards can begin effective monitoring, they should adopt a 36-60-month schedule that describes which goals will be monitored during which month. The board will typically have the superintendent draft a calendar since the administration knows when student performance data is freshly available throughout the year. Nevertheless, it remains the board's monitoring calendar, not the superintendents. Qualities to look for include:

- Spans the entire length of the goals. If the goals are five years long, the calendar should also be five years long
- Includes all of the board's goals and guardrails
- Often includes all board trainings, board-led community trainings, board-led community listening sessions, board self-evaluations, board-led superintendent evaluations, and statutory votes
- Schedules each goal to be monitored at least four times throughout the year, and each guardrail at least one time per year (on 12-month calendar)
- Schedules one or two interim goals to be monitored each month, no less and definitely no more than three
- Can schedule as many interim guardrails to be monitored during a month as the board wants
- Never suggests that goal monitoring reports be placed on the consent agenda, but guardrail monitoring reports may be on consent
- Clarifies that boards will monitor goals during every month of the year that the board meets

It's often helpful for the board to put all of its monitoring and non-monitoring items on a single calendar so the board's priorities are easily tracked while also keeping other obligations of the board in perspective. Such a calendar can still typically be 36-60 months in span given the cyclical nature of district activities. Here's an example:

	Goal Monitoring	Guardrail Monitoring	Leadership PD & Evaluation	Statutory Obligations	Out-of-Meeting Reports
January					
February					
March					
April					
May					
June					
July					
August					
September					
October					
November					
December					

Effective Monitoring Reports

Below are four qualities to ask about the **1-5-page monitoring report** before the board can begin progress monitoring. If the answer to any of these questions is “no”, hand the report back to the Superintendent and have them complete it before proceeding -- likely at the next regularly scheduled board meeting.

1. **The Goal:** Does it clearly show which specific goal / interim goal is being monitored?
2. **The Data:** Does it clearly show data for the 3 previous reporting periods (preferably on a line graph)? Does it clearly show the current reporting period? Does it clearly show the target reporting periods (annual targets and deadline target)?
3. **The Interpretation:** Does it clearly show the Superintendent’s understanding of system performance relative to the goal?
4. **The Evidence & Plan:** Does it clearly show supporting documentation that evidences the Superintendent’s understanding of system performance? If the school system is not at target or the Superintendent’s understanding of system performance indicates implementation is not on track, does the monitoring report clearly describe systemic root causes, strategic responses (including rationale), and any needed next steps?

Effective Superintendent Participation

How superintendents show up in the monitoring conversation has a huge impact on the conversation’s effectiveness. Some guidelines include:

- **Don’t Hide the Data:** The student performance data being presented during the monitoring conversation should be easy for most parents to understand. As such, monitoring reports should be only 1-5 pages at most, and should be written at no more than an 6th grade reading level.
- **Don’t Sugar Coat the Data:** The data is the data. Whatever it says is what it says -- good, bad, or ugly. Never suggest that the data is saying anything other than what you believe it to be saying. If the school system is off track, say that; don’t talk around that. Sugarcoating loses trust.
- **Align Monitoring with Managerial Action:** Data in monitoring report should reflect what staff are looking at to gauge the district's effectiveness. There should be no need to create data for a monitoring session that isn't otherwise being considered by the superintendent and cabinet.
- **Be Prepared:** Many superintendents rehearse for monitoring conversations by having their teams throw every conceivable question at them before the board meeting. This is a wise practice not only because it helps with the monitoring conversation but because it can help surface managerial issues and solutions that might not otherwise come up.
- **Don’t Be Defensive:** If the student performance data is disappointing, then it’s natural that board members would be disappointed. Unfortunately, not all of them will manage their disappointment in a mature, adult, and effective manner. Even if this happens, don’t get defensive.

Effective Board Member Participation

Goal monitoring, like board governance in general, is not always intuitive. It is easy to inadvertently conduct monitoring in an ineffective manner. Here are a few guidelines to follow to increase the likelihood of effectiveness:

- **Do Your Homework:** Board members should arrive at board meetings having already read the monitoring report, having already shared technical and tactical questions with the superintendent, and having already come up with at least 3 or 4 [SMART Questions](#) each about the monitoring report (see *During Goal Monitoring* below).
- **Understanding Reality:** The desired result of monitoring is to understand the current reality for your students as compared to the vision you’ve adopted for them (goals). Whether you enjoy the current reality isn’t the point of monitoring; whether or not you fully know the current reality is.
- **Keep the Conversation Going:** If the superintendent presents a monitoring report that is missing the prerequisites (see *Before Goal Monitoring* above) or that fails to clarify for board members the extent to which reality matches the goals, consider tabling the conversation and giving the superintendent a chance to fix it and re-offer it at a subsequent meeting, instead of choosing not to accept it and ending the discussion.
- **No Gotcha Governance:** Adopt a monitoring calendar that shows which goals will be monitored during which months and that spans the full term of the goals -- for 5-year goals, the calendar should be 5 years. Then ensure board members adhere to the monitoring conversation rubric below.
- **Don’t Offer Advice:** Monitoring is never an opportunity for board members to provide advice to the superintendent regarding what should/shouldn’t be done about student outcomes. It’s also not about liking/not liking the superintendent’s strategies.

DURING GOAL MONITORING

Monitoring is about understanding the extent to which reality matches the Board's adopted goals / interim goals. Monitoring is never about offering advice or recommendations; most of monitoring is about understanding where we are and how we got here. The Board's attention is focused on what's true for students, not on what adults are/aren't doing. Here are observations to look for / questions to ask that support effective progress monitoring. Notice that none of these questions offer advice concerning which inputs/outputs the Superintendent should select; these are SMART monitoring questions, not managing questions. (see [Effective Question Asking](#))

Current Performance Questions			Future Performance Questions
What do we know about the students mentioned in the report?	What do we know about the data mentioned in the report?	What do we know about the root cause of the student data in the report?	What adult behaviors need to change in response to the student data?
- Who is struggling the most? - Who is growing the most? - Who is not moving? - Which students are not included in this data?	- What is currently happening? - What else do we need to know about this? - How and what did we learn? - What are the strengths? - What are the limitations? - Where do gaps exist between student groups? - What's working? Not working? - What do you see as accounting for <anomalous data in report>?	- Why is it working in this area? - Why is it not working in this area? - How did we learn about this issue? - Why such significant growth? - Why was there no growth? - What do we need to know about? - Why do gaps between student groups exist? - Why is <data point a> so much <higher or lower> than <data point b>?	- How can we replicate what is happening in ___? - Given what we know about ___, what are you going to do to speed up the progress? - What evidence suggests that your new strategy is going to work? - How are we going to address ___ (issue not resolved)? - How might changes show up in the future (budget, etc.)? - How can the board help?

Ineffective Questions
- Any statements or questions that are really just board member opinions or recommendations about what the superintendent should do - Any statements or questions that don't reference the data mentioned in the monitoring report. - Any statements or questions that aren't actually questions but that are just statements or opinions - Any statements or questions about what will happen next that aren't grounded in previously asked questions about where students currently are and how students got there

AFTER GOAL MONITORING

To Accept the Monitoring Report or Not?

Once the board has completed the monitoring conversation, it must choose whether to accept or not accept the report based on three questions: 1) does reality match the vision, 2) is there growth toward the vision, and 3) is there a strategy and plan sufficient to cause growth toward the vision?

- If the answer to all three is yes, then the board can accept the monitoring report confident that data is accurate & the superintendent is performing.
- If the answer to only one or two of these questions is yes, the board may opt to table the matter (see *Keep the Conversation Going* above).
- If the answer to all three is no, the board should consider voting to not accept the report. Note: This vote informs the superintendent that they have failed to meet the expectations of monitoring.

To Change Goals or Not?

Once the board has completed the monitoring conversation, it's also appropriate -- though not required -- to ask whether or not the goal is still an appropriate representation of the community's vision for what students should know and be able to do. This inquiry should not be taken lightly; goal monitoring is most effective when the goals only change after their term has expired or they are accomplished. Frequently changing goals makes it almost impossible to adequately assess superintendent performance and to avoid wasting school system resources. If the board believes that the goal is no longer appropriate, it should create a plan to start the board-led community listening process over again and then begin the community listening process from scratch. Because goals represent the vision of the community, no new goals should be adopted without first going through this process.

Evaluating the Quality of Goal Monitoring

While the board is in the process of conducting the monitoring conversation:

1. Use the "Evaluation Rubric" to evaluate every individual question on its SMARTness: Strategy, Measure, Ask-oriented, Results, Time-bound.
2. Use the "Tally Sheet" below to track whether each individual question is focused or not ("Yes" or "No"). This will provide data that describes the percentage of all of the questions that are focused ("% Focused") on each of the SMART characteristics.
3. Average the individual ratings for the SMART characteristics together to get the overall rating of monitoring quality. Some behaviors -- ineffective monitoring practices, not being prepared in advance, or not participating -- will give automatic zeroes.
4. Total Monitoring Scores of 90 and above indicate highly effective monitoring, 80-89 indicate effective, 70-79 indicate approaching effective, and 69 or less indicate ineffectiveness. (see [Effective Goal Monitoring spreadsheet](#); it automatically performs these steps)

Monitoring Conversation Evaluation Rubric			
	Ineffective Monitoring		Effective Monitoring
Strategy-Focused	Any conversation is focused on or offering advice about technical or tactical issues. (see Technical/Tactical/Strategic)		Is the question about strategic issues rather than technical or tactical issues?
Measure-Focused	Any comments are focused on data not in the report.		Does the question reference specific metrics/data that has been provided at the request of the board?
Ask-Oriented	Any conversation is focused on accusatory yes/no questions or statements.		Is the question open ended?
Results-Focused	Any comments are focused on blaming or shaming.		Is the question focused on understanding data rather than sharing opinions?
Time-Bound	Any conversation is offering advice about future action.		Is the question focused on current performance (past actions) rather than future performance (future actions)?
	Automatic 0%		
Preparation & Participation	Read: The monitoring conversation is automatically scored 0% if any Board members have not completely read any of the monitoring report prior to the monitoring conversation.	Participate: The monitoring conversation is automatically scored 0% if there is non-participation by any board member present during the monitoring conversation.	Share: The monitoring conversation is automatically scored 0% if any board members failed to share questions with the Superintendent at least three working days prior to the monitoring conversation.

Monitoring Conversation Tally Sheet (or use Effective Goal Monitoring spreadsheet)												
Strategy-Focused		Measure-Focused		Ask-Oriented		Results-Focused		Time-Bound		Preparation & Participation		
# Yes:	# No:	# Yes:	# No:	# Yes:	# No:	# Yes:	# No:	# Yes:	# No:	Read?:	Part?:	Share?:
% Focused:		% Focused:		% Focused:		% Focused:		% Focused:				