|  |  |
| --- | --- |
|  | **District Assessment System Design Toolkit**  *Project Report* |
| *Version 0.9 (beta) released June 29, 2018* |

**Developed By**

|  |  |  |
| --- | --- | --- |
| **Joseph Martineau** | Senior Associate |  |
| **Kathy Dewsbury-White**  **Ed Roeber** | President & Chief Executive Officer  Assessment Director |  |
| **Ellen Vorenkamp** | Assessment Consultant |  |
| **Steven Snead**  **Jonathan Flukes** | Supervisor, Curriculum & Instruction Unit  Research, Evaluation, & Assessment Consultant |  |

|  |  |
| --- | --- |
|  | **Anytown Public Schools**  **District Assessment System Design Report** |
| Anna Principal  Independent Consultant  June 29, 2018 |

# Executive Summary

Anytown Public Schools made a decision to invest in examining its existing assessment system and designing a revised system. The district did so by selecting a team of district administrators, school administrators, and classroom teachers from all levels of schools in the district (elementary, middle, and high school) to engage in a principled process carried out in three workshops of six hours each. The workshops were conducted at the Anytown Regional Service Agency on three dates in January of 2018.

## Why the Work is Important

Assessment is a key component of education, supporting instruction, instructional programming, program and policy evaluation, placement, eligibility for programs, grading, and eligibility for honors or awards. However, it is rare that there is a carefully constructed system that is able to provide each stakeholder in a district the information they need in which each type of assessment complements the other and sends a clear and consistent signal about the kinds of knowledge and skill the district values. Because there are so many stakeholders that need information about student learning for different purposes, and any given assessment is only good for a small set of purposes, a district assessment system can easily become out of balance and even bloated.

The work the Anytown Public Schools team did through this principled process addressed many challenges that can make a district assessment system feel incoherent and like an obstacle rather than a tool to improve learning and instruction.

## The Principled Process and the Challenges It Addressed

The three workshops were conducted using the District Assessment System Design Toolkit[[1]](#footnote-1), which was developed to address several challenges to developing and maintaining a district assessment system that serves students, their educators, and district policymakers with the appropriate balance. The specific challenges addressed by the toolkit are presented in the table below.

**Table 1. *Challenges Addressed in the Principled Design Process*.**

| **Challenge** | **Associated Need** |
| --- | --- |
| Assessment Literacy | Shared vocabulary/understanding of types, characteristics, and purposes of assessment. |
| Appropriate Use | Shared understanding of what assessment types/characteristics match intended purposes |
| Parsimony | A system that minimizes duplication (and thus time devoted to formal test taking) |
| Coherence | All components form a coherent whole to create a balanced picture of student learning |
| Power Sharing | A system in which educators share power and collaborate across all levels in the district |
| “Systemness” | An approach to developing an *assessment system* rather than a *collection of assessments* |
| Implementation | A plan for implementation that attends to potential hurdles |
| Maintenance | A plan for thoughtful revisions so that the **system** remains a **system**. |

## The District Assessment System Design Team

The district selected a team of 25 members to participate in the process, with broad representation of many roles in the district, as follows:

* 4 participants from the district central office
* 7 participants representing elementary school
* 7 participants representing middle school
* 7 participants representing high school
* 8 administrators
* 15 teachers
* 2 school counselors

The complete list of participants is provided in Appendix A.

## Results of the System Design Process

In the three workshops, the district team prioritized the purposes of the district assessment system, identified the types and characteristics of assessment that will be used to fulfill those purposes, and developed a timeline and plan for implementing the system they designed. The results are provided in three sections: a narrative vignette describing the system, a schematic representation of the system, and a high-level outline of the plan for moving from the existing system to the newly designed system.

### Design Team Recommendations

The district design team made the following recommendations:

* The district design team prioritized assessment FOR learning (or formative assessment) for informing instruction. This means that it will be necessary to invest in professional learning in the area of formative assessment.
* Formative assessment practices are intended to take place entirely under teacher control and is to occur in-the-moment in the classroom, before, during, and at the end of lessons. Thus, the formative assessment professional learning opportunities will need to specifically address very frequent formative assessment for the purpose of maintaining a constant feedback loop between teacher and student.
* In addition, teachers will be fully in control of evaluating mid-marking period achievement to inform short-term instructional grouping and next-unit planning by using end-of-unit interim assessments (OF learning) for that purpose. This also means that professional learning should be focused on the use of end-of-unit interim assessment data for effective and appropriate instructional grouping. The design team focused on this use of interim assessment data for large units (e.g., midterms or monthly tests/papers).
* The design team also prioritized using summative assessments covering either marking-period prerequisites or end-of-marking period assessments for coarser-grained instructional programming decisions such as grade or course placement.
* Similarly, the design team prioritized end-of-large-unit (e.g., midterm) and end-of-marking-period (e.g., end-of-course) assessments for awarding course credit to high school students without taking a course.
* The design team prioritized end-of-small-unit interim assessments, (e.g., weekly quizzes and homework) and end-of-large-unit interim assessments (e.g., midterm exams, papers, or project and monthly exams) along with final exams, papers, and projects for grading.
* The design team also prioritized screener exams to identify English learners eligible for program services, using the state EL screener exam
* The district also prioritized district-selected assessments for identifying students eligible for special education services, as well as for identifying students eligible for the honor roll and other district honors.
* The district also prioritized using district selected or developed assessments covering complete marking periods or complete courses of study (e.g., all of high school) for evaluating readiness for college and/or career training, whether students are on or off track for success at the next level, and for evaluating district programs and policies.
* Finally, the district determined to use state developed or state selected assessments for measuring student growth for both program and policy evaluation and for teacher evaluation.

Note that professional learning will likely be necessary for most of the system, as shown in the schematic below. Some professional learning resources may be available from the state.

# Workshop 1

In Workshop 1 the Design Team first reviewed key assessment vocabulary. The purpose of the review was to address the issue of a fractured vocabulary in the field of assessment, where the same term is used to mean different things by different people. Sometimes this is just a natural outgrowth of a developing field, and sometimes this is because a term has been used inappropriately to market assessment products. The definitions of key terms are provided in Appendix B.

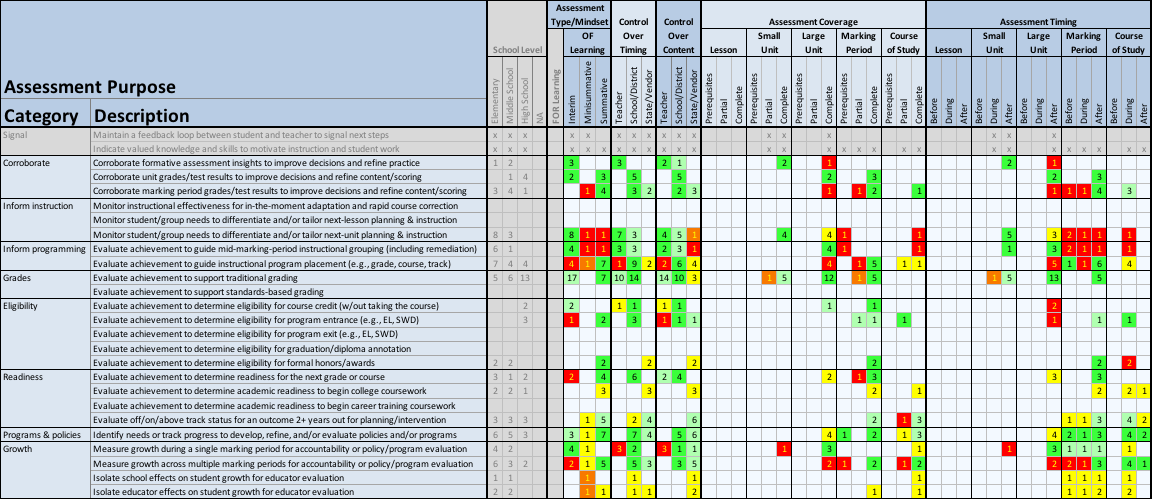
The team next reviewed a considerable list of purposes for assessment, which are provided in Appendix C.

To finish the first workshop, the team began performing an audit of the existing district assessment system using a Workbook that used the vocabulary and purposes reviewed in the workshop. This was completed between Workshop 1 and 2. The results of the audit are provided in Tables 3 and 4 below. The table was developed by evaluating the characteristics of the assessments entered into the audit against the purposes entered into the audit for the same assessment. The ratings of how well the characteristic of assessments matched the purposes listed is based on an expert rating of the degree of match by staff at the Center for Assessment.

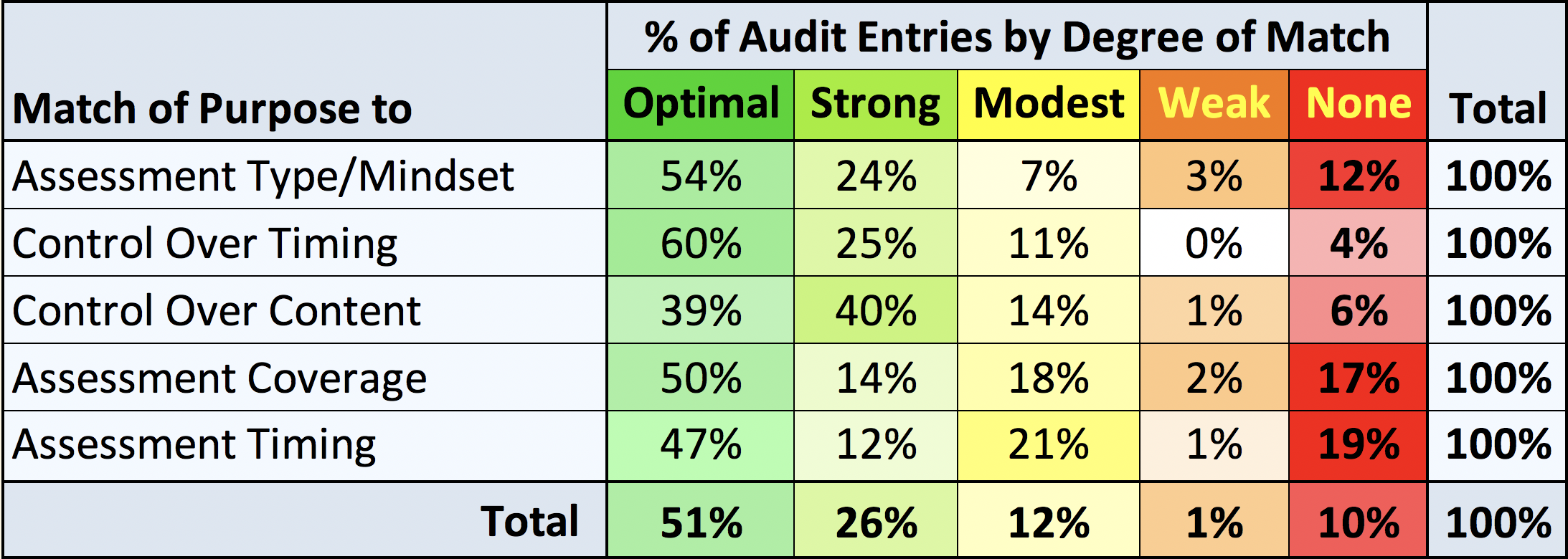
# Workshop 2

In Workshop 2, the district team reviewed the results of the district assessment audit, using the figures on the following page. In the first figure, the numbers represent how many entries in the audit there were that fell into each combination of purpose (the rows) and assessment characteristic (the columns). The colors represent how well the combination of purposes and characteristics were matched. As can be seen in both figures, most of the characteristics of assessment were at strongly to optimally matched to the purposes the assessments are used for, but there are some important exceptions. Particularly in the case of mindset (or type of assessment) and in what the assessments cover and when they occur, there are some characteristics that are not at all matched to the purposes for which the assessments are used. This indicates a need to review at least some existing assessments for modification or replacement.

### Detailed Results of the District Assessment Audit



### Summary Results of the District Assessment Audit

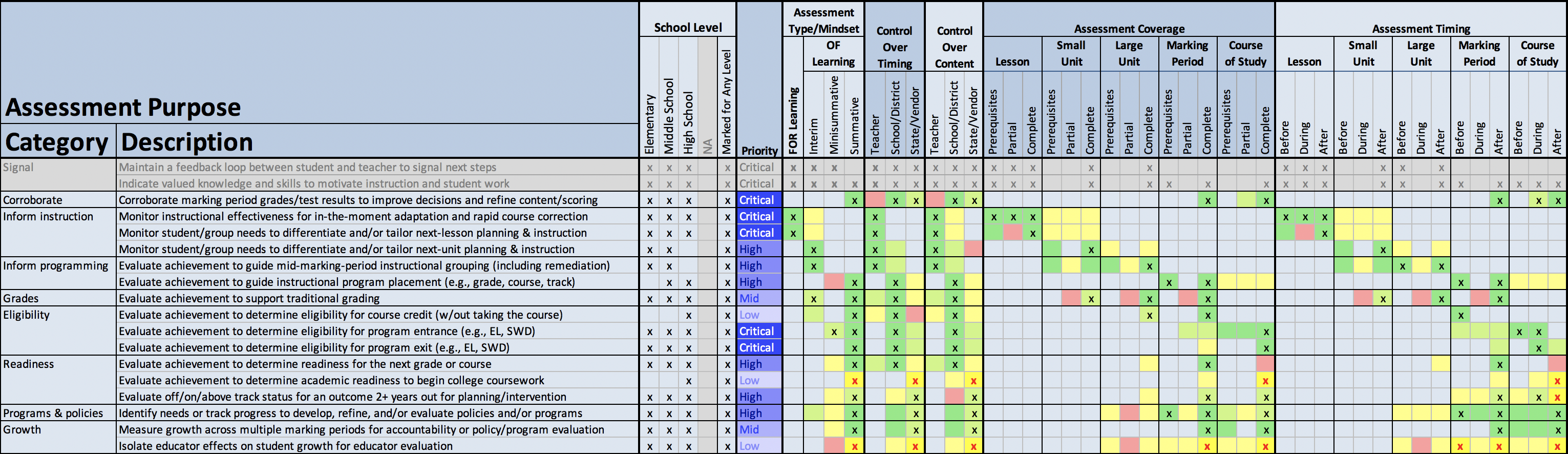


Following the review of the results of the district assessment audit, the Design Team selected a subset possible purposes of assessment to design a system that is simple enough to be feasible but comprehensive enough to serve the needs of the various stakeholders. From that subset deemed important enough to include, they then prioritized those purposes relative to each other on the scale *low*, *mid*, *high*, and *critical* (where the prioritization is relative to each other; the entire set was deemed important in the first step). Based on that ranking, the Project Workbook showed what mindsets/types and characteristics of assessment were best matched with each intended purpose, giving greater weight to *critical* priority purposes than *high* priority purposes, which were in turn given greater weight than *mid* priority purposes, and so on.

By starting with *critical* and *high* priority purposes, the team was able to select mindsets/types of assessment and characteristics of assessment that were ideally-matched or strongly-matched to the most important purposes as determined by the team. The team was then able to move on to *mid* and *low* priority purposes and were able to use some of the assessments already selected for higher-priority purposes if it was felt that a stronger match could make the system more complex and burdensome that was justified.

The results of that activity produced a design matrix giving a strong depiction of what the final recommended system would include. The design matrix is shown in the figure on the next page. That figure was interpreted by the External Facilitator to produce the narrative and graphical vignettes of the recommended system as given in the Executive Summary and on the following pages. The graphical vignette bears some explanation. A complete slide deck explaining how all of the components work together has also been created. It builds up the system piece by piece to show the roles of various stakeholders, the kinds of assessment that are intended to be used, and how they fit together. This schematic can be helpful in evaluating whether any new proposed components of the district assessment system will fit within a coherent system of assessments. It also gives a sound basis for periodically evaluating whether there are modifications needed and doing so in a principled manner.

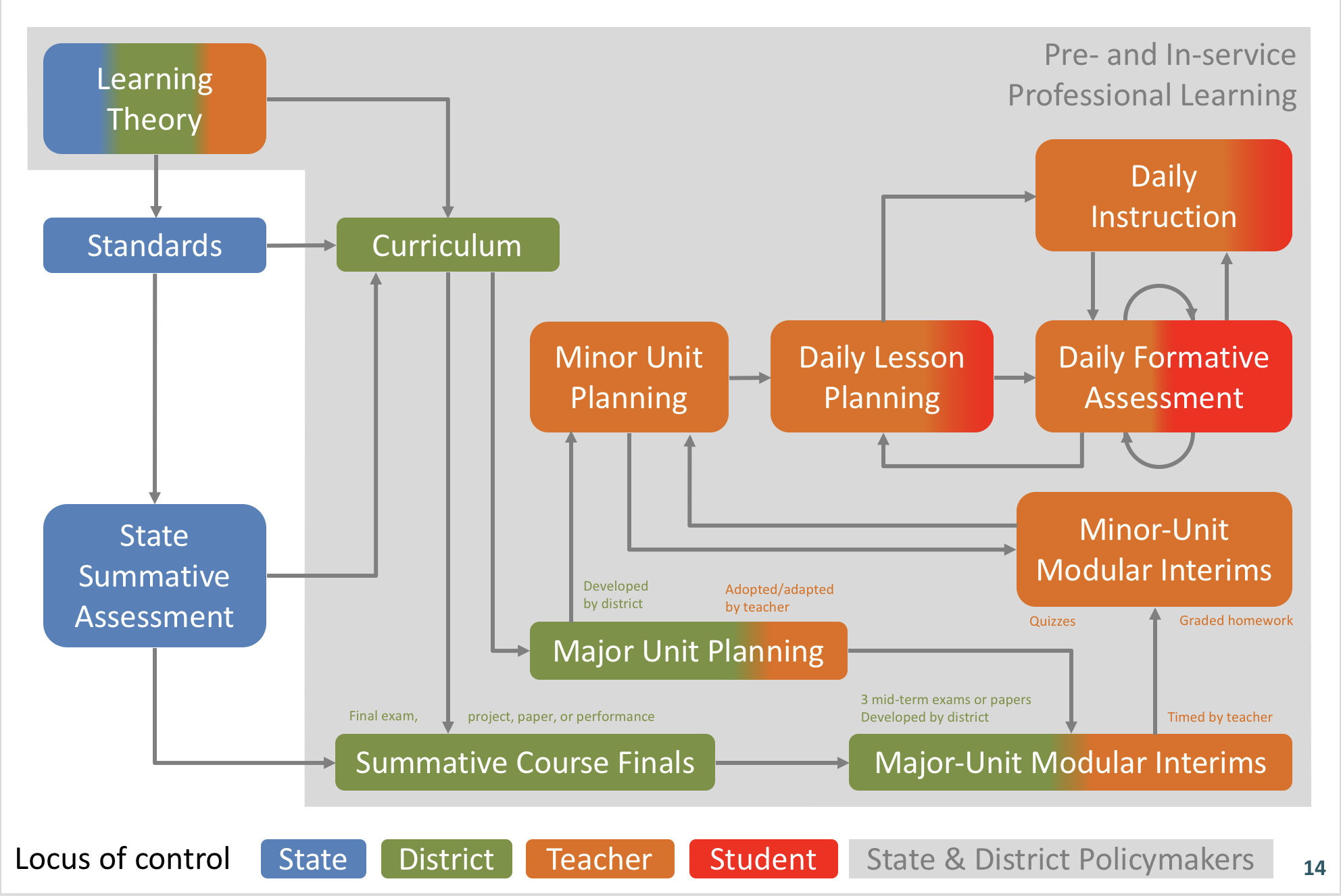
### Design Matrix Resulting from The Project.



In this figure, the following can be seen:

* The district design team prioritized assessment FOR learning (or formative assessment) for informing instruction. This means that it will be necessary to invest in professional learning in the area of formative assessment.
* Formative assessment practices are intended to take place entirely under teacher control and is to occur in-the-moment in the classroom, before, during, and at the end of lessons. Thus, the formative assessment professional learning opportunities will need to specifically address very frequent formative assessment for the purpose of maintaining a constant feedback loop between teacher and student.
* In addition, teachers will be fully in control of evaluating mid-marking period achievement to inform short-term instructional grouping and next-unit planning by using end-of-unit interim assessments (OF learning) for that purpose. This also means that professional learning should be focused on the use of end-of-unit interim assessment data for effective and appropriate instructional grouping. The design team focused on this use of interim assessment data for large units (e.g., midterms or monthly tests/papers).
* The design team also prioritized using summative assessments covering either marking-period prerequisites or end-of-marking period assessments for coarser-grained instructional programming decisions such as grade or course placement.
* Similarly, the design team prioritized end-of-large-unit (e.g., midterm) and end-of-marking-period (e.g., end-of-course) assessments for awarding course credit to high school students without taking a course.
* The design team prioritized end-of-small-unit interim assessments, (e.g., weekly quizzes and homework) and end-of-large-unit interim assessments (e.g., midterm exams, papers, or project and monthly exams) along with final exams, papers, and projects for grading.
* The design team also prioritized screener exams to identify English learners eligible for program services, using the state EL screener exam
* The district also prioritized district-selected assessments for identifying students eligible for special education services, as well as for identifying students eligible for the honor roll and other district honors.
* The district also prioritized using district selected or developed assessments covering complete marking periods or complete courses of study (e.g., all of high school) for evaluating readiness for college and/or career training, whether students are on or off track for success at the next level, and for evaluating district programs and policies.
* Finally, the district determined to use state developed or state selected assessments for measuring student growth for both program and policy evaluation and for teacher evaluation.

### Design Schematic Resulting from the Project.



This figure is helpful to show how the various components of the system are intended to work together and who will be responsible for what. It also makes clear that in order for the system to work as intended, considerable professional learning will be needed, and that that professional learning will need to include a clear and coherent conversation about whether and/or to what degree the learning theories upon which state, district, and teacher work is based are compatible with one another.

Note that it is more helpful to view this schematic in the context of the complete slide deck where the system is built up piece by pieces to show where each set comes from and how it is related to all other pieces that are already displayed.

# Workshop 3

In Workshop 3, the Design Team reviewed the narrative and graphical vignettes of the district assessment system described by the design matrix developed in Workshop 2. The Design Team reviewed both products and finalized them for inclusion in this report.

The Design Team also reviewed a comparison between the existing district assessment system as described by the assessment audit and the new design resulting from Workshop 2. The difference is shown in the figure on the following page.

The Design Team also began the process of identifying potential barriers to implementing the new system design, identifying potential strategies to address the potential barriers, and began the process of planning the next steps, which include the following:

* Developing a timeline for implementation
* Identifying who will be responsible for selecting and/or developing assessments that meet the design criteria from the project.
* Identifying the professional learning needed to implement the new system at the classroom, school, and district level in a way that is consistent with the narrative and graphical vignettes described in this report.
* Identifying who will be responsible for identifying and coordinating the needed professional learning activities.
* Developing a plan to periodically check in on progress toward implementation (and, when necessary, reviewing implementation for quality and consistency with the high-level design criteria identified in this project).

# Appendix A – Members of the District Assessment Design Team

All members of the District Assessment System Design Team are listed in the table shown below, demonstrating the diversity of roles represented by the team membership.

**< INSTRUCTIONS (TO BE DELETED)**

Groups that could be named in the right column include but are not limited to district leadership; central office staff; school administrators; kindergarten teachers, first grade-teachers, …, middle school teachers, high school teachers; special education teachers, Title I teachers, English learner teachers, career and technical teachers, mathematics teachers, English teachers, social studies teachers, science teachers, performing arts teachers, fine arts teachers, P.E. teachers, school counselors, school psychologists)

Make sure to include the District Liaison in the table >

| **Team Member Name** | **Team Member Title** | **Group(s) the Team Member Represented** |
| --- | --- | --- |
| Person A | Title A | Teacher, English, Grade 4 |
| Person B | Title B | Principal, Middle School |
| Person C | Title C | Central Office, Assessment |
| … | … |  |
| Person Y | Title Y | School Board Member |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

In addition to participating as a team member Rob Anybody also served as the District Liaison for this project. Finally, to ensure that all district team members could fully participate in the project, Barb Somebody (Independent Consultant) served as an external assessment expert to facilitate the work of the district team.

# Appendix B – Key Terms and Definitions

|  |  |  |  |
| --- | --- | --- | --- |
| Vocabulary Term | Definition | Example | Example |
| Control over Timing | Who (teacher, school, district, state, vendor) has the authority to specify when assessment will take place | A teacher plans her lessons for the week and makes an end-of-week quiz for the weeks’ worth of lessons. | A teacher plans her lessons to include mid-unit checks after each topic in the day’s lesson. |
| Control over Content | Who has the authority to control what content assessment covers | A teacher uses the district’s unit plan, but she spends a couple of additional days to ensure students are on track, and then administers the common end-of-unit assessment. | The district assigns each department to create common end-of-unit tests with flexibility to reorder units if appropriate and to complete them slower or faster depending on student needs. |
| Lesson | A single day (or no more than a few days of instruction covering a narrow set of content) |  |  |
| Small Unit | A unit of curriculum and associated instruction no larger than approximately three weeks | Five lessons | Two weeks |
| Large Unit | A unit of curriculum and associated instruction larger than three weeks and smaller than a marking period | Half a semester | A month |
| Marking Period | A unit of curriculum and associated instruction in which a student gets a final evaluation of performance | A semester | A course |
| Course of Study | A unite of curriculum and associated instruction larger than a marking period. | A group of courses culminating in an advanced placement course. | Grades K-12 |
| Assessment FOR Learning Mindset  (formative assessment) | A state of mind that approaches assessment as a tool to maintain a feedback loop between student and teacher. | A teacher plans several mid-lesson check-ins to receive feedback and provide feedback about learning and instruction to make rapid course corrections if needed. | A teacher provides several rounds of ungraded feedback on a unit writing assignment. |
| Assessment OF Learning Mindset | A state of mind that approaches assessment as a tool for evaluating student achievement. | Any student work that is graded | An orchestra teacher grades sight reading and practiced musical performance to assign chairs to students for the semester. |
| Formative Assessment  (assessment for learning) | Assessment FOR Learning | A teacher develops a small-group classwork project, shares the rubric with the class, observes each group based on the rubric, and asks each group to identify with her where their project falls on the rubric. | A teacher creates a mid-lesson quiz on addition with regrouping with half of the items requiring regrouping. She identifies students who got regrouping items wrong and the others right for small-group instruction during the second half of the lesson. |
| Summative Assessment (OF Learning) | A final measure of learning. | A final paper | A college admission exam. |
| Minisummative Assessment OF Learning) | An interim measure of learning that is based on the complete set of content of a marking period or course of study. | A commercial K-12 assessment that given three times per year in Reading and Mathematics | Grades K-12 |
| Large-Unit Interim  Assessment (OF Learning) | An interim measure of learning before the final measure, based on a relatively large subset of the content of an entire marking period | A midterm presentation. | An assessment addressing a month’s worth of units. |
| Small-Unit Interim Assessment (OF Learning) | An interim measure of learning before the final measure, based on a relatively small subset of the content of an entire marking period. | A weekly writing assignment | Tests administered every two weeks. |

1. Developed by the Center for Assessment (nciea.org), the Michigan Assessment Consortium (michiganassessmentconsortium.org), Wayne RESA (www.resa.net), and Oakland Schools (www.oakland.k12.mi.us). [↑](#footnote-ref-1)